
Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via http://linguistics.berkeley.edu/bls/.

The Annual Proceedings of the Berkeley Linguistics Society is published online via eLanguage, the Linguistic Society of America's digital publishing platform.
PROCEEDINGS OF THE TWENTY-FIRST ANNUAL MEETING
OF THE
BERKELEY LINGUISTICS SOCIETY
February 17-20, 1995

GENERAL SESSION
and
PARASESSION
ON
HISTORICAL ISSUES IN
SOCIOLINGUISTICS/
SOCIAL ISSUES IN HISTORICAL
LINGUISTICS

Berkeley Linguistics Society
Berkeley, California, USA
PROCEEDINGS OF THE TWENTY-FIRST ANNUAL MEETING
OF THE
BERKELEY LINGUISTICS SOCIETY
February 17-20, 1995

GENERAL SESSION
and
PARASESSION
ON
HISTORICAL ISSUES IN
SOCIOLINGUISTICS/
SOCIAL ISSUES IN HISTORICAL
LINGUISTICS

edited by
Jocelyn Ahlers
Leela Bilmes
Joshua S. Guenter
Barbara A. Kaiser
Ju Namkung

Berkeley Linguistics Society
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table of Contents</td>
<td>v</td>
</tr>
<tr>
<td>Table of Contents of BLS 21S</td>
<td>ix</td>
</tr>
<tr>
<td><strong>GENERAL SESSION</strong></td>
<td></td>
</tr>
<tr>
<td>Functional Verbs in Predicate Formation: Event-Type Hierarchy and</td>
<td>1</td>
</tr>
<tr>
<td>Grammaticization</td>
<td></td>
</tr>
<tr>
<td>MENGISTU AMBERBER</td>
<td></td>
</tr>
<tr>
<td>Spanish <em>Casi</em> as a Scalar Operator</td>
<td>12</td>
</tr>
<tr>
<td>RAUL ARANOVICH</td>
<td></td>
</tr>
<tr>
<td>Neurological Evidence for a Functional Basis for Lexical Categories</td>
<td>24</td>
</tr>
<tr>
<td>BARBARA BIRCH</td>
<td></td>
</tr>
<tr>
<td>Generic Demonstratives</td>
<td>32</td>
</tr>
<tr>
<td>BRIAN F. BOWDLE AND GREGORY WARD</td>
<td></td>
</tr>
<tr>
<td>On a Scalar Operator</td>
<td>44</td>
</tr>
<tr>
<td>YOON-SUK CHUNG</td>
<td></td>
</tr>
<tr>
<td><em>One</em>-Anaphora and Residual-DRS’s</td>
<td>60</td>
</tr>
<tr>
<td>PIROSKA CSMABI</td>
<td></td>
</tr>
<tr>
<td>The Interaction of the Binding Principles and the Chinese Reflexive</td>
<td>72</td>
</tr>
<tr>
<td><em>Taziji</em></td>
<td></td>
</tr>
<tr>
<td>JUN DA</td>
<td></td>
</tr>
<tr>
<td>Ethnolinguistic Loyalties among Barcelona’s Teens</td>
<td>83</td>
</tr>
<tr>
<td>HOPE N. DOYLE</td>
<td></td>
</tr>
<tr>
<td>The Lexical Representation of Light Verb Constructions</td>
<td>94</td>
</tr>
<tr>
<td>MARTIN EVERAERT AND BART HOLLEBRANDSE</td>
<td></td>
</tr>
<tr>
<td>Synchronic and Diachronic Typology: The Case of Ejective Voicing</td>
<td>105</td>
</tr>
<tr>
<td>PAUL D. FALLON</td>
<td></td>
</tr>
<tr>
<td>The Phonological Composition of Personal Pronouns: Implications for</td>
<td>117</td>
</tr>
<tr>
<td>Genetic Hypotheses</td>
<td></td>
</tr>
<tr>
<td>MATTHEW J. GORDON</td>
<td></td>
</tr>
<tr>
<td>Vowel Phonotactic Positions in Australian Aborigine Languages</td>
<td>129</td>
</tr>
<tr>
<td>PHILIP HAMILTON</td>
<td></td>
</tr>
<tr>
<td>The Function of F0-Peak Delay in Japanese</td>
<td>141</td>
</tr>
<tr>
<td>YOKO HASEGAWA AND KAZUE HATA</td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Conjunction as a Case Feature-Checker</td>
<td>351</td>
</tr>
<tr>
<td><strong>ED ZOERNER</strong></td>
<td></td>
</tr>
<tr>
<td><strong>PARASESSION</strong></td>
<td></td>
</tr>
<tr>
<td>Diachronic Aspects of Russianisms in Siberian Turkic</td>
<td>365</td>
</tr>
<tr>
<td><strong>GREGORY D. S. ANDERSON</strong></td>
<td></td>
</tr>
<tr>
<td>Pragmatic Markers in a Diachronic Perspective</td>
<td>377</td>
</tr>
<tr>
<td><strong>LAUREL J. BRINTON</strong></td>
<td></td>
</tr>
<tr>
<td>How Recent Contact Erased Ancient Traces in the Gender Systems of the Oromo Dialects</td>
<td>389</td>
</tr>
<tr>
<td><strong>ROBBIN CLAMONS</strong></td>
<td></td>
</tr>
<tr>
<td>Grammaticalization in AAVE</td>
<td>401</td>
</tr>
<tr>
<td><strong>PATRICIA CUKOR-AVILA AND GUY BAILEY</strong></td>
<td></td>
</tr>
<tr>
<td>The Death of “Prefixing”: Contact Induced Typological Change in Northern Australia</td>
<td>414</td>
</tr>
<tr>
<td><strong>IAN GREEN</strong></td>
<td></td>
</tr>
<tr>
<td>A New Hypothesis of the Origin of the Eastern Andalusian Vowel System</td>
<td>426</td>
</tr>
<tr>
<td><strong>JOSÉ IGNACIO HUALDE AND BENJAMIN P. SANDERS</strong></td>
<td></td>
</tr>
<tr>
<td>Yokuts as a Target Language in a Shift from Miwok</td>
<td>438</td>
</tr>
<tr>
<td><strong>MARVIN KRAMER</strong></td>
<td></td>
</tr>
<tr>
<td>Creole Studies and Historical Linguistics: Renewing our Vows</td>
<td>450</td>
</tr>
<tr>
<td><strong>JOHN MCWHORTER</strong></td>
<td></td>
</tr>
<tr>
<td>Social Issues in Historical Linguistics in Africa</td>
<td>465</td>
</tr>
<tr>
<td><strong>DEREK NURSE</strong></td>
<td></td>
</tr>
<tr>
<td>Internal vs. External Factors in Socio-Historical Explanations of Change: a Fruitless Dichotomy</td>
<td>478</td>
</tr>
<tr>
<td><strong>SUZANNE ROMAINE</strong></td>
<td></td>
</tr>
<tr>
<td>The Declension of Ethnonyms in English</td>
<td>491</td>
</tr>
<tr>
<td><strong>KEVIN TUITE</strong></td>
<td></td>
</tr>
<tr>
<td>Variation in Modern Dutch D-Weakening: A Historical Perspective</td>
<td>503</td>
</tr>
<tr>
<td><strong>NELLEKE VAN DEUSEN-SCHOLL</strong></td>
<td></td>
</tr>
</tbody>
</table>
The following papers were presented at the conference but do not appear in this volume:

Optimality Theoretic Syntax, Economy and the Lexicon
   JANE GRIMSHAW

Negative Concord and Logical Form
   ERIC JACKSON
# Table of Contents for Special Session to be Found in BLS 21S

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table of Contents</td>
<td>v</td>
</tr>
<tr>
<td>Table of Contents of BLS 21</td>
<td>vii</td>
</tr>
<tr>
<td>Preface</td>
<td>ix</td>
</tr>
<tr>
<td><strong>SPECIAL SESSION</strong></td>
<td></td>
</tr>
<tr>
<td>Giving Distance Its Due</td>
<td>1</td>
</tr>
<tr>
<td>A.L. BECKER</td>
<td></td>
</tr>
<tr>
<td>Dividing the Rice II: Reaching a Compromise in Northern Thai</td>
<td>16</td>
</tr>
<tr>
<td>JACK BILMES</td>
<td></td>
</tr>
<tr>
<td>The Grammaticalization of Thai ‘Come’ and ‘Go’</td>
<td>33</td>
</tr>
<tr>
<td>LEELA BILMES</td>
<td></td>
</tr>
<tr>
<td>Causal Relation and Information Sequencing in Mandarin</td>
<td>47</td>
</tr>
<tr>
<td>YUNG-O BIQ</td>
<td></td>
</tr>
<tr>
<td>Functions of Marked Aspect in Expository Discourse in Upper Tanudan</td>
<td>61</td>
</tr>
<tr>
<td>Kalinga (Philippines)</td>
<td></td>
</tr>
<tr>
<td>SHERRI BRAINARD</td>
<td></td>
</tr>
<tr>
<td>Two Pathways to Identifiability in Cerebon Javanese</td>
<td>72</td>
</tr>
<tr>
<td>MICHAEL C. EWING</td>
<td></td>
</tr>
<tr>
<td>Social Responsibility and Deixis in Tagalog Demonstratives</td>
<td>83</td>
</tr>
<tr>
<td>STEPHEN C. FINCKE</td>
<td></td>
</tr>
<tr>
<td>Creating the Middle Ground Register in Thai Conversation</td>
<td>95</td>
</tr>
<tr>
<td>SHOICHI IWASAKI AND PREEYA HORIE</td>
<td></td>
</tr>
<tr>
<td>Not yes, Not no: The Zen of Khmer Discourse Particles</td>
<td>107</td>
</tr>
<tr>
<td>ERIC SCHILLER</td>
<td></td>
</tr>
<tr>
<td>On the Degrees of Givenness: An Analysis of Noun Phrases in Some</td>
<td>114</td>
</tr>
<tr>
<td>Chinese Constructions</td>
<td></td>
</tr>
<tr>
<td>YILI SHI</td>
<td></td>
</tr>
<tr>
<td>Parallelism in Kayah Li: Elaborate Expressions and Beyond</td>
<td>127</td>
</tr>
<tr>
<td>DAVID SOLNIT</td>
<td></td>
</tr>
<tr>
<td>Local and Global Coreferences: Third-Person Zero Pronoun in Chinese</td>
<td>141</td>
</tr>
<tr>
<td>Narratives</td>
<td></td>
</tr>
<tr>
<td>MINGLANG ZHOU</td>
<td></td>
</tr>
</tbody>
</table>
The following papers were presented at the conference but do not appear in this volume:

Framing Achenese: Language Choice and Discourse Structures in Aceh
MARK DURIE
GENERAL SESSION
FUNCTIONAL VERBS IN PREDICATE FORMATION: EVENT-TYPE HIERARCHY AND GRAMMATICALIZATION

Mengistu Amberber
McGill University

1. INTRODUCTION. The purpose of this paper is to examine the syntactic and semantic status of a class of verbs which combine with other categories to form predicates. I will refer to such verbs as FUNCTIONAL verbs. I assume that a functional verb can be derived from its lexical variant by a process of GRAMMATICALIZATION. I use the term grammaticization in a narrow sense to refer to a synchronic process which makes some components of LEXICAL CONCEPTUAL STRUCTURE (LCS) oblique to the grammar.1

I will further assume that functional verbs encode relational content in the sense of Sapir (1921). Two types of relational content can be identified with respect to verbs. The first is an internal-relational content that encodes EVENT-TYPES not unlike 'aktonsarten' familiar from Vendler (1967) and Dowty (1979). In broad terms, these event-types include state, change-of-state, causation, and activity. The second one is an external-relational content that locates the position of the event relative to the time of speech.

Thus, a lexical verb typically consists of three types of information: (a) idiosyncratic-lexical information, (b) event-type information and (c) temporal information. The first and the second types of information are often morphologically fused. Temporal information, on the other hand, is typically encoded by inflectional morphology. However, in some constructions, the lexical and event-type content can be dissociated. For instance, in English, event-types can be realized either by bound elements, as in (1a), or by independent forms, as in

(1) a. -en 'inchoative/causative'; -ize 'causative'
   b. be 'state'; become 'inchoative'; make/cause/have 'causative'

In this paper I will focus on independent forms which encode event-types. In section 2, I present data from Amharic (Ethiosemitic) which employs the verb say to form a complex verb. I argue that there is an open MANNER component in the LCS of the verb say which must be spelled out syntactically. In section 3, I extend the proposed analysis to other cases of functional verbs; I examine the status of the verb take in Fon (Kwa) which functions as a closed-class member in a serial verb construction. In section 4, I discuss some consequences of the proposal.

2. THE QUOTATIVE VERB say

2.1. BASIC FACTS. In Amharic the verb እ猟 'say' occurs as a lexical verb just like the English quotative verb say:

(2) Lemma y1-hedal እ猟
    Lemma 3MS-go.imp.3MS say.pf.3MS
    Lemma said: "He will go".

The basic word order in Amharic is Subject-Object-Verb and the quotation clause occupies the object position. The Amharic quotative construction has only a direct quote structure in that the subject of the quote clause in (2) must be disjoint in reference from the original speaker.
Now the interesting fact is that a form which is the same as alə 'say' appears in structures such as (3):

(3) a. t'ormusu səbbər alə
glass-DEF break 'say'.pf.3MS
The bottle broke.
b. ləmma rot' alə
L. run 'say'.pf.3MS
Lemma ran.
c. ləmma təzən alə
Lemma sad 'say'.pf.3MS
Lemma became sad.
d. ləmma biʃəh alə-w
Lemma cold 'say' pf.3MS-1MO
Lemma became cold.
e. məʃət alə
night 'say'.pf.3MS
It became night.
f. ləmma zəm alə
Lemma quiet 'say'.pf.3MS
Lemma kept/became quiet.

Notice that the object position which is occupied by the quotation in (2), is occupied in (3) by a form which I will simply call a VERBAL NOUN (VN). Notice also that all grammatical features including tense and agreement are encoded by the say verb and not by the VN.

There is an obvious phonological similarity between the alə in (3) and the quotative verb alə in (2). However, whether there is an interesting structural relationship between the two verb forms needs to be determined.

I argue that alə in (3) is not simply homophonous with the quotative verb alə in (2) but is rather systematically related to it. One piece of evidence is that the alə of (3) has retained the morphological properties of the quotative form. The verb alə is historically derived from the tri-radical *bəʔələ. However, it has the curious property of exhibiting its older first radical /b/ only in the gerundive; this is true for both instances of alə as can be seen in (4) and (5) respectively:

(4) ləmma +-hədələ-hu biəo ...
Lemma 1MS-go.imp.-1MS saying ...
Lemma saying: "I will go"...

(5) t'ormus-u səbbər biəo ...
glass-DEF break 'saying'...
The glass broken ...

If the alə form in (3) is simply homophonous with the quotative verb it would be very difficult to account for the similarity in the gerundive, particularly since the phenomenon is not part of a regular phonotactic process of the language but rather an idiosyncratic fact that must be learned about this verb.

Secondly, it has been noted in the literature that the use of a verb meaning say as some sort of verbal formative, strange as it may seem, is in fact quite common in Nilotic and Cushitic languages (cf. Armbuster 1960). Again, this would be difficult to explain if the forms were accidentally homophonous with their respective lexical variants.

Third, a number of authors have observed that the construction may have evolved from an onomatopoeic source (see Armbuster 1960, Dawkins 1960). Synchronously, there are a number of constructions in which the form that co-
occurs with alə 'say' can be cognized as an imitation or reproduction of a physical sound emitted by an object. Examples are given in (6):

(6)  
a.  zınab-u  təb təb alə  
    rain-DEF  təb təb say.pf.3MS  
The rain dripped.

b.  bər-u  kəwə  kəwə  alə  
    door-DEF  kəwə  kəwə  say.pf.3MS  
The door knocked.

c.  hısan-u  ou  ou  alə  
    child-DEF  ou  ou  say.pf.3MS  
The child screamed.

Although the fact that the onomatopoetic item can appear with the verb alə 'say' is not perhaps surprising, as the verb is after all a quotative verb profiling vocalization, it is equally obvious that the items which obligatorily co-occur with alə in (3) are not onomatopoetic at all. Consequently, one needs to gain more insight into the nature of this verb.

Thus I assume that the two alə 'say' forms are polysemous. This assumption immediately raises one important question: how are the related senses encoded? In other words, how is it that a verb which primarily profiles vocalization can also encode abstract events?

Different approaches can be taken to address this issue. One way would be to list each use of the form in the lexicon. Another approach would be to list only one of the senses as basic and derive the related senses by some principle or algorithm. Still another approach would be to deny the discrete property of categories and argue for some kind of continuum upon which the various senses of the verb can be plotted, ranging from the most lexical to the most functional.

It is not my purpose here to compare the various ways of addressing this question. Rather I will suggest one possible direction which is consistent with some contemporary views of grammar. Within the context of generative grammar in general and the Principles and Parameters approach (cf. Chomsky and Lasnik 1991) in particular, the availability of multiple lexical entries cannot be the correct option because it lacks the explanatory power necessary for making generalizations (see also Pustejovsky 1993).

Thus I would like to argue for a theory which can account for the relationship between a lexical verb and its functional variants without assuming the availability of multiple lexical entries. My strategy is first to probe into the LCS of the lexical verb based on Jackendoff (1983), (1990), and then determine whether the same LCS can capture the functional use of the verb.

2.2. THE LCS OF say. What then is the LCS of the verb say? Jackendoff (1990) (following Gruber 1976) has suggested that the verb say has a Theme argument which belongs to the ontological category INFORMATION. The rationale behind this proposal is that when one says something to someone, what is said can be conceived of as an entity moving from the speaker to the receiver, that is as Theme. Beyond this however, it is not clear what the LCS of the verb say should be. In particular, Jackendoff (p.c.) observed that it is unclear whether say is or is not a conceptual primitive (SAY) analogous to GO, CAUSE, BE, etc.

However, implicit in the works of Jackendoff and Gruber is the idea that say is a two-place predicate, with Agent and Theme arguments. Notice that if the entity which would surface in the syntax as a quotation is given the status of Theme, then
the subject must be some kind of Agent. This follows from standard assumptions about the organization of events as articulated either in terms of the Thematic Hierarchy (see Grimshaw 1990, Jackendoff 1990 among others) or in terms of the Causal Chain of Croft (1991). In other words, if we take the localistic definition of Theme, as an entity which moves in a trajectory, the eventuality it participates in is that of a change in location. Any sub-event which comes before a change of location in a chain must have a causal function. This leads to the conclusion that the agent of say is an argument of CAUSE or an external instigator. Then translated into a Jackendovian notation, the LCS for say may be represented as in (7):

(7) say

\[ \text{[CAUSE ([Thing α] A, [ GO ([Info ] A, [FROM [Thing α ]])]})]\]

\[ \text{TO [Thing ]}<A> \]

According to (7), the verb say denotes an event in which the first argument is a thing, i.e. a causer and the second argument is an event. The embedded sub-event - designated by GO - has two arguments: the entity (i.e. Info(mation)) that moves and the trajectory it traverses or the Path (elaborated by the Source function From and the Goal function To). The A in sub-script stands for arguments which will be linked into syntactic positions. The end-point of the trajectory - the argument of TO - is optionally A-marked (<A>) capturing the fact that the goal argument of say is optional in the syntax.²

Despite its initial plausibility, there are non-trivial conceptual problems with the representation in (7). First, notice that if one accepts a force dynamic notion of causation as developed in Talmay (1985) where causation implies some degree of coercion, it is easy to see that the agent of say cannot be an external instigator. That is say cannot be in the same class with typical causative verbs such as kill, bring etc.

Second, it appears that the agent of say behaves like the single argument of the so-called unergative verbs (cf. Perlmutter 1978), or the INTERNALLY-CONTROLLED intransitive verbs in the sense of Smith (1970) and Levin and Rappaport (1992). According to Levin and Rappaport (1992), verbs like laugh, dance, tremble, denote eventualities that are internally-controlled, i.e., the eventuality can be controlled only by the person engaging in it. But verbs such as break, build, kill, denote eventualities that are under some sort of EXTERNAL CONTROL even when used intransitively.

This dichotomy has some genuine syntactic and semantic ramifications which have been amply demonstrated in Levin and Rappaport (1992). Consider (8) for instance: internally-controlled verbs typically lack a causative variant in transitivity alternation, whereas externally controlled verbs are not subject to this restriction:


b. The glass broke.  John broke the glass.

Verbs such as dance are at times referred to by the term MANNER OF MOTION. Jackendoff (1990:90) argues that Manner of Motion verbs cannot be considered as a variety of GO because they do not encode Path. Hence, for verbs such as dance, and also for verbs like laugh, sneeze, and cry, Jackendoff suggests the function MOVE as in (9):

(9) \[ \text{[MOVE ([Thing ]])} \]

Suppose that the event denoted by the verb say is internally controlled, and thus its Agent argument has the same status with the argument of verbs such as laugh, or dance. Notice that one crucial property which distinguishes the Manner of Motion
class of verbs from say is that the latter cannot occur on its own: John laughed is a complete sentence but *John said is not. Thus, here is a verb, the subject of which has the properties of internally-controlled intransitive verbs, and yet it must appear with another obligatory element (the quotation). Then the question is this: what is the syntactic and semantic status of the complement of say?

Determining the status of the quotative construction is a classical problem that has been addressed by a number of researchers over the years including Partee (1973) and Munro (1982), among others. For the sake of space I cannot go into the detail of the many interesting issues related to this question (but see Amberber 1995). I will instead propose a possible LCS for say which will account for its use as a functional form. Consider the dilemma once more: if we adopt a structure like (7), it may represent the non-agent argument, the entity which surfaces as the object, but goes against what we know about standard causation. On the other hand, if we adopt a representation such as (9) and claim that say is like any other internally controlled unergative verb, we account for the agentive status of the subject, but leave no room for the representation of the quotation whatever its thematic status turns out to be.

I would like to argue that say is best considered as a variety of GO. I will use the subscript SAY to indicate the particular semantic field within GO. The first argument of the event GO\textsc{say} is the sayer and its second argument is MANNER with an optional trajectory (Path). This can be schematized as in (10):

\[
(10) \quad \text{say} \quad \vee
\]

\[
\text{GO}_{\text{SAY}} ([\text{Thing}]A, [\text{MANNER}]A, [\text{Path}]<A>)
\]

I am attributing a semi-primitive status to SAY by identifying it as a semantic field of GO. The notion of a semantic field is independently motivated in Jackendoff's system to account for conceptual domains such as Possession (GO\textsc{poss}). What I would like to argue for is that the verb say is lexically unspecified for manner or more precisely for what I will call INNER MANNER. I assume that Inner Manner is part of the (idiosyncratic) lexical meaning of every verb.\(^3\)

Thus, suppose that say is a verb which lacks Inner Manner. In order to be interpreted, this component must be spelled out by another means, namely by the syntax. The function of the entity which must co-occur with the verb say is then to spell-out Inner Manner. This entity occupies a canonical object position in the syntax, but is different from other thematic arguments. While thematic arguments are true things which can have independent existence, the complement of say is an elaboration of the event itself. In fact, it is natural to assume that syntactically the quotation is some kind of cognate not unlike objects of unergative verbs\(^4\) such as dance as in dance a ceremony (see Austin 1982).

Now granted that the LCS of the lexical say is as represented in (10), how can we account for its functional variant? Suppose that there is a process available in universal grammar which suppresses parts of the LCS of a verb. Descriptively, this process can be referred to as BLEACHING, a common trend in grammatization through which some lexical items acquire grammatical character (see Hopper and Traugott 1993). Since, by hypothesis, grammatical items express topological notions (see Talmy 1987), when a lexical verb acquires a grammatical character it loses its specific reference. What this means with respect to say is that the specific GO\textsc{say} becomes the topological GO after grammatization. Once the verb say is underparsed (see note 1) as GO, then it would be free to encode event-types that are non-causative. Recall that this is exactly what we find in Amharic; in its functional
realization, the verb alɔ spells-out only non-causative events. For example consider (3a) repeated as (11) and the LCS of the functional alɔ presented in (12):

(11) tɔrmusu s^bbiri alɔ
    glass-DEF break 'say'.pf-3MS
    The bottle broke.

(12) a] alɔ

The first argument of GO will be realized by the entity which undergoes change of state (the bottle), whereas the manner component will be spelled-out by the entity which we call Verbal Noun (glossed as break in (11)).

It is interesting to note that the VN has one peculiar property: in tri-radical roots the form typically appears in two different templates, identified in the literature (see Beyene 1972) by the features INTENSIVE and ATTENUATIVE.

(13) a. CtCCxC [+Intensive] :=complete, :=instant
    b. CǝCǝC [+Attenuative] :=slight

In general, the plus-intensive template, as exemplified in break of (11) can have a range of related meanings - denoting that the eventuality comes about rather intensively, instantly, completely, or in a manner which has the property of all these. This feature clearly has a flavour of manner and its occurrence with the functional alɔ is quite natural as it is the spell-out of the Manner component. When the verb is used as a functional form, a new predicate is borne - a predicate which is composed of two morphologically independent items as in (14):

(14) [s^bbiri VN alɔ V]

The VN provides the idiosyncratic aspect of the verb meaning (including Inner Manner) whereas the verb alɔ remains constant spelling-out non-causative event-types.

Therefore, to summarize, the verb say which appears to be unusual as a functional element turns out to have the right properties to encode relational content. In the next section, I would like to show that certain other verbs, described in the literature as closed-class may be profitably analyzed as functional verbs. The case I will focus on involves a closed-class verb in a language with a serial verb construction.

3. CLOSED CLASS VERBS AS FUNCTIONAL VERBS. In the literature of serial verb languages a number of authors make a distinction between CLOSED and OPEN classes of verbs (see Sebba 1987). According to Lefebvre (1991) the verb take in Fon (a Kwa language of West Africa) belongs to a closed class. This verb co-occurs with verbs of an open class, as the example in (15) shows (from Lefebvre 1991):

(15) kɔkú só  àsɔ yɛ/wá ãxì
    Koku take crab go/come market
    Koku brought (direction away/towards the speaker) the crab to the market. (p.39)

This verb, of course, can also occur on its own in structures such as (16).

(16) kɔkú só ãkwé/ àsɔ / nu
    Koku take money/crab/something
    Koku took money/a crab/something. (p.55)
Lefebvre argues that the verb *take* has a causation component which requires an Agent and proposes the LCS represented in (17):

(17)  
\[ [x \text{ cause } [y \text{ undergo change of location}]] \] / TAKE (p. 55)

According to Lefebvre the open class verbs which co-occur with *take* in the serial structure are motion verbs such as *go, come,* or verbs such as *give, teach/study, show/learn.* The question is then whether the open class verbs can be defined in a unitary way. Lefebvre argues that all the open class verbs which can appear with *take* can be described by the LCS in (18):

(18)  
\[ [y \text{ undergo a change of location in a direction away from/towards the speaker to (a location) } z] \] (p.56)

Lefebvre further argues that the operation of serialization is defined as a process which involves the association of the LCS of two verbs. Note that both verbs come together with their full LCS, as shown in (19a) and (19b) respectively:

(19)  
a. *Take*  
\[ [x \text{ cause } [y \text{ undergo change of location}]] \]

b. *Go*  
\[ [y \text{ undergo a change of location to } z]] \] (p.59)

The problem with the association of these two verbs is that there will be two Theme arguments that need to be mapped onto syntax but there is only one position available in the syntax. Thus one Theme must be blocked. Crucially for Lefebvre, among the two *y* variables or Themes, the one which must not be realized in the syntax is that of the *take* verb. Thus, the following properties of the construction must be accounted for:

(20)  
a. Two *y* variables referring to a single Theme of location;

b. Only one argument represented by a *y* variable is realized;

c. Embedding of the LCS of the verb of the open class within that of the *Take* verb. (p.60).

It is also suggested that the construction involves a substitution operation which is stated in (21):

(21)  
Serialization is a substitution process which substitutes "a LCS of type [y undergo change of location to z] for the [y undergo change of location] of the *Take*-verb." (ibid).

To the extent that Lefebvre's analysis of the serial sequence as involving closed-class and open-class with the properties presented above is correct, the analysis which stipulates an asymmetrical relationship in the substitution operation such that the internal argument of the derived LCS must be that of the open class verb and *not* that of the *Take*-verb is particularly problematic; that is why substitution should apply in this asymmetrical fashion has been left unmotivated.

But suppose that *take* which participates in the serial construction is actually a functional verb which spells out a single event-type, i.e. CAUSE. I propose that this can be achieved by assuming some kind of Transparency Condition along the lines of (22):

(22)  
If a lexical verb denotes a complex event, only its highest sub-event will be transparent to syntax when used as a functional verb.5

For the present purpose, I assume the Event-type hierarchy in (23), in which the highest event-type is CAUSE.

(23)  
*Event-type hierarchy*

Causation > Change of state/Activity/State

Note that within our proposal none of the properties listed in (20) will be assumed as the closed class *take* does not have the full LCS attributed to its polysemous lexical variant. In other words, when the verb *take* is underparsed, only the highlighted portion of (17'), will be visible to the syntax.6
(17') \[ x \text{ cause } y \text{ undergo change of location}] / \text{TAKE}

As a result, there are no two \( y \) variables corresponding to one internal argument. The stipulative nature of the substitution operation also disappears as there is no need for embedding the LCS. The term \textit{closed} class then makes sense as a descriptive term for the functional \textit{take} as it has more or less the same status as morphological causative markers in non-serial verb languages.

4. CONSEQUENCES. Before concluding I would like to discuss some consequences of the proposal. Although the present study has focused on verbs, it would be desirable if the basic insight can be extended to other types of predicates. I believe that the proposal can be extended to other predicates particularly to adpositions.

It is well know that in some languages verbs are grammaticized to function as (directional) adpositions. This can be accounted for quite naturally within our proposal once the Transparency Condition is elaborated along the lines in (24):

(24) Only the Path argument of a verb will be transparent to syntax when the verb is used as an adposition.

In other words, the category Adposition is \textit{blind} to any other argument in the LCS except the Path argument. The minimal requirement would then be that verbs which lack Path in their LCS cannot function as adpositions. A cursory examination of the relevant literature provides evidence consistent with this claim. According to Lord (1993:10) in a number of languages the most common verbs which can have prepositional meaning are as shown in (25):

(25) \begin{align*}
\text{Verb} & \quad \text{Meaning} & \quad \text{Prepositional Meaning} \\
go & \quad \text{to, into,} & \quad \text{towards} \\
come & \quad \text{to} & \quad \text{through} \\
fall & \quad \text{to} & \quad \text{through} \\
\text{pass through} & & \\
\end{align*}

Consider, for instance, the LCS of a verb meaning \textit{go} represented in (26):

(26) \[ \text{GO}([\text{Thing }], [\text{Path TO }]) \]

When this verb is underparsed as an adposition, only the Path argument will be transparent to syntax. Thus the verb effectively acquires an adpositional LCS to be realized as \textit{to} or its equivalent.

For a more complicated case, consider verbs such as \textit{give} which can be employed as the adposition \textit{to} or the benefactive \textit{for} in several languages (Lord 1990). Suppose the LCS of \textit{give} is as in (27), (adopted from Jackendoff 1990:194):

(27) \[ \text{give} \]

\[ \text{V} \]

\[ [\text{CAUSE ([ ] } \alpha \text{A }, [\text{GPOSS ([ ]A, [ FROM [ [ ]A])])} ] \text{ Path TO [ ]A} \]

The Path function has two arguments \textit{FROM} and \textit{TO} which are essential components in the meaning of the verb. This is to capture the fact that the object of the \textit{giving} event - the Theme - originates from the \textit{giver} (i.e., the causer) and not from elsewhere. This is shown by co-indexation of the arguments of \textit{CAUSE} and \textit{FROM} in (27). When the verb is underparsed as an adposition, only Path becomes visible to syntax. Notice, however, that there are two possibilities in Path: \textit{FROM} and \textit{TO}, which may give rise to adpositions with the Source meaning \textit{from} or the Goal meaning \textit{to}, abstracting away from the multiple roles of \textit{TO} as Goal and Beneficiary.
I argue that the Source adposition cannot be derived, because FROM is not A-marked in the verb form as it is co-indexed with the CAUSE argument. That is, underparsing is relevant only with respect to A-marked arguments. This would predict that a verb meaning give can function as Goal adposition but never as Source adposition, an empirical claim worth investigating.

5. CONCLUSION. In this paper I attempted to show that functional verbs are systematically related to their respective lexical variants. Even uncommon instances of functional verbs such as say turn out to behave naturally once the correct LCS is motivated. Needless to say, the theory of grammar must be able to account for the role of functional verbs - whatever they may be called in different descriptions and theories, light verbs, vector verbs, co-verbs, etc. - because it is to be expected that some linguistic forms will have multiple functions since language must employ its finite means economically to represent infinite expressions.

NOTES

* This research was supported by SSHRCC grant # 410-93-0897 to Lisa Travis. I would like to thank Lisa Travis, Mark Baker, Mark Durie, Amanda Miller and O.T. Stewart for the valuable suggestions they made on an earlier version of this paper. I also thank members of the audience at BLS Twenty-First, particularly Sarah Taub and Rebecca Wheeler, for their insightful comments. All remaining errors are of course mine. In the transcription, in general IPA symbols are used. Ejective consonants are indicated by the symbol [ʼ]. The abbreviations used are DEF, definite; imp, imperfect; pf, perfect; S, subject; O, object; ACC, accusative. The agreement morphemes indicate person and gender (1/3, M, F).
1. Grammaticization as understood here is close to Grimshaw's (1995) idea of un (der) parsing of the LCS; Grimshaw argues that this is a generally available UG process which languages are free to make use of. In Grimshaw's view, the English light verb do has unparsed LCS, whereas the heavy verb do has parsed LCS. Although I cannot address the issue here, I think unparsing (as opposed to underparsing) is too strong even for the case of the English do . In section 2, I suggest one possible way which will constrain the process of underparsing.
2. Note that this notation is consistent with the rather minimal difference between say and the so-called information transaction verbs such as tell in that with the latter the Goal argument is obligatorily A-marked and must be realized syntactically.
3. The notion of Inner Manner may be similar to the notion of natural actions discussed in Jackendoff (1990:88) following Peterson (1985); Jackendoff points out that 'natural actions are difficult to describe in words' but are relatively easy to demonstrate. Thus, for instance, the distinction between wiggle and dance cannot be adequately captured by features in Conceptual Structure but rather must be encoded in some other modality. I would argue that the distinction between wiggle and dance is due to Inner Manner which must be learned as part of the idiosyncratic aspect of the verbs' meaning. It is also conceivable that Inner Manner may be analogous to Higginbotham's (1985) 'event' argument.
4. In fact, in some languages certain unergative verbs cannot occur without cognate objects (see Austin 1982, Hale and Keyser 1991). This means that unergative verbs are syntactically transitive, though they are single argument verbs in Conceptual Structure.
5. Why the highest event-type is privileged in the sense postulated here is of course something which must be motivated. I will not pursue this issue here, though, I
would like to point out that for independent reasons a similar asymmetry is present in a number of other studies as well. For instance, in the theory of Argument Transfer, developed in Grimshaw and Mester (1988) higher thematic roles can transfer without lower ones but not vice versa. 6. From this it follows that any functional verb denotes only a single event-type. When the lexical source denotes a single event then that will be retained as the event-type of the functional variant as we have seen in the case of alə 'say', though the specific semantic field is unparsed.

REFERENCES


Davidson, Donald. 1968/69. On saying 'that'. Synthese 19. 130-146.


Spanish *Casi* as a Scalar Operator

RAUL ARANOVICEH
UNIVERSITY OF CALIFORNIA, SAN DIEGO

1. Distribution of *casi*

The Spanish adverb *casi* ‘almost’ can be used as a sentential modifier, as in (1a), as a modifier of adjectives, as in (1b), or as a modifier of quantifiers, as in (1c):

(1a) Andy casi tradujo el poema.
    ‘Andy almost translated the poem.’

(1b) Andy está casi sordo.
    ‘Andy is almost deaf.’

(1c) Andy entiende casi todo.
    ‘Andy understands almost everything.’

The distribution of *casi* as a sentential modifier depends on a combination of the aktionsart of the verb and the polarity of the sentence. *Casi* is acceptable with an accomplishment verb in an affirmative sentence, as in (1a), but not in an affirmative sentence with an activity verb, as shown in (2a). If the polarity of (2a) is flipped, however, as in (2b), *casi* becomes acceptable.

(2a) * Andy casi respira.
    ‘Andy almost breathes.’

(2b) Andy casi no respira.
    ‘Andy almost does not breathe.’

In this paper I will argue that *casi* is a scalar operator, and that it selects a sentence that denotes the origin of a scale to yield a sentence that denotes the closest point on the scale. I will also show that the polarity of the sentence and its aktionsart determine whether a sentence can be associated to the origin of a scale. This will account for the data in (2). Finally, I will show that some well-known facts about the meaning of *casi* also follow from an analysis in terms of scales.

2. Two senses of *casi*

As an aspectual modifier, *casi* is ambiguous between an ‘outcome’ interpretation and a ‘phase’ interpretation (Binnick (1991) for English *almost*). The outcome interpretation is related to the telicity of the verb, whereas the phase interpretation is related to the process. Accomplishments, like the sentence in (1a), are ambiguous
between the two interpretations: (1a) can be translated as (3a), which gives the outcome reading, or as (3b), which gives the phase reading.

(3a) ‘Andy was going to translate the poem, but she didn’t.’
(3b) ‘Andy is almost done with the translation of the poem.’

Achievement verbs, on the other hand, which are telic but do not refer to a process, can take casi only under an outcome reading. The sentence in (4a) can only have the meaning in (4b):

(4a) Andy casi ganó la lotería.
‘Andy already almost won the lottery.’
(4b) ‘the event of Andy winning could have happened.’

Casi can be disambiguated by paraphrasing it as ya casi, literally: ‘almost already’, which can only take a phase interpretation. Casi cannot be paraphrased as ya casi with achievements, as (5b) shows. The paraphrase is possible with accomplishments, however, as (5a) shows, since accomplishments admit of a phase reading.

(5a) Andy ya casi tradujo el poema.
‘Andy already almost translated the poem.’
(5b) * Andy ya casi ganó la lotería.
‘Andy already almost won the lottery.’

Casi can also appear with some activities, as in (2b), repeated under (6a). The ya casi test shows that this occurrence of casi is related to the phase reading of the adverb. Casi can be paraphrased as ya casi yielding (6b).

(6a) Andy casi no respira.
‘Andy almost does not breathe.’
(6b) Andy ya casi no respira.
‘Andy already almost does not breathe.’

Both senses of casi are sensitive to negation and to the aktionsart of the verb, with some differences. I will discuss the two reading of casi separately.

3. The phase reading of casi

3.1. Where can phase casi appear?

In affirmative environments, phase casi is acceptable only with accomplishment verbs. (7c), which is an accomplishment, can take casi. (7a), (7b), and (7d)
which are an activity, an achievement, and a state, respectively, cannot take *casi* in its phase reading.

(7a)   * Andy casi respira.  
       ‘Andy almost breathes.’

(7b)   * Andy casi ganó la lotería.  
       ‘Andy almost won the lottery.’

(7c)   Andy casi tradujo un poema.  
       ‘Andy almost translated a poem.’

(7d)   * Andy casi tiene un auto.  
       ‘Andy almost has a car.’

With negation, accomplishment verbs can no longer take *casi*, as in (8c), but activities can, as in (8a). Achievements and states, on the other hand, do not improve under negation, as seen in (8b) and (8d), respectively.

(8a)   Andy casi no respira.  
       ‘Andy almost does not breathe.’

(8b)   * Andy casi no ganó la lotería.  
       ‘Andy almost did not win the lottery.’

(8c)   * Andy casi no tradujo un poema.  
       ‘Andy almost did not translate a poem.’

(8d)   * Andy casi no tiene un auto.  
       ‘Andy almost does not have a car.’

When the data in (7) and (8) are set against the results of standard analyses of the internal structure of states and eventualities, like the ones in Vendler (1967) and Mourelatos (1981), a generalization arises: *casi*, in its phase reading, can only apply to a sentence denoting the final stage of a process. In the next section I will give a more precise version of the generalization regarding the distribution of phase *casi* using scalar models.

3.2. Modelling the generalization

An accomplishment can be modelled as a sequence of moments with a culmination or ‘outcome’. Accomplishments are non-homogeneous in their internal structure, since all the moments in the process can be distinguished from one another by the distance that separates them from the outcome. An affirmative sentence like (9a) denotes the outcome, whereas a negative sentence like (9b) denotes any moment previous to the outcome.
(9a) Andy tradujo el poema.
    'Andy translated the poem.'

(9b) Andy no tradujo el poema.
    'Andy did not translate the poem.'

The moments that make up an accomplishment like 'translating a poem' can be ordered according to a measure of completion of the action, which, for the sake of argument, I will assume is the number of words that remain to be translated. The outcome of the process is reached when there are no more words to translate. This model can be represented in a scale in which the outcome corresponds to the zero point or point of origin, and all the other moments are in the body of the scale. This is shown in (10).

(10) \[
\begin{array}{c}
{\ldots -P \ldots} \\
{M_n \ldots \quad M_2 \quad M_1 \quad \ldots \quad M_0}
\end{array}
\]

I have drawn a dotted line between $M_1$ and $M_0$ because the entailment relations between the propositions denoting the moments on the solid line are different from the entailment that connects them to the outcome. If there are two words that remain to be translated, for instance, then it can be said that there is one word that remains to be translated. But if there is one word that remains to be translated, it cannot be said that there are zero words remaining. This is why the outcome is denoted by a proposition $P$, which is what sentence (9a) expresses, whereas any moment previous to the outcome is denoted by its polar opposite -$P$, which is what (9b) expresses. I will refer to scales of this sort, i.e. scales that start at the zero point, as zero-on scales. In this I am departing from the classical definition of a scalar model (Fauconnier 1975, 1978, Fillmore, Kay and O' Connor 1988, Kay 1990), which is based solely on the implication relations that connect a set of propositions. In these classical scalar models, the zero of the scale is left out, i.e. the point of origin is equal to 1.

The difference between (9a) and (9b), then, is that only (9a) denotes the origin of the scale that models the accomplishment. The main claim of this paper is that this is the factor that decides the distribution of casi: casi will not combine with (9b), only with (9a). The generalization about the distribution of casi can then be made explicit with the notion of a zero-on scale, as follows:

**The Casi Condition.** Casi can only apply to sentences that denote the origin of a scale.
3.3. Testing the *Casi* Condition

I will show next how this analysis explains the remaining data in (7) and (8). Achievements, being instantaneous, lack all internal structure, and cannot be mapped onto a scalar model. So, sentences like (11a) and (11b) can never take *casi*.

(11a) Andy ganó la lotería.
    ‘Andy won the lottery.’

(11b) Andy no ganó la lotería.
    ‘Andy did not win the lottery.’

Activities have a homogeneous structure, but unlike accomplishments, they have no culmination, no outcome. The moments that make up the activity ($M_a ... M_e$), then, cannot be arranged on a scale. A representation of an activity would be as in (13), where $P$ is (12a) and $-P$ is (12b).

(12a) Andy respira.
     ‘Andy breathes.’

(12b) Andy no respira.
     ‘Andy does not breathe.’

(13)
\[
\begin{array}{cccccc}
\text{-P} & \{ & \ldots & P & \ldots & \} & \text{-P} \\
M_a & M_b & M_c & M_d & M_e
\end{array}
\]

So, (12a) does not denote the origin of a scale and, as predicted, it cannot combine with *casi*.

The schema in (13) suggests that (12b) does not denote the end of a scale either. The fact that *casi* does combine with (12b), then, seems to be a counterexample for the *Casi* Condition. But activities can be mapped on a scalar model according to the intensity ($I_i$) with which they are performed. (12b) denotes the point at which the intensity of breathing is zero, whereas (12a) denotes any intensity of breathing. The representation that results from this is as in (14)

(14)
\[
\begin{array}{c}
I_n \\
I_2 \\
I_1 \\
I_0 \\
\text{---} & \text{---} & \text{---} & \text{---} & \text{---}
\end{array}
\]
In this model, (12b) does denote the origin of a scale, and therefore can be combined with casi. (12a), on the other hand, can only be mapped onto the body of the scale, and not on the origin, since the absence of activity does not have any discernible internal structure in terms of intensity. This mapping, then, will not license the use of casi either.

An interesting situation arises with states. States, like achievements, do not have any internal structure in terms of moments or intensity. Therefore, neither (15a) nor (15b) can be mapped onto the origin of a scale, and none of them can be combined with casi, as I showed in sentences (7d) and (8d).

(15a) Andy tiene un auto.
     ‘Andy has a car.’

(15b) Andy no tiene un auto.
     ‘Andy does not have a car.’

But some arguments may induce a measure of the state that makes it suitable to be mapped onto a scalar model. This is what happens with (16a) and (16b):

(16a) Andy tiene pelo.
     ‘Andy has hair.’

(16b) Andy no tiene pelo.
     ‘Andy does not have hair.’

These sentences can be represented in a scalar model that takes into account amounts \( (A_i) \) of hair. This is shown in the schema in (17), which is identical to the schema in (13). The affirmative sentence (16a) expresses a proposition \( P \) that denotes any point in the body of the scale, whereas the negative sentence (16b) denotes a proposition \(-P\) that denotes the origin of the scale, the point of zero hair.

(17)

\[
\begin{array}{c}
A_n \\
A_2 \\
A_1 \\
A_0
\end{array}
\begin{array}{c}
\leftarrow \\
\rightarrow P \\
\rightarrow \neg \\
\downarrow \neg P
\end{array}
\]

As predicted, casi cannot combine with (16a), as (18a) shows, but it does combine with (16b), as in the example in (18b).
(18a) * Andy casi tiene pelo.
   'Andy almost has hair.'

(18b) Andy casi no tiene pelo.
   'Andy almost does not have hair.'

To summarize, casi in its phase reading can modify affirmative sentences with accomplishment verbs, and negative sentences with activities or measurable states. I have shown that all these sentences express propositions that denote the origin (the zero point) of a scale associated with the measure of completion, intensity, or degree of an eventuality or state. The sentences that cannot take casi either denote the body of a scale, or cannot be mapped onto a scale at all. In this way, the complex interaction of aktionsart and polarity with casi is reduced to a very simple principle, expressed in the Casi Condition.

4. Casi in its outcome reading

4.1. Where can outcome casi appear?

The distribution of casi in its outcome reading is less puzzling than the distribution of phase casi. Only telic verbs can take casi in its outcome reading. This is illustrated by the sentences in (19). Neither activities (19a) nor states (19d) can take casi under this reading. Only achievements (19b) and accomplishments (19c) can combine with casi.

(19a) * Andy casi respira.
     'Andy almost breathes.'

(19b) Andy casi ganó la lotería.
     'Andy almost won the lottery.'

(19c) Andy casi tradujo un poema.
     'Andy almost translated a poem.'

(19d) * Andy casi tiene un auto.
     'Andy almost has a car.'

Unlike the phase reading of casi, the outcome reading does not seem to be influenced by negation. When the sentences in (19) are negated, the same generalization holds: only telic verbs can be modified by casi. So, (20a), with an activity verb, is ungrammatical, and so is (20d), which has a state verb. An achievement like (20b) and an accomplishment like (20c), on the other hand, can have casi as a modifier.

(20a) * Andy casi no respira.
     'Andy almost does not breathe.'
(20b) Andy casi no ganó la lotería.
    ‘Andy almost did not win the lottery.’

(20c) Andy casi no tradujo un poema.
    ‘Andy almost did not translate a poem.’

(20d) * Andy casi no tiene un auto.
    ‘Andy almost does not have a car.’

4.2. Extending the Casi Condition

The generalization for the outcome reading of *casi* is also driven by considerations of scalar values. Telic verbs, being bound events, bring about changes in the world. These changes can be associated with a probability value, since some changes are more likely to happen in one or another way. This value projects a scale. A sentence like (21a), for instance, which asserts that an event happened, also asserts that the probability of the event not happening was zero. Therefore, it can be mapped on the origin of a scale, as in the diagram in (22). The body of the scale, designated by (21b), contains the cases in which the probability of the event not happening was more than zero. At an intuitive level one can think of the points in the body of the scale as indicating how many things should have been different to produce the change.

(21a) Andy ganó la lotería.
    ‘Andy won the lottery.’

(21b) Andy no ganó la lotería.
    ‘Andy did not win the lottery.’

(22)

Probability scales like the one in (22) are ‘reversible’: since sentence (21b) indicates that the probability of the event happening was zero, it can also be mapped onto the origin of a scale, as in the schema in (23).
The body of the scale in (23) contains the cases in which the probability of the event happening was more than zero. At an intuitive level too, one can think of these numbers as indicating how many things should have been different to avoid the change.

Affirmative sentences with achievement verbs, as well as negative ones, then, denote the origin of a scale, and both can combine with casi in its outcome reading. The same is true for accomplishments. States and activities, on the other hand, cannot be mapped onto scales that take into account a measure of probability, since they do not represent changes in a situation. Neither states nor activities can be modified by casi in its outcome reading. The conclusion, then, is similar to the one for the distribution of the phase reading: only sentences denoting the origin of a scale can take outcome casi. The Casi Condition can therefore be generalized to the outcome reading of casi. This result has some interesting consequences regarding the ambiguity of casi, to which I will turn next.

5. The meaning of casi

5.1. Truth conditions for casi

The Casi Condition can be used to account for the distribution of casi because it is possible to show that different scales can be associated with various sentences depending on their polarity and their aktionsart. A consequence of this approach is that the ambiguity in the meaning of casi vanishes: Casi can be assigned a constant meaning because the contrast between the phase reading and the outcome reading is due to the association of different scales with a given sentence. Generally speaking, a sentence of the form ‘casi S’ designates the point on the scale closest to the origin, which is designated by ‘S’ itself, i.e. by the ‘argument’ of casi. The meaning of casi, then, is a function that goes from the origin of a scale to the next notch up.

I will suggest one way to formalize this analysis. As a functor that maps sentences onto sentences, the contribution of casi to the sentence it appears in is like that of a modal operator. The meaning of casi can thus be defined in possible-world semantics. Scales of the sort I have been considering in this paper impose a structure on a set of worlds, based on accessibility relations. In the case of accomplishments like ‘translating a poem’, for instance, each moment in the process corresponds to a possible world W_i (i.e. each moment defines a state of affairs), and only the worlds that are adjacent to each other in a scale like the one in (10) are ‘accessible’ to each
other. Given a structure like that, the question is which relation holds between the world in which ‘casi S’ is true, and the world in which ‘S’ (the argument of casi), is true. The answer is that ‘casi S’ is true in the first accessible world from the world in which ‘S’ is true. Assuming that sentences are functions from worlds (indices) to truth-values, the truth conditions for casi can then be defined as in (24):

\[
\text{casi } S^{W_i} = \begin{cases} 
T & \text{iff } S^{W_{i-1}} = T \\
F & \text{otherwise.}
\end{cases}
\]

An analysis of casi as a scalar operator, then, accounts for certain generalizations about its distribution and its meaning. An analysis in terms of scales uncovers a very simple condition behind a complex pattern created by the interaction of the polarity of the sentence and its aktionsart. By making reference to zero-on scales, a general feature of the meaning of casi that was hidden under a superficial ambiguity can be made explicit. All that is needed is to enrich the model of the language with zero-on scales, the existence of which is probably justified by enough independent evidence. What is left for further work is to see how this same analysis extends to the cases in which casi modifies quantifiers or adjectives, as in (1b) and (1c). Other features of casi, like its ‘orientation’ or argumentative value (as it comes out in the contrast between almost and barely in English) may also be amenable to an analysis in terms of zero-on scales. I will finish this paper by showing how one curious feature about the semantics of casi can be clarified by an analysis in terms of scales.

5.2. Sadock’s paradox

An analysis of casi as a scalar operator can also help in solving a paradox, which Sadock (1981) pointed out for almost, the English translation of casi. I will adapt Sadock’s discussion to the Spanish example. The problem that leads to the paradox is that a sentence of the form ‘casi S’, like (25a), conveys a sentence of the form ‘no S’, like the one in (25b).

(25a) Andy casi tradujo el poema.
‘Andy almost translated the poem.’

(25b) Andy no tradujo el poema.
‘Andy did not translate the poem.’

Is the relationship between these two sentences an implicature or an entailment? There is evidence to believe, Sadock says, that (25a) implicates (25b), since a sentence like (25a) can be combined with a sentence expressing ‘no S’ without redundancy, as in (26).
Sadock recognizes that this is a funny implicature, however, since it cannot be cancelled. This is shown in example (27).

(27)  Andy casi tradujo el poema, y lo que es más, lo tradujo.
     ‘Andy almost translated the poem, and what’s even more, she translated it.’

The noncancellability of the relationship between (25a) and (25b) can be accounted for if it is assumed that (25a) entails (25b). But given the lack of redundancy of an example like (26), it would be necessary to conclude that it is a funny entailment. Hence Sadock’s paradox.

Sadock’s answer to the problem is to say that the relationship between (25a) and (25b) is a funny implicature. I suggest that the other way is to be preferred, given the meaning of casi as a scalar operator and the structure of zero-on scales. In a zero-on scale the origin and the body of the scale are designated by polar opposites. Since casi maps the origin to a point in the body of the scale, it is only natural that it will entail the polar opposite of the sentence which it modifies. How is the informativeness of an example like (26) going to be explained, then? I will suggest that (26) is not really redundant. The tag no del todo ‘not quite’ works as a cancellation of an implicature associated with certain uses of casi. Sometimes, by uttering a sentence of the form ‘casi S’ instead of ‘no S’, a speaker implicates that the current state of affairs should be taken as making ‘S’ true even though ‘S’ is not the case. So, for instance, an Italian soccer fan could have said something like (28a) after Brazil defeated Italy in the ’94 World Cup, implicating that Brazil’s victory is nothing to be really proud of. A Brazilian fan, on the other hand, could have very well answered with (28b).

(28a)  Italia casi le ganó a Brasil.
     ‘Italy almost defeated Brazil.’

(28b)  Casi, pero no del todo.
     ‘Almost, but not quite.’

By considering the meaning of casi as a scalar operator, then, it is possible to gain some insight on Sadock’s paradox.

Acknowledgments

Thanks to John Moore, Michael Israel, Adele Goldberg, and the audience at the Linguistics Forum in UC San Diego, for their help and their comments. Any shortcomings or mistakes are my own.
References


Neurological Evidence For A Functional Basis for Lexical Categories
Barbara Birch
California State University-Fresno

Formal syntactic theories generally hold that lexical categories like noun or verb are marked +N or +V in the mental lexicon reflecting a semantic or inflectional basis for the distinction (Croft 1991, Chomsky, 1965). However, Hopper & Thompson (1984) argue that the nominal or verbal meaning and inflectional characteristics that any root has in a context depend on the extent to which that root fulfills nominal or verbal functions in the discourse, making the discourse function primary and the inflectional marking or nominal/verbal meaning secondary and derivative.

The field of linguistics is now more open to cross-fertilization from neurology and I hope the opposite is also true. This paper is a somewhat preliminary attempt to discuss the nature of lexical categories in linguistic theory in a way that is informed by research done by neurologists with people with aphasia. It is also meant to outline what lexical categories are like so that neurologists acquire a more sophisticated view of them.

Aphasia research shows a very robust finding that in a number of languages, patients have differing problems with the production, but not necessarily the comprehension, of nouns and verbs both when they occur within contexts or without contexts. In general, verbs are more problematic for Broca's aphasics (McCarthy & Warrington, 1985; Miceli, Silveri, Villa, and Caramazza, 1984), and nouns cause more difficulty for Wernicke's aphasics (Caramazza & Hillis, 1991; Goodglass, et.al. 1966). There are several hypotheses to explain these findings: one is that of Miceli, et al., 1988, which we can call a Lexical Hypothesis and another is that of Bates, et al., 1991, which we can call the Semantic Hypothesis.

Lexical Hypothesis: The mental lexicon is divided into subcomponents based on lexical category. Broca's aphasics have more difficulty with verbs because there is a breakdown in the part of the lexicon where verbs are represented. Similarly, nouns are difficult for Wernicke's aphasics because they have a problem in the area where nouns are represented. Thus, dysfunction is lexical and not semantic or conceptual.

Semantic Hypothesis: Nouns and verbs differ in meaning. Noun meaning is related to object and verb meaning is
related to action. Kellogg, 1995, puts it thus: "...the anterior and posterior cortex is differentially involved in the representation and processing of action and object meanings. The cortical regions responsible for the semantic representation of actions lie near the motor cortex (commonly the site of damage for Broca's aphasias) and the cortical regions responsible for the semantic representation of object names draw input from sensory association areas, (which are areas implicated in Wernicke's aphasias)."

Both of these hypotheses are in accord with the traditional or formal syntactic view of lexical categories. First of all, they are distinguished from each other formally in the mental lexicon and second, the categories are meaning-based. However, as I have pointed out in previous publications, these two hypotheses about lexical categories find little support from other psycholinguistic literature (Birch, B. 1989, 1991, 1993; Cofer & Bruce, 1965; Wickens, 1970; Johnson-Laird, et.al., 1974.) In other words, the early and more recent psycholinguistic experiments provide little reason to think that the mental lexicon or semantic memory is partitioned by lexical category or that words are 'tagged' with lexical category information. Many linguists would also argue that there is no good semantic basis for distinguishing nouns and verbs--it is too simplistic to say that nouns 'mean' things and verbs 'mean' actions. There is a third hypothesis to explain the findings from aphasia research which we will come back to shortly.

Let us return to Hopper and Thompson's theory that lexical categories are discourse or function based. Is there any support from aphasia for that view? While most aphasia studies have involved speakers of European languages like English or Italian, Bates, et al., 1991, looked at evidence from Chinese. In Chinese, nouns and verbs are not differentiated by morphological form or inflectional processes. Discourse function is indicated by word order and context. Bates found the same noun/verb dissociation in Chinese aphasics as well, indicating that it is discourse function which is the primary source of lexical category distinctions, casting doubt on claims that lexical categories are acquired primarily through evidence from morphology (Maratsos, 1982).

Bates et al. also found evidence that the noun/verb dissociation also occurs within lexicalized verbal compounds with a V-N structure, like 'look-book', meaning 'read'. If we assume that common compound words are
listed in the lexicon, and I think we should, then the
distinction between noun and verb occurs somewhere other
than in the lexicon. (See Birch, forthcoming, for further
analysis of this finding.)

As one might expect, Bates used this evidence to argue
against the Lexical Hypothesis and for her own Semantic
Hypothesis, but there is some strong evidence against her
hypothesis as well. Miceli, et al. (1984) found that when
patients who have a deficit in verb retrieval are shown
pictures of actions, they often substitute an related
nominal form to describe the picture. Instead of 'cry'
for example, they will say 'tears' or instead of 'swim'
they will say 'flippers'. Miceli et al. argue
convincingly that the problem for these patients is not
in the meaning or the concept, but in the retrieval of
the word class.

To summarize, there is good evidence that the noun/verb
production dissociation found in aphasics is not
necessarily due merely to problems in the lexicon nor to
problems with conceptualizing. Bates et al. and Miceli et
al. have shown each other's theories to be lacking in
complexity. Also, there is evidence from Chinese that
lexical categories are functionally based not formally
based.

Now let me present the third hypothesis to explain the
noun/verb problem in aphasics, which I call the **Cognitive
Hypothesis**. Other recent evidence from neurolinguistics
is consistent with the view that for English speakers at
least, although there may be no difference in the lexicon
or semantic memory, there is some kind of cognitive
difference between nouns and verbs. Damasio & Damasio
(1993) conclude the following from a study of brain
damage and verbal ability in three aphasics patients:

> ....the systems which mediate access to
congrete nouns are anatomically close to
systems which support concepts for concrete
entities [and]....the systems which mediate
access to verbs are located elsewhere and are
anatomically close to those which support
concepts of movement and relationship in
spacetime....It seems plausible that the
systems which enact the two-way linkage
between concrete entity concepts or action
concepts, and the corresponding nouns or
verbs, should also be relatively separate
albeit interactive.
I understand this view to be that there is a dynamic lexicon which is undifferentiated with respect to noun and verb and that there are action/relation and thing concepts in a dynamic semantic memory. Mediating between the lexicon and the different types of concepts and accounting for lexical access are two different interactive systems, one which results in nouns and one which results in verbs. Deficiencies in retrieval of nouns or verbs could result from impairment in any of these three areas. Before discussing this hypothesis further, let me discuss some of Damasio & Damasio's findings.

Damasio & Damasio found that retrieval of common nouns was significantly impaired while retrieval of verbs was normal or close to normal in two of the three patients they described. However, of the common nouns, retrieval of animal and fruit/vegetable names was much more impaired than retrieval of the names of tools/utensils:

<table>
<thead>
<tr>
<th></th>
<th>Animals</th>
<th>Fruits/Vegetables</th>
<th>Tools/Utensils</th>
<th>Verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boswell</td>
<td>24%</td>
<td>25%</td>
<td>76%</td>
<td>92%</td>
</tr>
<tr>
<td>AN-1033</td>
<td>51%</td>
<td>54%</td>
<td>70%</td>
<td>96%</td>
</tr>
</tbody>
</table>

Although Damasio & Damasio do not tell us exactly which common nouns were used in the study, I point out that animals, fruits, and vegetables are prototypical nouns related to prototypical nominal concepts, but words for tools and utensils are frequently indeterminate in lexical category; hammer, nail, wrench, spoon, and fork can be either nouns or verbs, either related to entities or actions. I suggest that the patients with impairment in the retrieval of nouns had better success with these because they can retrieve some of them as verbs.¹

Damasio & Damasio argue that these neurological connections are acquired through habit and associative learning based on direct experience with words used in discourse. If, as Hopper and Thompson would have it, discourse function determines meaning and inflection, discourse function is the primary source of the unconscious knowledge that we have stored about words and their lexical categories. That is why we find the differential effects of names of fruits and vegetables and names of tools or utensils. The neurological findings support Hopper & Thompson's contention that lexical categories are functionally derived, but they also strongly suggest to me that function at some point becomes cognition.
The difference between nouns and verbs in English comes about because discourse function is translated through experience into cognitive pathways and connections. Now I can go back to the three level Cognitive Hypothesis as proposed by Damasio & Damasio.

Recall that there is evidence that words are not tagged with lexical category in the mental lexicon, nor does the lexicon have subgroupings based on lexical category. There is evidence that semantic memory is not subdivided in semantic fields by notions of 'action' or 'object.' Some noun/verb deficits in aphasics are not located in the lexicon. Some noun/verb deficits in aphasics are not located in the concepts in semantic memory.

So here is a slight revision of the Cognitive Hypothesis in its adaption to a linguistic theory of lexical categories: There is a mental lexicon undifferentiated with respect to lexical category. There is also a sea of concepts in semantic memory which are not organized by notions of 'action' or 'thing' but are organized into semantic fields like 'job,' 'proper name', 'fruit', 'color', etc. Ranging over these concepts are cognitive operators like NOUN and VERB, which combine with the meaning concept prior to lexical selection. These operators would be on the order of those proposed by Johnson-Laird, 1983:413.

The operators form part of Damasio & Damasio's interactive mediation systems; they are cognitive but derived from exposure to discourse. Our notion of lexical categories arises from learned preferences for certain operator/concept/word combinations stemming experience with words in various discourse functions; 'apple' generally used as a noun, 'fork' used sometimes as noun and sometimes as verb, and so on.

I suggest that some noun/verb dissociations in aphasics are due to impairment of either the noun or the verb operator. The verb operator is 'located' near the motor cortex and the noun operator 'draws input' (to use Kellogg's terms) from sensory association areas. Aphasics who have difficulty with the phonological or orthographic encoding of words may have deficits in the lexicon, aphasics who display marked difficulties with discrete semantic fields (like proper names) may have deficits in the concepts stored in semantic memory, and aphasics who retain knowledge of concepts but have either narrow or wide-ranging nominal difficulties may have impairment of the NOUN operator.
In conclusion, I would like to articulate two suggestions for neurologists. First, monolingual Salish or Wakashan speakers with aphasia (if there are any) could be tested. Distinctions between noun and verb in those language families occur at the predicate level, not at the level of lexical roots. If the same noun-verb dissociation were to be found, I believe it would be evidence against the Lexical Hypothesis and the Semantic Hypothesis. It would also provide a glimpse into the interplay of the noun-verb distinction, syntactic structure, and discourse, because data from those languages suggest that the NOUN and VERB operators can work at the level of predication and not roots.

Second, studies which have shown differential production of nouns and verbs in English-speaking aphasics must be replicated with test items which are more strictly controlled. Although the experimenter may say that walk or run is a verb, the linguist knows that they often occur as nouns as well. Although the experimenter may believe that hand or head are nouns, they also frequently occur as verbs. Neurologists must take into consideration that lexical categories in English and other languages are dynamic and not static, fuzzy and not discrete. These characteristics are to be expected in lexical categories that are functionally derived. More carefully crafted experiments may show that aphasia patients display a range of deficits which may indicate relative nouniness and relative verbiness of words in syntactic contexts.

Notes

'I am grateful to a participant at the BLS Conference for pointing out the continuum of "animacy" to be found in these figures.

Selected References


Birch, B. forthcoming. Lexical Categories within Compounds.


Cofer, C. and Bruce, D. (1965) Form-class as the basis for clustering in the recall of nonassociated words. JVLVB 4:386-389.


Generic Demonstratives*

Brian F. Bowdle and Gregory Ward
Northwestern University

Previous analyses of generics have focussed on three types of generic NPs: bare plural NPs, and singular NPs headed by either the indefinite or definite article (e.g., Burton-Roberts 1976; Carlson 1977a,b; Declerck 1991; Langacker 1991). However, existing analyses do not account for an additional type of generic reference, which we term GENERIC DEMONSTRATIVES, as illustrated in (1).

(1) A: My roommate just bought a Labrador.
    B: Those Labradors make great pets.

Here, B is using a demonstrative NP generically to refer to the kind ‘Labradors’. In this paper, we provide a cognitive account of such demonstratives. We will argue that generic demonstratives mark the kind being referred to as a relatively subordinate or homogeneous kind located among the speaker’s and hearer’s private shared knowledge. We also present the results of a preliminary experiment in support of our analysis.

Generic NPs

As is well known, bare plurals in English can receive either an ‘existential’ or ‘generic’ interpretation, as illustrated in (2):

(2) a. Spotted owls are flying overhead.
    b. Spotted owls have excellent vision.
    c. Spotted owls are rare.

In (2a), the bare plural spotted owls is interpreted existentially; that is, (2a) is true just in case there exist two (or more) spotted owls flying overhead. As Carlson notes (1977a), the existential interpretation of a bare plural is synonymous with the corresponding existentially quantified NP; thus, spotted owls is equivalent to some spotted owls in (2a). In contrast, both the object-level predicate has excellent vision in (2b) and the kind-level predicate rare in (2c) permit generic interpretations of the bare plural. The two types differ in that object-level predicates involve attributes that are characteristic of individual members of the kind in question, whereas kind-level predicates involve attributes that apply to the class as a whole.\(^1\)
In addition to bare plural NPs, other forms may receive a generic interpretation as well. For example, it is well known that singular NPs with the definite article may be used generically, as in (3):

(3) a. The spotted owl has excellent vision.
    b. The spotted owl is rare.

Again, both the object-level predicate in (3a) and the kind-level predicate in (3b) permit generic interpretations.

Similarly, singular NPs with the indefinite article may be used generically; however, their use is more restricted. As has been noted (Carlson 1977a), kind-level predicates are incompatible with the indefinite article, as seen in (4):

(4) a. A spotted owl has excellent vision.
    b. #A spotted owl is extinct.

Thus, while bare plural NPs and singular NPs with the definite article may be used to refer to the class as a whole, indefinite NPs may be used generically to refer only to a ‘representative instance’ of a class (e.g., Talmy 1988, Langacker 1991).

**Generic Demonstrative NPs**

In addition to bare plurals and singulars with the definite or indefinite article, generic NPs may also be realized with a demonstrative determiner. Examples of such generic demonstratives are provided in (1) above, and in (5) below:

(5) a. A: My roommate owns an IBM ThinkPad.
    B: Those IBM ThinkPads are quite popular.
    b. A: My roommate owns an IBM ThinkPad.
    B: That IBM ThinkPad is quite popular.

Here, both the single and plural demonstrative are being used generically, as a statement about the kind ‘IBM ThinkPad’. Note that demonstrative NPs share with bare plural NPs and singular NPs with the definite article the property of permitting both object-level and kind-level predicates, as seen in (1) and (5), respectively.

While the examples in (5) employ the distal form of the demonstratives, the proximal form is possible as well, as seen in (6):

(6) a. [in front of a computer] These IBM ThinkPads are amazing!
    b. [in front of a computer] This IBM ThinkPad is amazing!

The choice between the distal and proximal forms depends primarily on the spatial location of the demonstratum: in (5) the demonstratum is distant from the speaker while in (6) it is proximate.

The felicitous use of generic demonstratives is subject to a number of constraints. First, we have found that the predicate of a generic demonstrative is typically evaluative, as seen in (7):
(7)  A: My cousin just returned from Canada with an adorable Labrador retriever puppy.
    B₁: Those Labradors are extremely loyal, you know.
    B₂: #Those Labradors were first bred in Newfoundland, you know.
    B₃: Labradors were first bred in Newfoundland, you know.

Here, the generic demonstrative is felicitous with the evaluative predicate of B₁’s utterance, but not with the more factual predicate found in B₂’s. Note that this constraint does not apply to generics in general, as evidenced by the felicity of the corresponding bare plural generic in B₃’s utterance. This evaluative quality of generic demonstratives reflects the speaker’s emotional involvement with or reaction to the kind in question, and relates to a range of demonstrative uses that has been characterized as EMOTIONAL or EMPATHETIC DEIXIS (e.g., R. Lakoff 1974, Lyons 1977, Ariel 1990, Chen 1990).²

Generic demonstrative NPs are further constrained in that they must represent a kind assumed to be already familiar to the hearer. That is, the kind itself must constitute private shared knowledge (Joshi 1982). Consider (8):

(8)  a.  A: My brother just bought a small car.
    B₁: Those small cars are dangerous!
    B₂: Small cars are dangerous!

b.  A: My brother just bought a red car.
    B₁: #Those red cars are so garish!
    B₂: Red cars are so garish!

The difference in felicity between the use of those small cars and those red cars as generics can be attributed to the ease with which the former, but not the latter, can be construed as a ‘hearer-old’ (Prince 1992) category. In (8a), B is licensed to assume that ‘small cars’ is familiar to A as a category. Thus, the kind ‘small cars’ would be processed as a coherent concept, with members of the category possessing many properties in common. In (8b), however, the generic demonstrative is infelicitous because B cannot plausibly assume that ‘red cars’ constitutes a pre-existing category for A. Although the kind ‘red cars’ could easily be constructed, it is less coherent and relatively ad hoc. Note that the corresponding bare plural is felicitous for both kinds of cars (B₂’s response).

**Plural Demonstrative NPs**

In addition to the general constraints outlined above, the felicity of plural generic demonstratives is sensitive to a further restriction not shared by other generic NPs. Consider the data in (9):

(9)  a.  A: My roommate just bought a dog.
    B₁: Dogs make great pets.
    B₂: #Those dogs make great pets.
b. A: I'm thinking about buying a new car.
   B₁: Cars can be expensive.
   B₂: Those cars can be expensive.

In both (9a) and (9b), generic reference with a plural demonstrative NP is infelicitous. This difference in acceptability cannot be accounted for in terms of the generic context alone, since the predicates make great pets and can be expensive are both object-level and evaluative. What seems to distinguish generic demonstratives from other types of generic NPs is the specificity of the kind being referred to, as illustrated in (10):

(10) a. A: My roommate just bought a Labrador.
   B₁: Labradors make great pets.
   B₂: Those Labradors make great pets.

b. A: I'm thinking about buying a new sportscar.
   B₁: Sportscars can be expensive.
   B₂: Those sportscars can be expensive.

Note that the kinds being referenced here ('Labradors' and 'sportscars') are more specific instantiations of the kinds in (9) ('dogs' and 'cars'). Unlike other types of generic NPs, the felicity of plural generic demonstratives varies across analogous contexts depending on the specificity of the evoked kind.³

This notion of kind specificity is captured by folk taxonomies, which partition conceptual domains in terms of vertical levels of abstraction or category inclusiveness (e.g., Kay 1971, Berlin et al. 1973, Rosch et al. 1976). Rosch et al. (1976) distinguished between three levels of categorization: the superordinate, basic, and subordinate levels. Superordinate level categories are the most general, and are comprised of relatively heterogeneous sets of items. For example, the superordinate category 'animal' includes members as diverse as dogs, birds, and fish. Basic level categories exhibit an intermediate degree of inclusiveness. The basic level has been claimed to represent the most efficient level of categorization, as it maximizes within-category similarity and minimizes between-category similarity. For example, members of the basic level category 'dog' tend to be relatively similar to each other, but quite different from members of other kinds of animals, such as birds or fish. Finally, subordinate level categories are the most specific, and are comprised of relatively homogeneous sets of items. For example, members of the subordinate category 'Labrador' are far more similar to each other than members of the basic level category 'dog'.

A central feature of taxonomies is that, as one moves from superordinate to subordinate levels, there is a steady reduction in the number of salient internal contrast sets. That is, category variability is lowest for subordinates. We claim that it is this property of categories that accounts for the distribution pattern of plural generic demonstratives exemplified in (9) and (10). Unlike bare plurals, the use
of plural generic demonstratives seems to implicate that members of the kind in question are "all alike". In other words, the generic use of a plural demonstrative conveys that the predicate holds equally for all category members, rather than being merely typical or characteristic. Because plural demonstratives indicate a high degree of homogeneity, the kind being specified should possess few salient internal contrasts. Thus, plural demonstratives are most acceptable with relatively subordinate generic sets (e.g., 'Labrador' or 'sportscar').

The plausibility of such a relationship between the felicity of plural generic demonstratives and category variability naturally depends on the assumption that individuals are in fact sensitive to variability information. This assumption has been supported by a number of psychological studies. For example, judgments of the degree of variability among category members are positively correlated with the actual level of variability (e.g., Lathrop 1967, Nisbett & Kunda 1985, Park & Hastie 1987). Similarly, the perceived variability of a group has been shown to increase with the number of subordinates comprising the group (e.g., Park & Judd 1990). Category variability information has been found to influence a number of cognitive processes, including property induction (e.g., Nisbett et al. 1983, Park & Hastie 1987, Rips 1975) and categorization (e.g., Fried & Holyoak 1984, Rips 1989).

To investigate our hypothesis, we conducted a preliminary study in which subjects were asked to evaluate a series of generic statements. In this study, 24 undergraduate students were each presented with 12 brief conversations between two participants, A and B. In each conversation, A's utterance concerned a single specific member of a particular category, and B's utterance was a subsequent generic statement about that category. Two separate factors were manipulated in constructing these conversations. First, we varied the taxonomic level of the category in question. One version of each conversation involved a basic level category, as in (11). The other involved a subordinate level category, as in (12):

(11) A: My parents want to give me their cat for the summer.  
    B: Cats can be pretty destructive.

(12) A: My parents want to give me their Siamese cat for the summer.  
    B: Siamese cats can be pretty destructive.

For each conversation, half of the subjects were given the basic level version, and half were given the subordinate level version. Each subject was presented with an equal number of basic and subordinate level conversations.

Second, we varied the type of generic NP uttered by B. For half of the subjects, B's utterance always contained a bare plural (e.g., cats or Siamese cats). For the remaining subjects, B's utterance always contained a plural demonstrative (e.g., those cats or those Siamese cats).

The conversations were presented in booklets, with three conversations appearing on each page. Subjects were asked to read each conversation and rate the felicity
of B’s response on a five-point scale, where 1 indicated that the utterance was ‘very odd’, and 5 indicated that the utterance was ‘very natural’.

In accordance with our hypothesis, we predicted an interaction between taxonomic level and type of generic NP. For conversations containing bare plurals, the naturalness of the generic statements should be relatively insensitive to the taxonomic level of the evoked kind. In contrast, for conversations with generic demonstratives, statements about subordinate kinds should be rated as significantly more natural than those about basic kinds, due to the decrease in homogeneity associated with the latter.

The acceptability ratings were submitted to a 2 (taxonomic level) x 2 (type of generic NP) analysis of variance. As predicted, there was a significant interaction between the two factors (F(1,22)=25.31, p<.001). Whereas generic statements containing bare plurals were equally natural whether the evoked kind was subordinate (m=4.06) or basic (m=3.99), generic statements containing plural demonstratives were significantly more natural for subordinates (m=4.01) than for basics (m=2.83). There were also significant main effects of taxonomic level (F(1,22)=32.02, p<.001) and type of generic NP (F(1,22)=7.80, p<.025), due mainly to the relatively low acceptability of basic kinds with demonstratives. Thus, statements containing demonstratives were judged infelicitous only when the evoked kind was at the basic level. The data constitute strong evidence for the sensitivity of plural generic demonstratives to category variability information. Note that these results cannot be explained by positing a different relationship between the evoked kind and the predicate across taxonomic levels. If this were the case, the effect of taxonomic level found for demonstrative NPs should have obtained for bare plurals as well.

Although we chose to manipulate category variability in terms of taxonomic levels, where subordinates exhibit the greatest relative degree of homogeneity, it should be noted that the felicity of plural generic demonstratives is not strictly dependent on level of abstraction. What is crucial is that the number of salient internal contrasts be minimized. For example, ‘porcupine’ is a basic level category. However, unlike many basic categories, ‘porcupine’ is not further divisible into subordinate kinds for most American English speakers. Thus, the perceived variability of this basic level category is comparable to that of typical subordinates, as indicated by the felicity of the plural demonstrative generic in (13):

(13)  A: My dog was attacked by a porcupine yesterday.
       B: Those porcupines are very territorial.

It is the degree of homogeneity – the extent to which members of the kind are “all alike” – that is central to determining the felicity of plural generic demonstrative NPs, rather than the actual taxonomic level of the kind per se.
Singular Demonstrative NPs

Singular generic demonstratives are similar to plural generic demonstratives in their sensitivity to category variability. In the case of singular demonstratives with mass nouns, this analogy is especially transparent. Consider (14):

(14)  A:  I'll be serving ice cream at the party.
       B1:  Ice cream is a real crowd pleaser.
       B2:  #That ice cream is a real crowd pleaser.

Here, generic reference with a singular demonstrative NP is quite odd. However, when a subordinate substance kind is evoked, as in (15), the singular demonstrative NP is felicitous:

(15)  A:  I'll be serving Ben & Jerry's ice cream at the party.
       B1:  Ben & Jerry's ice cream is a real crowd pleaser.
       B2:  That Ben & Jerry's ice cream is a real crowd pleaser.

It is the additional specificity provided by a brand name in this example that renders the singular generic demonstrative acceptable.

In the case of count nouns, the range of distribution exhibited by singular generic demonstratives is even more restricted than that exhibited by plural generic demonstratives. Consider (16):

(16)  A:  My roommate owns a laptop computer.
       B1:  Those laptop computers are pretty versatile.
       B2:  #That laptop computer is pretty versatile.

Here, the subordinate kind 'laptop computer' is sufficiently homogeneous to allow the felicitous generic use of a plural – but not singular – demonstrative. However, this contrast between singular and plural demonstratives disappears as the variability of the evoked kind approaches zero, as in (17):

(17)  A:  My roommate owns an IBM ThinkPad.
       B1:  Those IBM ThinkPads are pretty versatile.
       B2:  That IBM ThinkPad is pretty versatile.

In contrast to plural demonstrative generics, whose felicitous use requires that the kind in question be relatively homogeneous, singular demonstrative generics involving count nouns further require that the individual exemplars of the evoked kind are conceptually identical or functionally indistinguishable. This requirement explains why singular generic demonstratives are generally odd when the evoked kind is a natural kind, even when the category is sufficiently subordinate to not allow further subdivision. Consider (18):
(18)  A: My parents want to give me their Siamese cat for the summer.
      B: #That Siamese cat is pretty destructive.

Here, even though it is difficult for most individuals to differentiate the kind ‘Siamese cat’ into further subordinates, the singular demonstrative is infelicitous. This difficulty can be attributed to our knowledge of natural kinds: in general, no two members of any given species will be exactly alike. In contrast, the exemplars of the lowest level categories in artifact taxonomies (e.g., ‘IBM ThinkPad’, ‘Honda Civic’) tend to be functionally indistinguishable by design. For such categories, the singular generic demonstrative will be felicitous.

Discussion

The observed pattern for generic demonstrative NPs may be summarized as follows. In all cases, the felicitous use of generic demonstratives requires that the evoked kind be relatively homogeneous. That is, these forms require a minimum of salient internal contrasts. Singular generic demonstratives involving count nouns further require that the individual exemplars of the evoked kind are conceptually identical. Thus, the interpretation of generic demonstratives is related to the variability of the category in question.5

Of course, the variability associated with a given category is not a static property. First, increasing familiarity or expertise with a kind generally leads to an increase in perceived variability (e.g., Linville et al. 1989). For example, a cat expert who has learned to distinguish among various types of Siamese cats would presumably find generic statements containing those Siamese cats to be quite odd. On the other hand, an individual possessing unusually limited experience with or knowledge of cats might find a generic statement containing those cats to be perfectly natural.

Second, if the initial impressions of a category are based on idealized summary descriptions (such as stereotypes or prototypes), the perceived variability will tend to be lower than if no such information is available (e.g., Park & Hastie 1987, Smith & Zarate 1990). Indeed, a central component of most stereotypes is the belief that the group in question is a relatively homogeneous entity (Kashima & Kashima 1993). For this reason, generic demonstratives are frequently applied to racial or ethnic groups for which widely-held stereotypes exist (e.g., Those Japanese..., Those Mexicans...).

Finally, our data on generic demonstratives call into question two existing claims about the use of definite determiners in English generic NPs. First, it has been argued that definite plural generics are unacceptable in English: whereas the dog may receive a generic interpretation, the dogs may only be interpreted existentially (Declerck 1991, Langacker 1991). Second, it has been argued that definite mass nouns in English (e.g., the water) are infelicitous as generic NPs (Langacker 1991). However, the results of our investigation indicate that the generic interpretation of plural and mass demonstrative NPs is indeed possible, subject to the conditions outlined above.
Extensions

Thus far, all of our examples of generic demonstratives have involved evaluative predicates that have been explicitly mentioned. However, a generic interpretation of demonstratives is possible even when the predicate is implicit, as in (19):

(19)  
   a.  Labrador!
   b.  Those Labrador!

Both of the utterances in (19) could be used in the presence of a single Labrador behaving in a prototypical manner. Such utterances can be analyzed as a case of conversational R-implicature (Horn 1984), instructing the hearer to supply the relevant object-level predicate. Like demonstrative NPs in generic contexts, this construction is sensitive to category variability, as illustrated in (20):

(20)   
   a.  Dogs!
   b.  #Those dogs!

Whereas (20a) can be used generically in the presence of a single dog, (20b) can only be interpreted existentially.

Our analysis of generic demonstrative determiners extends straightforwardly to demonstrative pronouns. Consider (21) and (22):

(21)  
   A: My sister owns a Honda Civic.
   B₁: Those are great cars.
   B₂: That’s a great car.

(22)  
   A: My sister is a philosopher.
   B: Oh, she’s one of those...

In each case, the demonstrative pronoun may receive a generic interpretation. Given that the felicity of such uses is constrained by the variability of the evoked kind, this has important implications for the interpretation of pronouns in ambiguous TOKEN-FOR-TYPE METONYMIES (cf. Gibbs 1994). For example, if a speaker points at a single isolated car and says I want one of those or That’s what I want, the speaker will most likely be understood as referring to a specific type of car (e.g., ‘Honda Civic’), rather than to a higher level category (e.g., ‘cars’ or ‘vehicles’).

Finally, our claims concerning the sensitivity of demonstratives to category variability extend to proper name demonstratives in habitual statements, as illustrated in (23):

(23)  
   That Dr. Williams is always falling asleep in surgery.
Demonstratives with proper names are acceptable just in case the predicate is time-stable or denotes an essential trait of the referent. Even in the absence of such grammatical markers of habituality as the present progressive in (23), felicitous use of a proper name demonstrative nonetheless requires a context in which the relevant predicate is time-stable. Consider (24):

(24) That Dr. Williams fell asleep in surgery today.

Here, if Dr. Williams has never before fallen asleep in surgery and is believed to be an otherwise competent physician, the use of a demonstrative would be odd. In this way, proper name demonstratives are analogous to generic demonstratives, with the former being sensitive to the variability of an individual across time, and the latter being sensitive to variability across category exemplars.

Conclusion

In this paper, we have investigated the use of demonstratives in generic contexts. We have shown that demonstrative NPs may be used generically, and are closely aligned with other definite generics in that they refer to kinds rather than representative instances. Such uses are subject to a number of general constraints, including that a) the predicate be evaluative; and b) the kind in question constitute private shared knowledge. More importantly, we have shown that generic demonstratives are sensitive to the variability of the evoked kind. Specifically, their felicitous use requires that the category be relatively homogeneous. We take the pattern of generic demonstratives that we have identified to be a grammatical reflex of this aspect of underlying category knowledge.

These findings have implications for the analysis of a variety of phenomena, including the interpretation of ambiguous token-for-type metonymies and the use of proper name demonstratives in habitual statements. More generally, these findings are consistent with a growing body of research claiming that many principles of cognitive representation (e.g., frames, prototypes, reference points) are projected onto linguistic structure. To our knowledge, however, the present results constitute the first real evidence that a particular grammatical device is sensitive to category variability information.
Notes

* We would like to thank Betty Birner, Janet Pierrehumbert, Beatrice Santorini, and Jeff Sherman for their comments and suggestions. The research described in this paper was supported by NIDCD grant R01-DC01240 (Ward).

1 Carlson (1977a:56 ff.) provides a useful, yet imperfect, diagnostic for distinguishing generic uses of bare plurals from existential ones: in extensional non-negative contexts, a superordinate bare plural can be substituted for a hyponym under an existential interpretation, but not a generic one. Thus, for example, if it is true that spotted owls are flying overhead, then it is true that birds are flying overhead. However, if it is true that spotted owls have excellent vision, it does not follow that birds have excellent vision. Thus, only the existential interpretation permits the substitution of hyponymous expressions with no change in truth-value.

2 Interestingly, the same constraint applies to the use of the demonstrative with proper names, as illustrated in (i):
   (i) a. That Jonathan is a pain.
   b. #That Jonathan is an engineer.
Here, we see that the demonstrative with a proper name is felicitous only with a relatively evaluative predicate. See below.

3 We’re not considering cases involving contrastive accent, e.g. THOSE dogs make great pets, or Those LABRADORs make great pets. While a full examination of the role of accent must await further study, we nonetheless note that accented demonstratives induce a contrast among subordinates of the evoked kind, while accented kind terms induce a contrast among categories at the same taxonomic level as the evoked kind.

4 Note that the infelicity of B2’s response in (16) cannot be attributed to the singular number of the NP alone. By replacing the demonstrative with the definite article, the generic interpretation of the singular NP becomes felicitous, as seen in (i):
   (i) The laptop computer is pretty versatile.

5 We have remained relatively silent on the question of whether the homogeneity of the kind in question is defined locally, in terms of the predicate, or globally, in terms of the set of salient attributes associated with the kind irrespective of the particular predicate. That is, do generic demonstratives specify kinds for which the predicate applies invariently to all members, or more generally to kinds whose members do not differ from one another along multiple dimensions? We leave this issue for future research.

6 Interestingly, the constraint that the predicate of a generic demonstrative be evaluative appears to be relaxed in the case of demonstrative pronouns, as demonstrated by the felicity of B’s response in (i):
   (i) A: My sister owns a Toyota Camry.
   B: Those are made in Kentucky.

References


On a scalar operator

Yoon-Suk Chung
University of California, Berkeley

0. Introduction

It is well-known that many linguistic expressions are inherently scalar and
directional in nature. Such expressions not only include quantifiers (all, some,
none, etc), and time-denoting adverbs (always, frequently, sometimes, etc) but also
includes predicates (adore, love, like; must, should, can, etc) and non-temporal
adverbs (almost, even, only, etc).

In this paper I discuss one such Korean word *khenyeng*, which has been
traditionally analyzed as a delimiter with two 'distinct' senses, each of which
triggers a different syntactic environment. I provide a unified syntactic and semantic
analysis of these two allegedly distinct senses, arguing that *khenyeng* is a scalar
conjunction. In other words, *khenyeng* is a conjunction and a negative polarity
trigger (hence, negative conjunction). I also argue that *khenyeng* is a scalar
operator, and that a uniform treatment of the semantics of the *khenyeng*
construction can be formulated in terms of scalar semantics developed in Fillmore et

The organization of the paper is as follows. Section 1 reviews Choy's
(1929/1961) descriptive work, the only serious work on *khenyeng* to date. Section
2 introduces more *khenyeng* examples, showing that Choy's description is not
correct. Section 3 and 4 give a unified analysis of the form and meaning of
*khenyeng*, respectively. The conclusions and directions for further study are
provided in section 5.

I. Previous studies

The *khenyeng* sentence begins with a fragment, followed by *khenyeng*,
followed by a full clause, as in (1) and (2).

(1)  fragment     khenyeng     full clause
     sipppul khenyeng     ilpwul-to mos patassta
     ten dollars let alone one dollar-even not received
     'I did not receive one dollar, let alone ten dollars.'

(2)  fragment     khenyeng     full clause
     sang khenyeng     pel-ul patasseyo
     prize instead of punishment-Acc got
     'Instead of being rewarded, I got a punishment.'

In his seminal descriptive grammar of Korean, Choy (1929/1961)
characterizes *khenyeng* as an 'auxiliary particle' whose semantic function in a
sentence is to add certain auxiliary meanings, such as focus, emphasis, etc. to a
word or a phrase to which it is attached. In the subsequent works in the
transformational generative tradition, however, the term 'auxiliary particle' has been
replaced by 'delimiter' since its semantic function is reinterpreted as delimiting or
specifying the meaning of the element it follows.

According to Choy (1929/1961), *khenyeng* is used, loosely speaking, when
the result of one event compared to that of another event in the sentence, is so
obvious that the speaker does not need to mention it. Choy draws a distinction between the two senses of *khenyeng*.

As illustrated in (1) above, the first sense of *khenyeng* (K1, henceforth) concerns the case in which the speaker is explicit about the implausibility of the more likely state of affairs expressed in the post-*khenyeng* negative full clause, implicating the impossibility of the less likely state of affairs expressed in the pre-*khenyeng* fragment. As represented in (2) above, the second apparent sense (K2, henceforth) concerns the case in which the expected state of affairs in the fragment is denied and the unexpected contrary state of affairs is affirmed in the positive full clause.\(^7\)

Choy (1929/1961) goes on to say that NPs flanked by K1 are of the 'same' kind, and a predicate in the K1 sentences has a predicative scope over NPs in the pre-*khenyeng* as well as in the post-*khenyeng*. On the other hand, the NPs conjoined by K2 are of a 'different' kind and the predicate in the K2 sentence has a predicative scope only over the NP in the post-*khenyeng* part of the sentence.

Choy's (1929/1961) description of *khenyeng* is summarized as follows:

(3) analysis of *khenyeng* by Choy (1929/1961)

*khenyeng* is a lexical item with two different meanings and syntactic environments such that

- a in the K1 case, the impossibility of the less likely state of affairs is followed by the implausibility of the more likely state of affairs, while in the K2 case, denial of the expected state of affairs is followed by affirmation of the unexpected contrary state of affairs,
- b words conjoined by K1 are of the same kind, while words conjoined by K2 are of different kinds,
- c the verb of the K1 sentence is predicative of the word preceding *khenyeng* as well as the one following *khenyeng*, while the verb of the K2 sentence is only predicative of the word following *khenyeng*.

In the next sections, I will show that the distinction of two forms and meanings associated with *khenyeng* expressions is unnecessary and may be replaced by a unitary account.

II. More examples

To arrive at a fuller understanding of the formal and semantic properties of *khenyeng*, we need to consider more examples that contain *khenyeng* expressions. Although Choy (1929/1961) only provides the *khenyeng* examples that conjoin NPs, as shown in (1) and (2) above, the range of syntactic categories connected by *khenyeng* is not so limited. In addition to NPs, as illustrated by another examples in (4), *khenyeng* can usually conjoin any two like categories including adverbial phrases (5), postpositional phrases (6), and verb phrases (7).

(4) a tayhakko-nun\(^8\)-khenyeng cwunghakkoyo-to colep mos haysseyo
college-Contr-let alone middle school-even graduation cannot did
'He could not graduate from middle school, let alone college.'
b chingchan-un-khenyeng kkwucwung-ul tulesssupnita
praise-Contr-instead of scolding-Acc got
'Instead of being praised, I got a scolding.'
(5) a manhi-nun-kenyenɡ cokum-to epsta
   a lot-Contra-let alone a little-even do not have
   'I do not have a little, let alone a lot.'

   b cacwu-nun-kenyenɡ acwu ittakumssik nathanakonhaysseyo
      frequently-Contra-instead of very once in a while would show up
      'Instead of frequently, he would show up once in a great while.'

(6) a hayswuyokcang-ey-nun-kenyenɡ swuyengcang-ey-to an kasseyo
      beach-to-Contra-let alone swimming pool-to-even not went
      'I did not go to a swimming pool, let alone a beach.'

   b tayhaksayng-mankhum-un-kenyenɡ kotunghaksayng-pota-to yenge-lul
      mos hanta
      college student-as-Contra-let alone high school student-as-even English-Acc
      cannot speak
      'He cannot speak English as well as a high school student, let alone as well
      as a college student.'

(7) a ssu-ki-nun-kenyenɡ ilk-ci-to mos haysssta
      write-Nml-Contra-let alone read-Nml-even cannot did
      'He could not read, let alone write.'

   b ku-ka ka peli-ese sepsepha-ki-nun-kenyenɡ siwenhayyoo
      he-Nom go away-since sorry-Nml-Contra-instead of glad
      'Instead of feeling sorry, I am glad that he has gone.'

   c ku-nun pap-ul mek-umyense-nun-kenyenɡ cha-lul masi-myense-to TV-lul
      mos ponta
      he-Top meal-Acc eat-while-Contra-let alone tea-Acc drink-while-even TV-Acc
      cannot watch
      'He cannot watch TV while drinking his tea, let alone while eating his
      meal.'

   In addition, contrary to Choy's (1929/1961) claim that the K1 type verb has
   a predicative scope over the NPs preceding as well as following khenyenɡ, there
   are also cases of K1 sentences in which the verb does not serve as a predicate of the
   pre-khenyenɡ NP. Compare (8) with (9).

(8) a pwule-nun-kenyenɡ yenge-to mos hanta
      French-Contra-let alone English-even cannot speak
      'He cannot speak English, let alone French.'

   b pwule-lul mos hanta
      French-Acc cannot speak
      'He cannot speak French.'

   c yenge-lul mos hanta
      English-Acc cannot speak
      'He cannot speak English.'

(9) a pi-nun-kenyenɡ kwulum-to kkici anhassta
      rain-Contra-let alone cloud-even cloud up did not
      'It was not cloudy, let alone rainy.'

   b * pi-ka kkici anhassta
      rain-Nom cloud up did not
      'It was not rainy.'

   c kwulum-i kkici anhassta
cloud-Nom cloud up did not
'It was not cloudy.'

d * pi-to kwulum-to kkici anhassta
rain-also cloud-also cloud up did not
'It was neither rainy nor cloudy.'

For example, unlike the verb in (8a), the verb *kkita 'cloud up' in (9a) cannot serve as a predicate of the pre-*khenyeng* NP *pi 'rain'. This is shown by the ungrammaticality of (9b) in which the verb *kkita 'cloud up' cannot have *pi 'rain' as a possible argument. The grammaticality of (9a) thus shows that even when the NPs require different verbs, they can be connected in a *khenyeng* sentence by the verb that subcategorizes for the post-*khenyeng* NP. The point may be sharpened by the comparison of the examples (9d) and (9a). The difference in grammaticality between (9d) and (9a) may be attributed to the predication of the verb whose scope is assigned by the construction involved. That is, the verb in (9d) is predicative of the first NP as well as the second NP, while the verb in (9a) is only predicative of the post-*khenyeng* NP. As we will see in section 3.1, this observation is significant since it is directly relevant to one of our arguments that *khenyeng* is a proposition conjunction, not merely a constituent conjunction.

*Khenyeng* can thus conjoin pairs of maximal phrases of a wide range of syntactic categories, and not merely NPs. And when conjoining NPs, *khenyeng* does not require that the first NP be an argument of the verb which subcategorizes for the second NP.

Having briefly shown, by considering more examples, that Choy's description of *khenyeng* is not correct, I will take up the syntactic and semantic features of *khenyeng* in turn.

III. Grammar of *khenyeng*

3.1 *Khenyeng* is a conjunction

The first argument that I want to make on the syntactic side is that *khenyeng* is a conjunction, not a delimiter as is widely assumed in the literature (cf. Choy (1929/61), Martin (1992)). Even though *khenyeng* may not be a canonical conjunction, there is evidence both for the claim that *khenyeng* is a type of conjunction, and that it is not a delimiter.

*Khenyeng* shares some syntactic properties with the focus delimiters such as *nun 'only concerned', to 'also, even', and *ya 'when it comes to' in that it is neither preceded nor followed by the nominative, accusative or genitive case markers, and in that it can replace them in appropriate syntactic environments. But *khenyeng* shows a remarkable difference from the focus delimiters in several ways.

The first difference can be addeduced from a general constraint concerning delimiters, namely, that there is a restricted ordering among them. Yang (1972) classifies delimiters into three sub-categories based on their distributional properties and their mutual co-occurrences: X-lim (*mace 'even, indeed, including', *mata 'each', *kkaci 'up to, even', *pwaithe 'from')

Y-lim (*man 'only, exactly', *pakk-ey 'only'), Z-lim (*nun 'only concerned', to 'also, even', (i)ya 'when it comes to', *na 'and, or', (ii)lato 'even the last choice or recourse'). Yang then argues that when the three kinds of delimiters co-occur, X-lim always precedes Y-lim, which always precedes Z-lim. According to Yang (1972), the focus delimiters including *nun* are Z-lim, coming last in a string of delimiters. No other delimiters are permitted to follow Z-lim, as shown by (10a). If *khenyeng* is a delimiter, it should not be able to
follow *nun* either. The grammaticality of (10b), therefore, demonstrates that *khenyeng* cannot be considered a delimiter.\(^{12}\)

(10) a * Mimi-nun-to yeyppu-ci anhta
     Mimi-Top-also pretty-Nml not
     'As for Mimi, she is not pretty, either.'

 b Mimi-nun-khenyeng Swumi-to yeyppu-ci anhta
    Mimi-Contra-let alone Swumi-even pretty-Nml not
    'Swumi is not pretty, let alone Mimi.'

Secondly, *khenyeng* is never immediately preceded by an adverb or an infinitive form of verbs, as is common for the focus delimiters.\(^{13}\) For example, the adverb *manhi* 'much' and the infinitive of verbs *-e* can be followed by the focus delimiter *to* as in (11a) and (12a) but not by *khenyeng* as in (11b) and (12b), respectively.

(11) a manhi-to cokum-to epseyo
     much-also a little-also not have
     'I have neither a little nor much.'

 b *manhi-khenyeng cokum-to epseyo
     much-let alone a little-even not have
     'I do not have a little, let alone much.'

(12) a ilk-e-to tul-e-to po-ci mos hayssta
     read-Inf-also listen to-Inf-also try-Nml cannot did
     'I had the experience of neither reading nor listening to (it)'

 b *ilk-e-khenyeng tul-e-to po-ci mos hayssta
     read-Inf-let alone listen to-Inf-even try-Nml cannot did
     'I did not have the experience of listening to, let alone reading (it).'</n
Thirdly, as is the case for the focus delimiters, *khenyeng* cannot be immediately preceded by most postpositional phrases such as locative, instrumental, etc. For example, the delimiter *to* can follow the locative case marker, *eyse* 'in, at' in (13a), but *khenyeng* cannot in (13b).

(13) a i kos-eyse-to ce kos-eyse-to tampay-lul phiwul swu epssupnita
    this place-in-also that place-in-also tobacco-Acc smoke possibility not exist
    'You are not permitted to smoke here or there.'

 b *i kos-eyse-khenyeng ce kos-eyse-to tampay-lul phiwul swu epssupnita
    this place-in-let alone that place-in-even tobacco-Acc smoke possibility not exist
    'You are not permitted to smoke there, let alone here.'

The last reason to reject the delimiter account concerns the fact that focus delimiters can follow a wide variety of verbal connectives, while *khenyeng* cannot. Consider (14), in which one of the verbal connectives, *-ulyeko* 'in order to', can precede the delimiter *to* but not *khenyeng*.\(^{14}\)
(14) a  cenyek-ul mek-ulyeko-to chyeta po-lyeko-to ha-ci anhassta
dinner-Acc eat-in order to-also look at try-in order to-also do-Nml did not
'He did not intend to eat or look at the dinner.'
b  *cenyek-ul mek-ulyeko-khenyeng chyeta po-lyeko-to ha-ci anhassta
dinner-Acc eat-in order to-let alone look at try-in order to-even do-Nml did not
'He did not intend to look at, let alone eat the dinner.'

Observing that there exists evidence to doubt khenyeng's status as a
delimiter, we turn our attention to some positive observations available, suggesting
treatment of khenyeng as a type of conjunction. First, as shown by the wide variety
of examples given in (4)-(7), khenyeng usually serves to connect two
grammatically equal phrases such as NPs, ADVPs, VPs, etc. Secondly, the
khenyeng construction shows properties that are typically associated with
coordination constructions. For the khenyeng construction, for example,
topicalization as in (15a), relative clause formation as in (15b), and clefting as in
(15c) are possible.15 The examples in (16) show comparable sentences containing a
canonical coordinate conjunctive marker.

(15) a  i chayk-khenyeng ku sinmwun-to Mimi-nun han sikan-ey mos ilkeyo
that book-let alone this paper-even Mimi-Top one hour-in cannot read
'Mimi cannot read a paper in an hour, let alone a book.'
b  Mimi-ka tayli-ki-khenyeng yokha-ci-to anhun namca
Mimi-Nom hit-Nml-let alone yell at-Nml-even not man
'The man who Mimi did not yell at, let alone hit.'
c  cikum Mimi-ka mekko issun kes-un soykoki-khenyeng twaycikoki-to anita
now Mimi-Nom eating thing-Top beef-let alone pork-even not
'What Mimi is eating now is not pork, let alone beef.'

(16) a  i chayk-kwa ku sinmwun-ul Mimi-nun han sikan-ey mos ilkeyo
book-and-paper-Acc Mimi-Top one hour-in cannot read
'As for this book and that paper, Mimi cannot read them in an hour.'
b  Mimi-ka tayli-kena yokha-ci-to anhun namca
Mimi-Nom hit-or yell at-Nml-also not man
'the man who Mimi neither hit nor yelled at.'
c  cikum Mimi-ka mekko issun kes-un soykoki-ntonun twaycikoki-ita
now Mimi-Nom eating thing-Top beef-or pork-be
'What Mimi is eating now is pork, or beef.'

Yet there are other cases which resist a strict coordinate conjunction
account, forcing us to admit that khenyeng is a more tolerant type of conjunction.
First, recall the case seen in (9), which is repeated as (17):

(17) a  pi-nun-khenyeng kwulum-to kkici anhassta
rain-Contr-let alone cloud-even cloud up did not
'It was not cloudy, let alone rainy.'
b  *pi-ka kkici anhassta
rain-Nom cloud up did not
'It was not rainy.'
c kwulum-i kkici anhassta
cloud-Nom cloud up did not
'It was not cloudy.'
d * pi-to kwulum-to kkici anhassta
rain-also cloud-also cloud up did not
'It was neither rainy nor cloudy.'

Since they do not share the same verb, as shown in (17b) and (17c), the two NPs in (17a) cannot be literally said to be connected by khenyeng. Moreover, khenyeng sentences sometimes comprise two independent verb phrases with their own verbs. An example of such a case can be seen in (18) in which the pre-khenyeng part is composed of its separate verb and object:

(18) ton-ul pel-ki-nun-khenyeng sonhay-lul pwassta
money.Acc earn-Nml-Contr-instead of damage.Acc suffer
'Instead of making money, I suffered damage.'

This perplexing counterexample to the constituent conjunction account points to the possibility that khenyeng is better viewed as combining two propositions with the first proposition syntactically realized as a clause fragment. There is evidence that khenyeng sentences are composed of two independent propositions in the sense that the pre-khenyeng part of a sentence is not part of the clause headed by the verb which occurs after khenyeng. The first piece of evidence comes from the clause-bound subject-honorification agreement phenomenon. Our assumption predicts that the honorification of the verb, for example, must agree with a post-khenyeng NP, not with a pre-khenyeng NP. The following sentences bear out our prediction.

(19a) halapenim-kkeyse-nun-khenyeng tongsayng-to theynis-lul an chiko isseyo
grandfather-Nom[hon]-Contr-let alone younger brother-even tennis-Acc not
playing be
'My younger brother is not playing tennis, let alone my grandfather.'

(19b) *halapenim-kkeyse-nun-khenyeng tongsayng-to theynis-lul an chiko
kheyseyo
grandfather-Nom[hon]-Contr-let alone younger brother-even tennis-Acc not
playing be[hon]
'My younger brother is not playing tennis, let alone my grandfather.'

(19a) is grammatical since the non-honorific form isseyo 'be' agrees in honorification with the non-honorific NP tongsayng 'younger brother' in the post-khenyeng part, but not with the honorific NP halapenim 'grandfather' in the pre-khenyeng part, while (19b) is ungrammatical since the honorific from kheyseyo 'be [hon]' cannot agree in honorification with its non-honorific subject tongsayng.

Secondly, the appearance of tense or a passive morpheme in the fragment shows that the pre-khenyeng part belongs to a separate clausal fragment from the post-khenyeng full clause. Consider (20)-(21).
(20) i chayk-un cal ssu-ess-ki-nun khenyeng mwusun malinci al swu-ka epsta
this book-Top well write-Pst-Nml-Contr-let alone what speech know
possibility-Nom not exist
'This book is not clear, let alone well-written.'

(21) chenpwul-i ket-hi-ki-nun-khenyeng paykpwul-to mos kethyesseyo
thousand dollars-Nom collect-pass-Nlm-Contr-let alone hundred dollars-
even cannot be collected
'One hundred dollars were not collected, let alone one thousand dollars.'

To sum, with the negative and positive evidence presented above, I have
shown in this subsection that there is some doubt as to the status of khenyeng as a
delimiter. It is more appropriate to treat khenyeng as a conjunction, albeit a non-
canonical one.

3.2. Khenyeng is a negative polarity trigger
In this subsection, I claim that khenyeng is a negative polarity trigger with
scope only over its pre-khenyeng clausal fragment. Note first, as illustrated by the
(a) examples of (4)-(7) above, sentences containing K1 usually have explicit
negative adverbs mos 'cannot', or an 'do not' before the verbs. K1 sentences can
also have such intrinsically negative verbs as eps- 'not exist', molu- 'not know',
silphayha- 'fail', tteleci- 'fail', pwucok- 'short of', and elyep- 'doubtful', etc.
Consider the examples in (22).

(22) a ku-nun pwule-nun-khenyeng yenge-to molunta
he-Top French-Contr-let alone English-even not know
'He does not know English, let alone French.'

b kummeytal-un-khenyeng unmeytal-to tanun tey-ey silphayhayssta
gold medal-Contr-let alone silver medal-even win opportunity-in failed
'He failed in winning a silver medal, let alone a gold medal.'

c ponkosa-nun-khenyeng yeypikosa-to tteleyesseyo
college entrance exam-Contr-let alone preliminary exam-even failed
'He failed in a preliminary exam, let alone a college exam.'

d kyelsung-un-khenyeng cwunkyelsung-to elyepkeyssta
final-Contr-let alone semifinal-even doubtful
'He will not make the semifinals, let alone the finals.'

Notice also that an inequality of comparison as in (23a), and a rhetorical
question as in (23b), which conveys a negative connotation throughout a sentence,
can be expressed in the K1 sentence.

(23) a onyen-khenyeng sipnyen-to te cinassta
five years-to say nothing of ten years-even more passed
'As many as ten years have passed, to say nothing of five years.'

b nwuka ne-eykey chenpwul-un-khenyeng paykpwul-ilato cwukeyss-nunya?
who you-to one thousand dollars-Contr-let alone hundred dollars-even give-
Q?
'Who would give you one hundred dollars, let alone one thousand dollars?'

Since all (a) examples in (4-7), (22) and (23) are negative affect sentences, it is tempting to say that K1, hastily extending to K2, is a syntactically negative polarity item which has the entire sentence as its affective domain. The apparent argument that K1 is a negative polarity item seems to be supported, for example, by the fact that amwuto 'anyone', a representative negative polarity item in Korean, can occur in a sentence, as illustrated in (24).

(24) Mimi-nun-khenyeng amwuto to an wasseyo
    Mimi-Contr-let alone anyone-even not came
    'No one came, let alone Mimi.'

This claim, however, turns out to be wrong when we consider K2 examples. As illustrated in the (b) examples of (4)-(7), K2 occurs without any accompanying negative form. Or, to be more exact, morphologically explicit negative morphemes must not appear in K2 examples. Compare (25a) and (25b).

(25) a chingchan-un-khenyeng kkwucwung-ul tulesssupnita
    praise-Contr-instead of scolding-Acc got
    'Instead of being praised, I got a scolding.'
    b * chingchan-un-khenyeng kkwucwung-ul an tulesssupnita
    praise-Contr-instead of scolding-Acc not got
    'Instead of being praised, I did not get a scolding.'

If K2 is also a negative polarity item, it must appear within the scope of an appropriate trigger. But there is no negative polarity trigger in (25a). Rather K2 is a negative polarity trigger with scope over the preceding clausal fragment. The evidence is as follows.

If, as we assume, K is a proposition conjunction, then the K2 sentence, such as (26a), is semantically combined from the two sentences (26b) and (26c).

(26) a sang-un khenyeng pel-ul patasseyo
    prize-Contr-instead of punishment-Acc received
    'Instead of being rewarded, I got a punishment.'
    b sang-ul mos patasseyo
    prize-Acc not received
    'I was not rewarded.'
    c pel-ul patasseyo
    punishment-Acc received
    'I got a punishment.'

Comparison between (26a) and (26b-26c) shows that khenyeng semantically corresponds to the negative morpheme. It follows that a morphologically explicit negative form cannot normally appear in the pre-khenyeng part of a sentence. If a morphologically explicit form occurs in the pre-khenyeng part, the propositional
meaning of the pre-*khenyeng part is in conflict with that of the post-*khenyeng part. Consider example (27).

(27) * chingchan-ul mos tutki-nun-khenyeng kwuwucwung-ul tulesseyo praise-Acc cannot hear-Nml-Contr let alone scolding-Acc heard 'I got a scolding, let alone not a praise.'

(27) is ungrammatical since under normal contextual circumstances, for example, 'being praised' implies 'not receiving a punishment.'

Returning now to K1 examples, I find evidence that suggests K1 to also be a negative polarity trigger. First, unless we can find any convincing argument that the forms and meanings of K1 and K2 are unrelated, diachronically or synchronically, it would not make much sense to claim that K1 is a negative polarity item, while K2 is a negative polarity trigger.

Second, if the negative polarity phenomenon in Korean is clause-bound as is generally assumed (cf. Choe (1988)), and the pre-*khenyeng part is a clause-reduced fragment separated from a full clause of the post-khenyeng part, as argued earlier, (24), repeated as (28), should be ruled out, since khenyeng occurs outside the scope of its potential negative polarity trigger, an 'not'.

(28) [Mimi-nun-khenyeng] [amwu-to an wasseyo]
    Mimi-Contr-let alone anyone-even not come
    'No one came, let alone Mimi.'

Thirdly, viewing khenyeng as a negative polarity trigger with its preceding clausal fragment in its scope predicts the variation that K1 and K2 cases differ superficially in that the full clause is negative only in the former. Furthermore, the recognition of this difference between K1 and K2 is significant since it, together with the argument that khenyeng is a conjunction, simplifies the schema of the syntax of khenyeng as in (29):

(29) [[X K] [Y]]

In the formula (29) in which K represents khenyeng, X is a variable representing a clausal fragment and Y is another variable representing a full clause, and X is always a negative polarity environment while Y is not.

IV. Semantics of khenyeng

Now that we have looked at the syntactic properties of khenyeng, we turn our attention to the semantics of khenyeng. We must first note that the semantic interpretation of the khenyeng construction ties in with the syntactic arguments that we made above, namely, khenyeng is a conjunction and a negative polarity trigger. Since khenyeng serves as a negative conjunction, the syntactic schema of (29) must be interpreted as (30).

(30) 'not X' and 'Y'
Formally, this interpretation can be represented by the semantic schema (31), which, together with the syntactic schema (29), reads as 'the sentence meaning of [X K] [Y] is the sum of the denotation of 'not X' and the denotation of 'Y'.

(31) \(-\|X\| & \|Y\|\)

It should be emphasized that the notation of \(\|X\| & \|Y\|\) is taken to represent propositions not syntactic forms, like clause or clausal fragments. For example, \(\|X\|\) is merely syntactically chosen as a fragment in order to highlight the focused element in the contextually present or assumed discourse. The argument that \(\|X\|\) is a focused element accords with the distributional fact that no other particles except for the subdued focus delimiter nun can precede khenyeng.

The suitable semantic interpretation therefore requires the interpreter to reconstruct a semantic clause from a fragment, constructing two semantic clauses. For example, the following sentences (32a) and (33a) must be reconstructed as (32b-c) and (33b-c), respectively, since one of the potential preceding contexts of (32a) and (33a) may be, for example, (34a) and (34b), respectively.

(32) a chenpwul-un-khenyeng paykpwl-to mos patasseyo
one thousand dollars-Contr-let alone one hundred dollars-even not received
'I did not receive one hundred dollars, let alone one thousand dollars.'
b chenpwul-ul mos patasseyo
one thousand dollars-Acc not received
'I did not receive one thousand dollars.'
c paykpwl-ul mos patasseyo
one hundred dollars-Acc not received.
'I did not receive one hundred dollars.'

(33) a chingchan-un-khenyeng kwucwung-ul patasssupnita
praise-Contr-instead of scolding-Acc received
'In stead of being praised, I got a scolding.'
b chingchan-ul mos patasssupnita
praise-Acc not received
'I was not praised.'
c kwucwung-ul patasssupnita
scolding-Acc received
'I got a scolding.'

(34) a chenpwul-ul patasssupnii-kka?
one thousand dollars-Acc received-Q
'Did you receive one thousand dollars?'
b chingchan-ul patasssupnii-kka?
praise-Acc received-Q
'Did you get a praise?'

It should, however, be noticed at the same time that the meaning of the whole khenyeng sentence is not merely the sum of the meanings of its conjoined propositions. For example, (32a) is more than a sum of (32b) and (32c), even though they are not different truth-conditionally. It seems obvious that the meaning
difference is attributable to the presence of *khenyeng*. Now, in order to interpret the *khenyeng* sentence, it is necessary to elucidate the semantic function of *khenyeng*.

I argue that the semantic function of *khenyeng* is to serve as a scalar operator which has the entire sentence under its scope. In other words, *khenyeng* requires the interpreter to construe two propositions as scalar such that the propositions expressed correspond to distinct points on a scale. In other words, a scale must be evoked in interpreting the semantic structure of *khenyeng* in the sense of Fillmore et al. (1988), Kay (1990), and Kay (1992). In explicating the meaning of English expression of ‘let alone’, for example, Fillmore et al. (1988) motivates the notion of scale according to whom it is defined as a set of partially ordered propositions based on the degree of intensity for certain given properties. Unlike the multiply coordinated model in the discussion of ‘let alone’, however, I have in mind a simple one-dimensional scale, since what are put in contrast in the *khenyeng* construction are pair-focused propositions, not multiply paired-focused constituents.17

Let us now make clear by way of an example what we conceive as a one-dimensional scale. Suppose that there is a directed contextual scale along which there are two distinct scalar points $P_\alpha$ and $P_\beta$, such that $P_\beta$ outranks $P_\alpha$ in informativeness, and is located farther from the origin of the scale. The horizontal representation of this scale can be made diagrammatically as follows:

\[(35)\]

\[\begin{array}{cccc}
| & | & | & | \\
0 & P_\alpha & & P_\beta \\
less \text{ informative} & & & more \text{ informative}
\end{array}\]

The scale is now interpreted if some quantity has reached the point $P_\beta$ on the scale, then it has, a fortiori, reached the point $P_\alpha$. Now returning to the semantic schema (31), two propositions, $\ll X \ll$ and $\ll Y \ll$, correspond to $P_\alpha$ and $P_\beta$ on the contextual scale since the second asserted and stronger proposition unilaterally entails the first entailed and weaker proposition. In order to see how this scale can account for the actual data, I will repeat relevant examples below.

\[(36)\] a  chenpwul-un-khenyeng paykpwl-to mos patassta
one thousand dollars-Contr-let alone one hundred dollars-even not received
'I did not receive one hundred dollars, let alone one thousand dollars.'

b  chingchan-un-khenyeng kkwucwung-lul patasssupnita
praise-Contr-instead of scolding-Acc got
'Instead of being praised, I got a scolding.'

For example, in (36a), if I did not receive $100, I certainly did not receive $1,000. Likewise, in (36b), if one has reason to believe I got a scolding, he has stronger reason to believe that I was not praised. Hence, the general semantic function of the construction is to suggest that the first proposition expressed as the clausal fragment follows from asserting the second proposition expressed as a full clause.

As Fillmore et al. (1988) indicate, an advantage of the scale is that it can represent the relative entailment of scalar propositions, and provide semantic
constraints on the acceptability of the sentence types with a scalar operator. For example, the difference between (37a) and (37b) can be accounted for by a scale.

(37a) chenpwul-un-khenyeng paykpwul-to mos patassta
one thousand dollars-Contr-let alone one hundred dollars-even not received
'I did not receive one hundred dollars, let alone one thousand dollars.'

(37b) * paykpwul-un-khenyeng chenpwul-to mos patassta
one hundred dollars-Contr-let alone one thousand dollars-even not received
'I did not receive one thousand dollars, let alone one hundred dollars.'

(37b), in contrast with (37a), is ungrammatical since it meets interpretive problems as a result of an interchange between the pair of compared propositions. In other words, (37b) violates the semantic constraint that the weaker proposition in the khenyeng construction must precede an informationally stronger proposition. Even though khenyeng itself does not determine the nature of the scale, it requires, as part of its intrinsic semantic properties, that the interpreter sets up some scalar order of the compared propositions, when interpreting the sentence in which it appears.

Having said this, it seems that what we need in explaining the semantics of the khenyeng sentence is a kind of semantic entailment relation since it explains well, for example, the relationship that "I did not receive $1,000; a fortiori, I did not receive $100". However, the khenyeng sentence involves more than a simple logical entailment relation. A reader might notice that I have put a qualification 'contextual' in the expression 'contextual scale'. I will defend why the scale must be understood as pragmatic, not semantic.

As Fillmore et al. (1988) argue, the view that the scale is contextual in nature is found to be justified in the cases in which, while the semantic entailment relation holds between the two conjoined propositions, the entire khenyeng sentence is still unacceptable. Consider (38), which is bad regardless of context:

(38) # ywuksipsa-uy seycykopkun-un i-nun-khenyeng sosswu-to anita
sixty-four-of cube root-Top two-Contr-let alone prime-even not
'The cube root of sixty-four is not prime, let alone two.'

The interpretation of (38) fails, even though 'not being a prime number' entails 'not being the number two'. The reason that (38) is pragmatically anomalous is not that the entailment relation does not hold, but that it does not hold within the same scale. In other words, (38) implies that since we have reason to believe that the cube root of sixty-four does not enter the scale, we have all the more reason to believe that the cube root of sixty-four does not reach some non-lowest point on the scale. But since two is the lowest point, the sentence is odd.

As Fillmore et al. (1988) show, the importance of the contextual scale is also manifested explicitly by the following examples in (39).

(39a) Seoul-un-khenyeng Beijing-to mos wassta
Seoul-Contr-let alone Beijing-even cannot came
'We did not pass Beijing, let alone Seoul.'

(39b) Beijing-un-khenyeng Seoul-to mos wassta
Beijing-Contr-let alone Seoul-even cannot came
'We did not pass Seoul, let alone Beijing.'

The acceptability judgment of (39a) and (39b) depends on whether the interpreter can understand the speaker's perspective or not. If the trip was being made from the west to the east, only (39b) is correct. On the other hand, if the trip was being made from the east to the west, only (39a) is correct. The conclusion that can be drawn from examples such as (38) and (39) is that what is involved in a semantic interpretation of the scalar operator *khenyeng* is a kind of special pragmatic entailment relation that presupposes a set of contextual conditions shared by the speaker and the addressee.

VI. Conclusion

The traditional explanation considers *khenyeng* as a lexical item with two different meanings, each of which occurs in a different syntactic environment. In this paper, I have demonstrated that the difference is more apparent than real since the speakers may infer the two senses of *khenyeng* on the basis of its inherent scalar properties. This leads to a unified account which treats *khenyeng* as a negative conjunction and a scalar operator. On this account, *khenyeng* is a single lexical item, rather than two semantically unrelated homophonous lexical items.

Interestingly enough, there are also several other expressions similar to *khenyeng* in Korean. These include *mal hal nawiepsi 'needless to say', mwullon 'needless to say', kosahako 'apart from, let alone, needless to say', ppwun mananila 'not only-but also', and hamwulmye ... ilya 'much more/less, not to mention, let alone', etc.

Time and space preclude a discussion of these operators here. But the study of each of these scalar operators, preferably in comparison with *khenyeng*, will shed light on the issues that might have been neglected in this exclusive study of *khenyeng*, and will give a more encompassing and complete explanation of scalar operators in general.

---

1 I am grateful to Charles Fillmore, David Peterson, and especially Paul Kay for valuable comments on a previous version of this paper. All shortcomings that remain are, of course, my responsibility.
2 The definition of delimiter will be provided below.
3 Space does not allow me to investigate the pragmatics of *khenyeng* in this paper.
4 I mark space between a fragment and *khenyeng* to highlight the structure of the *khenyeng* sentence. In actual orthography, *khenyeng* is attached to the preceding fragment and no space is allowed between them. Abbreviations made in the paper include Contr 'Contrast', Nom 'Nominative', Nml 'Nominalizer', Acc 'Accusative', Pst 'Past', Top 'Topic', Q 'Question ending' and Inf 'Infinitive'.
5 The somewhat awkward 'auxiliary particle' and its congener 'delimiter' are due to lack of a suitable corresponding term in English.
6 For example, see Yang (1972). However, the term delimiter does not show any substantial difference from and is no better than the traditional one, auxiliary particle, since the meaning of a noun phrase, for example, is not 'limited' by its immediately following delimiter. Yet in the remainder of the paper I will continue to use the term delimiter since a better term does not at present suggest itself to me.
Choy (1929/1961) actually presents a third sense as well. However, this meaning is not available in current Korean. Following is one of the instances he cites:

(i) onyen-khenyeng sipnyen-i cinassta
    five years-to say nothing of ten years-Nom passed
    'Ten years have passed, to say nothing of five years.'

I am not sure why there is a discrepancy in acceptability between Choy and current Korean. It may be that the third sense of khenyeng which Choy points out disappeared in the course of time (in about half a century) or his observation may be influenced by the Japanese equivalent of khenyeng, dokoroka, which displays exactly these three senses. Incidentally, (i) becomes acceptable when a comparative adverb is attached to the verb, about which more will follow in section 3.2. Now compare (i) and (ii).

(ii) onyen-khenyeng sipnyen-to te cinassta
    five years-to say nothing of ten years-even more passed
    'As many as ten years have passed, to say nothing of five years.'

As (1) and (2) show, the examples Choy (1929/1961) provides lack the subdued focus delimiter nun before khenyeng. Korean dictionaries (say, Hankulhakhoy (1991)) say that nun comes in when the speaker wants to emphasize the word that precedes khenyeng. The idea is that nun is optional in this construction. However, this remark is incorrect since the delimiter nun must be present, for example, after the adverbiai or the postpositional phrases. Moreover, the presence of khenyeng sounds more natural even after the noun phrases, as in (1) or (2). It would be therefore preferable to say that nun drops in a very restricted syntactic environment, say, after the noun phrases. I will not go into more detail about nun since consideration of the issue involved is not relevant to the points I will make.

Verbal nouns are considered as a kind of verb phrase in this paper in order to give a consistent explanation of verb phrases.

Yang made a mistake in classifying kkaci 'to', pwuthe 'from' as X-lim. According to this classification, for example, the string kkaci-pwuthe 'even from' should be ruled out, since both elements of the string belong to X-lim, but it is okay. I will not go into more detail about this issue, since it is not related to any of my points.

Hong (1983) gives a different classification of delimiters, but the delimiter nun still occupies the last slot in a possible string of delimiters.

One may argue that khenyeng is a delimiter which comes in the last fourth position in an arrangement of delimiters, proposing another syntactic delimiter slot. I will reject the idea simply because khenyeng will then be the only delimiter that occupies the position. Proposing another slot for the delimiters is not economical for the overall organization of the grammar. But this idea might not be taken as a strong argument.

The rest of the negative evidence was already hinted at in the foregoing examples which show khenyeng usually following the delimiter nun. Thus, I omit a detailed explanation.

All the (b) examples of (13)-(14) are acceptable when the delimiter nun is inserted before khenyeng. This observation might suggest a possibility that nun-khenyeng is a compound delimiter in the making. The fact that nothing can intervene between nun and
khenyeng also seems to suggest its compound nature. However, there are independent positive properties of the khenyeng construction which are directly indicative of conjunction, as will be discussed momentarily.

15 However, the other canonical coordination tests - clause conjunction, gapping, and conjunction reduction - fail to apply because a khenyeng sentence cannot have more than one-paired focus.

16 In fact, Lee (1992) assumes, without any arguments, that khenyeng is a negative polarity item.

17 For a formal characterization of a more complex scale, see Fillmore et al. (1988).

References


Im, Hong-Pin, and Ik-Sep Lee (1983). Kwuke Mwunpeplon [Korean Grammar], Seoul, Hakyensa.


0. Introduction. Recent linguistic approaches to discourse (Webber:1979, Kamp:1981, Heim:1982) set out to unravel the behavior of anaphoric elements in a text. All of these frameworks attack the problem of anaphora by introducing discourse markers or discourse referents (following Kartunen:1976) as mediators between linguistic entities (NP's) and their model theoretical interpretation. Accordingly, all of the above frameworks concentrate on how to representation pronominal anaphoric relations, as in:

(1) Ed would like to resolve the situation one way or another before he leaves office...

In this paper I argue that, to give an adequate description of anaphora, one must consider another (equally common and important) type of anaphoric expression as well:

(2) The company opened its procurement office here last fall (another one, its first, is in San Francisco)...

I show that the anaphoric relations between the italicized elements in (1) and (2) exemplify two, markedly different types of anaphora.

Intuitively, the difference between the two anaphoric types consists in whether the anaphoric element in question reactivates (a) the discourse referent/token (as with pronominal anaphora in ex.1) introduced by the nominal expression construed as the antecedent or (b) the descriptive content/type/sense introduced by that nominal expression (as with one-anaphora in ex.2). I will call these two types of anaphoric relations Referential and Descriptive anaphora, respectively.

To empirically set the two types apart, I will use the test of modal subordination (see Roberts:1986 following Kartunen:1976). Given the distinction between the two types of anaphora, I propose a substantial modification of standard Discourse Representation Theory (see Kamp & Reyle:1993) to accommodate one-anaphora into that framework. For reasons that become clearer later, I will call the modified DRT structure a Residual Discourse Representation Structure, or RDRS. The overall moral of the paper, however, states that (no matter what framework we consider) all anaphoric expressions cannot be treated uniformly, as the different behavior of pronominal anaphora and one-anaphora demonstrates.
1. Reactivating a Discourse Referent v. Reactivating a Description. Let us then turn to the distinction that underlies the difference between the two types of anaphora.

1.1. The Ambiguity. Consider the following examples:

(3) Bill brought a ham sandwich with pickles for lunch today.
   a. ...It was the same sandwich he brought yesterday. /same token/
   b. ...Joe brought the same sandwich as well. /same kind/

(4) Bill drives the Honda over there.
   a. ...His wife drives another car, which is also a Honda. /different token/
   b. ...His wife doesn’t like Hondas, and drives another car. /different kind/

The sentences in (3)-(4) exemplify a systematic ambiguity. Namely, they show that a single linguistic expression (with the exact same phrase as antecedent) can appear with two distinct readings, depending on its context. In the (a) sentences the italicized phrase acts as anaphoric on the token introduced by its construed antecedent, while in the (b) sentences it is anaphoric on the kind introduced by that antecedent. For example, given the continuation in (3a), Bill is assumed to have brought the same token sandwich for lunch today as he did yesterday, let’s say because he did not have time to eat it yesterday. With the continuation in (3b), however, pragmatic factors tell us that Joe’s and Bill’s sandwiches are different tokens of the same kind of sandwich. A similar ambiguity holds for the two continuations in (4), whether we are talking about a different token or just a different kind from the car that Bill drives.¹

The two readings in examples (3)-(4) therefore arise depending on whether the antecedent is taken to be the discourse referent or simply the description introduced by the antecedent phrase of the anaphoric expression. If it is the discourse referent (token) introduced by the antecedent phrase that acts as the antecedent, I call the anaphoric relation REFERENTIAL ANAPHORA, while if the description (type or kind) acts as the antecedent, I call the anaphoric relation DESCRIPTATIONAL ANAPHORA. The most important point here is that the distinction between the two types of anaphora (that is, Referential v. Descriptive) cannot always be attributed to a given linguistic expression.² Although one-anaphora will always involve anaphoricity on kind, expressions such as the same sandwich or another car can act either as Referentially anaphoric, or as Descriptively anaphoric.

1.2. Comparison with Other Anaphoric Distinctions. The Referential v. Descriptive distinction made in this paper closely resembles a couple of
earlier proposals. In particular, it is parallel to the distinction between identity-of-reference v. identity-of-sense anaphora proposed by Grinder & Postal (1971), as well as the distinction between concrete entity v. concept anaphora advocated by Asher (1993). The most well-known distinction among different anaphoric types, however, is the distinction between deep and surface anaphora put forth by Hankamer & Sag (1976) (which was later recast in Sag & Hankamer:1984 as a distinction between model interpretive anaphora and ellipsis). I only consider the last of these three distinctions in any detail here.

Sag & Hankamer (1984) summarize the following three tests to distinguish deep anaphora from surface anaphora:

(a) only deep anaphora can be used deictically, can be ‘pragmatically controlled’;

(b) only surface anaphora requires parallelism in syntactic form between anaphor and antecedent;

(c) only surface anaphora exhibit the ‘missing antecedent’ phenomenon. In S&H’s system pronominal anaphora and sentential *it* are instances of deep anaphora, while VP-ellipsis, gapping, sluicing, and stripping instantiate the ellipsis type of anaphora.

Assuming the tests above, *one*-anaphora exhibits certain properties of both types of anaphora. First of all, *one* can be used deictically (pragmatically controlled), as the following example shows, which is assumed to be characteristic of model interpretive anaphora:

(5)  [...walking up to a balloon vendor...]  
Could I have a purple *one*?  

On the other hand, *one*-anaphora seems to require some sort of syntactic parallelism (or semantic contrast) between the antecedent and the anaphor, a property which supposedly characterizes ellipsis type anaphora:

(6)  I like the blue *balloon* more than the yellow *one*.
(7)  ?? Our *dog* mated with a black *one*.

It has been noted that so-called paycheck pronouns require syntactic parallelism, and it is easy to see as well that *one*-anaphora patterns with VP-anaphora in that they both involve reactivation of descriptive material. Given these facts and arguments (showing *one*-anaphora as well as pronominal anaphora as sharing certain properties of both deep and surface anaphora), I take that the Referential v. Descriptive anaphoric distinction is orthogonal to the deep v. surface anaphoric distinction.
2. **One-anaphora and modal subordination.** By now we have established the intuitive distinction in the nature of the anaphoric link for the cases of one-anaphora and pronominal anaphora. It is therefore time that we turned to some formal differences between the two anaphoric types. The crucial test I present here involves modal subordination phenomena.\(^3\)

2.1. **What is modal subordination?** Data pointing to effects of modal subordination was first introduced in Kartunen (1976) under the terminology of short term discourse referents, which was later taken up and generalized by Roberts (1986) within the DRT framework under the terminology of modal subordination. From the point of view of anaphora, modal subordination simply stands for the effect that modal contexts exercise on the licitness of anaphoric links. In particular, if a (non-specific) indefinite antecedent is introduced in a non-factual context (e.g., within the scope of a modal operator, or negation), it is inaccessible for a subsequent pronominal anaphor, unless the anaphor appears in a sentence that is a modal continuation of (i.e., is modally subordinated to) the sentence in which the antecedent appeared. To be more precise, consider the examples in (8) and (9):

(8)  
\[\begin{align*}
\text{a. } & \text{If Joe bought a book, he'll be home reading it by now.} \\
\text{b. } & \text{It'll be a murder mystery.}
\end{align*}\]

(9)  
\[\begin{align*}
\text{a. } & \text{If Joe bought a book, he'll be home reading it by now.} \\
\text{b. } & \text{#It’s a murder mystery.}
\end{align*}\]

In (8) (which is Roberts’ ex.4) a book is introduced within the scope of a conditional operator. (8b) is interpreted as continuing the modal context induced by if, thus antecedent clause of (8a) is part of the common ground (or presupposed material) for (8b).

On the other hand, when the second sentence introduces factual mood, as in (9) (which is Roberts’ ex.3), the picture looks different. Even though the conditional itself (9a) is presupposed (part of the common ground), its antecedent clause is not presupposed. Therefore the existence of a book is not presupposed (not part of the common ground for ex.9b), and pronominal anaphoric reference to it is illicit.

Similar facts hold for negation: If a (non-specific) indefinite (such as a power plant in ex.10) is introduced in the scope of negation, it is unavailable as an antecedent for a pronominal anaphor from a subsequent declarative sentence as well, because the existence of a power plant is not presupposed:

(10)  
Mike has never seen a power plant. He is visiting #it tomorrow.
To sum up, examples (8)-(10) therefore show that pronominal anaphora exhibits modal subordination effects.

2.2. Lack of modal subordination effects. As it turns out, one-anaphora does not show the same modal subordination effects as pronominal anaphora does. Let us see the evidence that demonstrates this point. Consider (11) and (12), which parallel (9) and (10), respectively:

(11) If Joe really bought a car today, he is driving to Yosemite now. He went to the dealer to buy one this morning.

(12) Mike has never seen a power plant. He is visiting one tomorrow.

In these examples the antecedent of one is introduced either within the scope of a conditional operator or within the scope of negation, while the sentence containing the anaphor one is not modally subordinated to the first sentence. Still, the anaphoric link between the intended antecedent and anaphor is licit. In other words, one-anaphora does not show modal subordination effects.  

3. DRT representation of pronominal anaphora. This takes us to the main question of the paper, namely how anaphoric relations should be represented, given the different behavior of pronominal anaphora and one-anaphora. Here I will take standard Discourse Representation Theory as my starting point, as presented in Kamp & Reyle (1993) and Roberts (1986). In Section 4. I offer a modification of their theory to accommodate one-anaphora. But first we have to review quickly how pronominal anaphora is generally represented in DRT.

The main innovation of Kartunen (1976), Webber (1979), and other discourse oriented approaches such as Kamp’s (1981) Discourse Representation Theory (DRT) and Heim’s (1982) File Change Semantics (FCS) consists in their treatment of nominal expressions. Nominals (such as indefinites, pronouns, etc.) are no longer taken to refer directly to real world entities or entities in a model; instead, they are assumed to introduce variables into the representation, so-called discourse referents. These discourse referents can be interpreted in mapping them into a model of the actual world, as it is done in DRT or FCS. Under such a view of discourse anaphoricity is expressed as a condition expressing the identity of the values of two variables. That is, for a mini-discourse such as (13) (after Heim:1982), we get the DRT representation below:
(13) Otto owns a sheep. Harry vaccinates it.

\[
\begin{array}{|c|}
\hline
x, y, z \\
\hline
\text{owns}(x,y) \\
\text{Otto}(x) \\
\text{sheep}(y) \\
\text{vaccinate}(z,v) \\
\text{Harry}(z) \\
z = y \\
\hline
\end{array}
\]

In the DRT representation discourse referents are introduced in the ‘universe’ of the Discourse Representation Structure (DRS), at the top of the representation. At the same time, predicates attributed to the discourse referents are listed in the body of the DRS as conditions. The last condition of the DRS is taken to express that the pronoun it refers back to the sheep Otto owns, by explicitly stating the identity of the two variables y and z, which were introduced as discourse referents for a sheep and it.\textsuperscript{5}

The other property of pronominal anaphora relevant for us is that of exhibiting modal subordination effects, as discussed in Section 2.1. In DRT, following Roberts (1986), modal contexts and negation are represented by subDRS’s, which are embedded into the main DRS representing the discourse. SubDRS’s come with their own universe, and the relevant operator can be affixed to the subDRS as presented here (modal subordination is expressed with the double line between the subDRS’s):

(14) If Joe bought a book, he’ll be home reading it by now. #It’s a murder mystery.

\[
\begin{array}{|c|}
\hline
x, r \\
\hline
y \\
\text{John}(x) \\
\text{book}(y) \\
\text{bought}(x,y) \\
\hline
\end{array}
\quad
\begin{array}{|c|}
\hline
z, w \\
\text{reading}(z,w) \\
z = x \\
w = y \\
\hline
\end{array}
\]

\text{murder mystery}(r) \\
r = ?
\]
Mike has never seen a power plant. He is visiting it tomorrow.

As shown, indefinites and pronouns introduce discourse referents in the universe of their own subDRS (in which they are introduced), while proper names introduce a discourse referent in the universe of the main DRS.

Given the above representations let us look at the intended anaphoric links. In (14) (which is Roberts' ex.9) the anaphoric reference \( w=y \) is licit, since the subDRS where \( w \) is introduced is subordinate to the subDRS where \( y \) is introduced. However, the universe of the antecedent clause is not accessible to \( r \), since \( r \) was introduced in the main DRS, which is not subordinate to the subDRS expressing the antecedent clause. This is why the intended reference \( r=y \) fails.

For (15), the intended anaphoric reference \( v=y \) is illicit because the discourse referent \( y \) is introduced into the universe of the embedded DRS, and therefore it is not accessible from the main DRS (where \( v \) is introduced), which is not subordinate to the subDRS in (15).

4. DRT and one-anaphora. Accepting the above as the proper treatment of pronominal anaphora, let us turn to the question how one-anaphora should be represented in the DRT framework.

4.1. Descriptions as property variables? First of all, recall the preliminary observation of Section 1.1. stating that one-anaphora reactivates a description, as opposed to pronominal anaphora, which reactivates an (individual) discourse referent. How can we express this intuitive difference in the formal representation?

We certainly cannot assume that one takes the condition power plant\( (x) \) as its antecedent.
(16) Mike has never seen a power plant. He is visiting one tomorrow.

\[
\begin{array}{|c|c|}
\hline
x, z, v \\
\hline
\hline
\text{see}(x,y) & \text{visit}(z,v) \\
\text{Mike}(x) & z=x \\
\text{power plant}(y) & \text{one}(v) \\
\text{one}=? \\
\hline
\end{array}
\]

The condition \textit{power plant}(x) is not a predicate, therefore it cannot be applied to \( v \) as a function. Adding the condition \( v=y \) would not help either, since it would lead to an incorrect interpretation of the discourse.

Now suppose that one-anaphora also involves an anaphoric link between some kind of discourse referents. These discourse referents, obviously, could not be of the type ‘individual’, rather they would need to be property variables introduced alongside individual variables.\(^7\) Then the representation for (12) in DRT terminology would look approximately as below:

(17) Mike has never seen a power plant. He is visiting one tomorrow.

\[
\begin{array}{|c|c|}
\hline
x, z, o, v \\
\hline
\hline
\text{y, p} \\
\hline
\hline
\text{see}(x,y) & \text{visit}(z,v) \\
\text{Mike}(x) & z=x \\
\text{P}(y) & \text{Q}(v) \\
\text{P}=\text{power plant} & \text{Q}=? \\
\hline
\end{array}
\]

Under this proposal there are two discourse referents associated with a power plant: the individual variable \( y \), and the property variable \( P \). Supposedly, the property variable would be introduced within the universe of the (sub)DRS alongside discourse referents for individuals. The anaphoric \textit{one} itself would stand for a property variable as well, and the anaphoric link would be expressed as a condition on the identity of the values of two property variables, just like in the case of pronominal anaphora and individual discourse referents.

Given the representation in (17), the problem is already obvious. In the second sentence \textit{one} licitly refers back to the description power plant. However, this description was introduced within the scope of negation, and thus the property variable was introduced in the universe of the embedded subDRS. Therefore, according to the general constraints on the accessibility
of antecedents, the anaphor *one* could not licitly access this (property) discourse referent as its antecedent, the intended $Q=P$ anaphoric reference is predicted to fail, even though in reality it is perfectly acceptable. The question remains, how should we represent *one*-anaphora then?

4.2. **One-anaphora and Residual DRS’s: The Proposal.** To remedy the problems mentioned in the previous section, I suggest the following modification of standard DRT. Since descriptions are available as antecedents across all modal contexts, there is no reason to include them within the modally structured DRS’s. Instead, let us assume that descriptions are introduced on a representational level parallel to the level containing the discourse referents in the DRS’s. Let us call these two levels Descriptive Tier (DT) and Referential Tier (RT), respectively. The specific proposal I make here is only partially motivated by the behavior of *one*-anaphora, hence the proposed representation might seem somewhat ad hoc. This representation nevertheless proves useful in explaining various phenomena such as quantification, distributivity, and specificity. Therefore from here on, to illustrate the full force of the proposal, I will include eventuality variables in the upcoming representations. Disregarding them in the paper so far, however, did not affect any of the arguments made.

Let us now see the actual proposed representation:

(18) Mike has never seen a power plant.

a. He is visiting #it tomorrow. [=10=15]

b. He is visiting *one* tomorrow.  [=12]

<table>
<thead>
<tr>
<th>DT -- List of functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>[le(seeing e)]1; [lx(called-Mike x)]2; [lx(power plant x)]3; [le(seeing e)]4; [one]5=[lx(power plant x)]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RT -- Residual DRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>x2, f4, z, v5</td>
</tr>
<tr>
<td>e1, y3</td>
</tr>
<tr>
<td>Agent(e,x) Patient(e,y)</td>
</tr>
<tr>
<td>Agent(f,z) Patient(f,v)</td>
</tr>
<tr>
<td>z=x v=z</td>
</tr>
</tbody>
</table>

The RT contains a kind of DRS that looks much like a standard DRS in that it contains discourse referents in a modally structured representation of the discourse. On the other hand, it looks very different from standard DRS’s in terms of what it contains as conditions in the body of the DRS. Instead of the usual predicates, it contains a neo-Davidsonian representation which associates individual variables with the eventuality variable, by designating them as
Agent, Patient, etc. Since this DRS is stripped of the conditions that are normally represented in it, I will call this a Residual DRS, or RDRS for short. In interpreting the representation, elements of the RT will carry existential claim (even if within the scope of a modal operator) in the same fashion as in standard DRT. This is the representational level which is responsible for quantification, distributivity, and specificity effects. And most importantly for us, (Referential) pronominal anaphora takes place on this level as well; again, observing the same conditions and constraints as in standard DRT.

The DT of the representation then contains a list of $\lambda$-extracted versions of the predicates which are normally listed as conditions in the body of a standard DRS. Elements on this tier therefore do not express any existential claim, nor are they arranged in a modally structured representation. Given that the RDRS now includes eventuality variables as well, the DT will also include eventuality predicates, $\lambda$-extracted of course, which specify the type of eventualities mentioned in the given discourse. This level of the representation will be the site of one-anaphora (and Descriptive anaphora in general), where the only condition on accessibility is ordering in the list.\(^9\)

The two levels of the representation are linked via a co-indexing mechanism. The coindexing is interpreted as function application. That is, coindexing a $\lambda$-extracted predicate from the DT and a variable from the RT boils down to ordinary predication after $\lambda$-conversion, and the effective interpretation of the representation is equivalent to that of standard DRT. The two representations, however, are not functionally equivalent, since the present proposal offers an unique organization of the representation which allows for a formal representation of one-anaphora, an anaphoric relation that standard DRT cannot handle.\(^10\) Therefore by reorganizing the representation we do not lose anything, but gain an explanation for an extended class of anaphoric phenomena.

Finally let us see how an RDRS representation derives the desired results, and solves the problems that arose in connection with one-anaphora. In the representation above I combined the two representation for (10) and (12), where the underlined material indicates the two intended anaphoric relations for (18a) and (18b).

The discourse referent $v$ still cannot access $y$ as its antecedent, since accessibility conditions for discourse referents within an RDRS are the same as for a standard DRS. Thus we derive modal subordination effects for pronominal anaphora. On the other hand, anaphors on the DT behave differently. Since the DT simply contains a list of functions without intervening modal structure, one is free to access $\lambda x$(power plant $x$) as its antecedent. That is, with the representation in (18) we derive the fact that one-anaphora is exempt from modal subordination effects, while maintaining an account of the modal subordination effects for pronominal anaphora.
5. Conclusion. In this paper I presented evidence from the behavior of one-anaphora to show two points. First, given the different behavior of one-anaphora and pronominal anaphora we are forced to conclude that no uniform treatment of anaphoric elements is possible, irrespective of what framework we consider. Second, I showed how standard DRT can be modified to extend its scope to one-anaphora, without losing any of the results for the treatment of pronominal anaphora.

*I want to thank Donka Farkas, Ray Jackendoff, Christopher Kennedy, Ewan Klein, Jason Merchant, Mariano Plotkin, and the audience at BLS for comments on different versions of this paper. For all remaining mistakes and problems I take sole responsibility.

1 For an indefinite NP to be anaphoric on an antecedent the anaphoric relation has to be a contra-indexing relation, as opposed to a co-indexing one. For more arguments that contra-indexing relations indeed show the characteristics of anaphoric relations see Csúri (1995).

2 If the anaphoric expression alone specified whether the anaphoric relation was Referential or Descriptive, the resolution of anaphoric expressions would be one degree less complicated.

3 In addition to modal subordination, however, there are other tests that distinguish Referential v. Descriptive anaphora, involving the presence or lack of E-type and bound variable readings. For a discussion of these tests see Csúri (in prep.).

4 Here I argue that the modal structure of the discourse does not constrain one-anaphora in any way. For a suggestion that modality might pose certain restrictions on one see Merchant (1994).

5 Note that such a representation produces the E-type reading of a pronoun discussed by Evans (1980) and others.

6 This argument is a simpler version of Klein’s (1987) argument as to why VP-ellipsis cannot simply be represented with a property variable.

7 A proposal along similar lines has been made by Hardt (1993). Using Groenendijk and Stokhof’s (1991) Dynamic Predicate Logic framework he proposed that VP-ellipsis be represented as anaphora on a property variable.

8 For details on additional motivation for, and further advantages of, the RDRS representation please stay tuned for Csúri (in prep.).
In Heim’s File Change Semantics the DT would be equivalent to a single file card which lists all the functions that have been mentioned in the discourse. That way expressions anaphoric on functions would have unlimited access to antecedents.

Klein (1987) offers a DRT analysis of VP-ellipsis, which Asher (1993) extends to other types of concept anaphora. For some problems with these analyses see Csúri (1995) and Csúri (in prep.).

REFERENCES

Merchant, Jason. 1994. ‘(In)definite ‘one’: DC anaphoric dependency and novelty,’ ms. UCSC.
The interaction of the binding principles and the Chinese reflexive taziji

Jun Da
University of Texas at Austin

1. Introduction

The basic approach of the Binding Theory, as set forth in Chomsky (1980, 1981 & 1986) is that index is first arbitrarily applied to NP's with the results then being checked at surface structures to see whether certain conditions are satisfied. While it has little problem to account for clear-cut cases of anaphors and pronominals, the theory turns out to have left open two complicating issues. The first one has to do with grammatical sentences which can violate the binding principles at the same time. The other is concerned with nominals such as the English PRO which can be subcategorized as both an anaphor and pronoun at the same time. Even though the PRO Theorem provides the licensing condition for its distribution, i.e., ungoverned positions, it is silent on when PRO should or should not be bound and if it does, how its antecedent is located.

The goal of this paper is to propose a tentative Optimality-theoretic solution to the above two problems with a special reference to the Chinese reflexive taziji. Language specifically, I will show that the binding of the Chinese reflexive is decided not by any single principle alone but the interaction of a set of binding conditions. In particular, I will show that given a grammatical input, we can rely on a hierarchy of binding constraints to predict when the reflexive should and should not be bound and in the latter case select an antecedent for the reflexive. Theoretically, this paper is intended as a first attempt to explore the advantage of applying a constraint-based approach to the study of binding relations.

2. The binding of taziji: Basic facts

The Chinese reflexive taziji is made up of two morphemes: the third person singular pronoun ta (he/she) and the bare reflexive ziji (self). Both forms of taziji share the same pronunciation, though they differ in their written forms. (In this paper, taziji will be used in the masculine sense unless otherwise noted.) In the literature, taziji has been considered either a free variant (Li and Thompson 1981) or a surface representation of the Chinese bare reflexive ziji (cf. Tang 1989). Whatever its relation to ziji, however, it suffices for us to focus on taziji here, since, as has been assumed in the literature, binding relations are checked at surface structures. For more information on the bare reflexive ziji as well as its relation to taziji, please refer to Li & Thompson (1981), Wang & Stillings (1990), Battistella & Xu (1990), Tang (1989), Huang & Tang (1991), Xu (1993 & 1994) and Pan (1995).

We now turn to the binding properties of taziji. In most cases, the reflexive will take a local antecedent, as shown in (1) and (2). In both sentences, it occupies the complement position of the verb xihuan (like) and has to be co-indexed with the sentential subject John and pengyou (friend) respectively:

(1) John xihuan taziji,

like he-self.

John likes himself.
(2) \[IP \{NP_1 [NP_2 John_{de}] pengyou_{j} [VP xihuan [NP_3 taziji_{i/j}]\} \]
\[John's \ friend \ like \ himself.\]

When the above simple sentences are embedded within other sentences, the reflexive will display the same kind of binding property, as shown in (3):

(3) \[IP_1 John_{i} [VP_1 suo [IP_2 Bill_{j} [VP_2 xihuan taziji_{i/j}]]].\]
\[say \ like \ he-self.\]
\[John \ says \ that \ Bill \ likes \ himself.\]

In (3), taziji is bound to the immediately available subject Bill of IP₂, even though another subject John in the higher IP₁ can serve as its potential antecedent (as we will see later).

Similar to its locality nature at the verb complement position, taziji will behave as a local anaphor even when it occupies either the subject position as in (4) or the Spec position of an NP as in (5):

(4) John_{i} suo [IP₁ Bill_{j} yiwei [IP₂ taziji_{i/j} bu hui zuo chunjuan]].
\[say \ think \ he-self \ no \ able \ make \ egg \ rolls.\]
\[John \ said \ that \ Bill \ thought \ that \ he-self \ was \ not \ able \ to \ make \ egg \ rolls.\]

(5) John_{i} bu xiangxin [IP₁ Bill_{j} hui suo [IP₂ [NP taziji_{i/j}/*k de taitai]k]
\[no \ believe \ will \ say \ he-self \ 's \ wife \ pretty \]
\[piaoliang]].\]
\[John \ doesn't \ believe \ that \ Bill \ would \ say \ that \ his \ wife \ is \ pretty.\]

In (4), there is no subject other than taziji within IP₂ and the reflexive finds its antecedent Bill in the higher IP₁. Within IP₂, a co-indexing between taziji and itself will constitute a violation of the so-called i-within-i condition (Chomsky 1981). In (5), the same filter disallows the co-indexing between taziji and the sentential subject taitai(wife). Instead, the reflexive has to be bound to the sentential subject Bill in the higher IP₁.

While the i-within-i condition must be observed all the time, the c-commanding condition, in contrast, turns out to be violable in the binding of taziji. Even though all the above examples satisfy the c-commanding condition, it is not the case with the following sentence (Pan 1995):

(6) \[IP \{NP John_{de} taidu_{j} [VP hail-le taziji_{i/j}]\} \]
\[John's \ attitude \ hurt-ASP \ himself.\]

In (6), John is at the Spec position of the NP (which serves as the subject of the sentence) and functions as an antecedent for taziji, and yet it does not c-command the reflexive.

The c-commanding condition, however, is not the only requirement that the binding of taziji will sometimes fail to meet if the reflexive is taken as a local anaphor. Recent evidence from Yu (1992) and Pan (1995) indicates that taziji can also bear the properties of a long-distance anaphor and hence creating an overall
violation of Principle A. (7a) and (8a) are two such examples ((7a) is (10) in Yu 1992. (8a) is (204b) in Pan 1995.):

(7) a. [IP1 Johni suo [IP2 Maryj bu xihuan ta-MAS-zijij/*j]].
    say not like he-self.
    John says that Mary doesn't like himself.
    b. [IP1 Johni suo [IP2 Maryj bu xihuan tai/*j/k]].
    say not like he.
    John says that Mary doesn't like him.

(8) a. [IP1 Johni yao [IP2 woj zuo zai [NP tazijij/*j de shenbiam]]].
    want I sit at he-self's side
    John wants me to sit at his side.
    b. [IP1 Johni yao [IP2 woj zuo zai [NP tai/*j-de shenbiam]]].
    want I sit at he's side.
    John wants me to sit at his side.

In (7a), the masculine taziji is governed by the verb xihuan (like). The subject of the lower IP2, Mary, is thus an accessible subject. However, taziji has to choose instead the subject John in the higher IP1 as its antecedent. This is in contrast to the pronominal ta (him) in (7b) which can optionally be free in the whole clause. In (8a), taziji occupies the Spec position of an NP in IP2 and has to be co-indexed again with the matrix subject John in the higher IP1. This is similar to the possessive pronoun ta-de (his) occupying the same position as in (8b). Whatever positions it occupies, both (7a) and (8a) demonstrate that taziji is no longer a pure local anaphor. Rather, it acts as a long-distance one which must be bound to an antecedent outside the minimal governing category IP2. Within the lower IP2, however, it can be optionally free.

Despite its anaphoric orientation, there are other cases in which taziji can be either A- or A'-free, as the following three examples illustrate ((9) is (13a) and (10) is (12) in Yu 1992, respectively. (11) is from my own informants):

(9) ni j wen taziji*i/*j.
    you ask he-self.
    You ask him(self).

(10) taziji zengmo suo?
    he-self how say
    What (or how) does he say?

(11) taziji shenghuo shifen jianku dan hai shi leli bangzhu taren.
    he-self life very hard but still be willing help others
    He(-self), though living a hard life, is still willing to help others.

In (9), taziji occupies the complement position of the verb wen (ask) and yet there is no qualified antecedent to which it can be bound, since the sentential subject ni (you) does not agree with taziji in person. In (10) taziji occupies the subject position and there are simply no nominals in the sentence which can serve as its potential binder. In (11), taziji functions as the topic of the whole sentence and occupies an A'-position. No antecedent is available for taziji, either. In all the three cases, taziji is free.

To sum up, we have seen that taziji displays a diversity of binding properties: It can be locally bound as a pure anaphor, though c-commanding
condition is not a strict requirement on its binding; it can be long-distance bound when no qualified local candidate is found within the minimal governing category; it can even be A- or A'-free when no qualified antecedent is available in the whole sentence. These properties are a clear reflection of the two complicating issues mentioned in the introduction. In the following section, I will review briefly how the binding of taziji has been accounted for in the literature.

3. Previous analyses

Past insights into the nature of taziji fall into two schools. One school treats the reflexive as a pure local anaphor subject to Principle A of the Binding Theory (Wang & Stillings 1990, Battistella & Xu 1990 and Huang & Tang 1991). For example, Battistella & Xu (1990) propose that:

(12) Taziji is bound to the closest accessible subject.

While it has no problem in dealing with the local anaphoric reading of taziji, the above principle is apparently unable to account for the other properties of the reflexive as shown in (6) to (11). For example, in (6) John, which occupies the Spec position of the subject NP, does not c-command the reflexive, even though the reflexive has to be bound to it. A further technical difficulty in their account is that in the relevant literature, it is not clear how the distance between the reflexive and its antecedent (for the measurement of closeness) is formally evaluated.

Problems in the above treatment were first reported by Yu (1992), though he provided no alternative solution. A comprehensive treatment came recently from Pan (1995) in which an attempt was made to formalize the concept of closest accessible subject.

Pan's insights are mainly based on the observation that whenever possible, the local anaphoric reading always takes over the long-distance reading. The latter will avail only when the local subject does not agree with taziji in phi-features. These ideas are formalized in his Compatibility and Closeness Conditions:

(13) The Compatibility Condition
α and β are compatible iff
a. α and β have compatible animacy features; and
b. α and β are syntactically, semantically and pragmatically compatible.

(14) The Closeness Condition
α is closer to X, the reflexive, than β iff the path from X to the minimal maximal projection dominating α is a subset of the path from X to the minimal maximal projection dominating β.

In essence, Pan's Compatibility Condition is no more than a restatement of the common assumption in the literature that a co-indexed pair should be interpretable. The Closeness Condition further specifies that the distance between the two should be kept minimal. With these two explicit conditions, Pan's Principle A is formulated as follows:

(15) Principle A
An anaphor must be bound to the closest compatible candidate, where a candidate for an anaphor is a noun phrase that does not dominate the anaphor.
When it is applied to (3), this principle is equivalent to Chomsky's (1981) Principle A in its judgment of the sentence. In (3), the path from John to tazi is [IP1, VP1, IP2, VP2], whereas that from Bill to tazi is [IP2, VP2] which is a subset of the former. In terms of Pan's Closeness Condition, Bill is closer to tazi than John and qualified as its antecedent. Note that Bill is within the minimal governing category and c-commands the reflexive.

Without the c-commanding condition, this principle has again the same effect as Chomsky's Principle A which predicts that in sentences such as (6) tazi will select John rather than anything else as its antecedent. In (6) we find that the subject NP taidu (attitude) cannot be an antecedent of tazi, because their co-indexing would otherwise violate Pan's Compatibility Condition on animacy agreement. Note that John, which occupies the Spec position of the sentential subject NP, is the only animate NP in the utterance, even though it does not c-command the reflexive.

Though Pan's theory caters for more binding properties of tazi, there are still two unsatisfactory aspects in his account of the reflexive. First, it is unclear how the concept of compatibility can be evaluated. For example, the exclusion of other phi-features such as gender, person and number in his Compatibility Condition fails to accommodate for the long-distance anaphoric reading of tazi as shown in (7a). According to Pan's Compatibility Condition, (7a) would be judged ungrammatical, since Mary does agree with ta-MAS-zaji in animacy and hence is a compatible antecedent for tazi. So is John in the same sentence. Given two compatible nominals as candidates for antecedency, his theory predicts that the closer one, i.e., Mary, would be the antecedent, though it is not true in this case.

The second problem in Pan's account is related to the exclusion of c-commanding condition from his principle. Even though we have seen that the condition is not a strict requirement on a potential antecedent for tazi, it nonetheless plays a role in certain situations such as the following:

\[(16) \text{John yiwei [IP1 Bill [VP1 PP dui Davidk] suo [IP2 tazi ijji/*k/*k]} \text{thinks to say he-self} \]
\[\text{[VP2 xihuan Sam]]}. \]
\text{like}
\text{John thought that Bill said to David that he liked Sam.}

In (16), the antecedent Bill c-commands tazi, whereas neither David nor Sam does. However, Bill is as distant as David and further away than Sam from the reflexive. It is unclear how Pan's Principle A would evaluate this since no specific c-commanding condition is included.

It is clear from the above brief review that the diversity of the binding properties of tazi creates a dilemma in the conventional application of the Binding Theory. We have seen that principles such as the c-commanding condition hold at one time but are violable at another. We have also seen that a competition for closeness is evident in the selection of an antecedent for the reflexive. The interaction and competition among those principles suggest the need for a new mechanism to accommodate for them. In this sense, the insights of the Optimality Theory may point to an alternative solution.
4. A constraint-based perspective of the binding theory

Optimality Theory (Prince & Smolensky 1993) is a model of constraints and constraint interaction on output representations. In OT, a grammar is a system of ranked, violable universal constraints. In formal terms, it consists of two functions: First, the function GEN maps an input representation, through some optional structure-building operations, into a (possibly infinite) set of candidate outputs. These candidates are then fed into the function EVAL for evaluation. EVAL in turn contains a hierarchy of constraints which rates parallely each member of the candidate set. The most optimal candidate, i.e., the one with least violations of the constraints, is selected as the well-formed or grammatical.

Apparently there is a parallel between the conception of the Optimality Theory and the basic approach to the Binding Theory. It is therefore beneficial for us to try to adopt the insights of the Optimality Theory to the study of binding relations. In this section I discuss in general terms how the notions and principles of the Binding Theory can be converted as constraints in the Optimality-theoretic framework.

The first step in accessing binding relations, i.e., the co-indexing between a nominal and its potential antecedent, can be defined as an optional operation in GEN: It assigns a binding relation to each pair of nominals in an input sentence. In formal terms, this will be a partial function. For example, given a set of NP's such as \{John, himself\} from an input sentence John likes himself, it will output a set of ordered pairs as in (17):

(17) GEN for binding: (Optional)
{John, himself} → \{<John, himself>, <himself, John>, <John, John>,
<himself, himself>, John, himself, ...\}

Note that in (17) the output of GEN is represented as a set of ordered pairs (as used in logic). Unlike its usage in phonology, the triangular bracket notation is used here to denote a binding relation: The first element in the ordered pair is to be considered a candidate antecedent (which usually occupies a higher position in the phrase structure, whether dominating or preceding) with which the second element is to be co-indexed. Further, an element outside an ordered pair in the candidate set is to be considered free.

With the set of candidate outputs, the next step is to determine which member of the set is the most optimal, i.e., well-formed. This is where the various binding notions and principles are brought into play.

To understand how they interact with each other, we need to take a look at the classification of nominals. In the literature, a nominal is classified in terms of two primary features, i.e., [anaphor] and [pronominal], with either a positive or negative value. Such a framework of nominal classification actually implies a basic assumption: The former must have an antecedent somewhere, whether implicitly or explicitly, while there is no such requirement for the latter. (We will call the former referentials and the latter non-referentials hereafter.) This distinction provides the basic conditions in a constraint-based approach to the study of binding relations:

(18) Binding Conditions
a. BIND: <..., α>  α, being a referential, must be present as the second element in an ordered pair.
b. *BIND: *<..., α>  α, being a non-referential, must not be the second element in an ordered pair.
In the above set, BIND requires that a referential such as a reflexive must be present in an ordered pair as its second element. *BIND, in contrast, forbids a non-referential nominal such as an R-expression to be present at the same position. The workings of this set of constraints are quite similar to that of the Faithfulness Conditions in phonology which seeks a correspondence between input and output. When the two constraints either dominate or being dominated by other constraints, we will find the diversity of the binding properties for such nominals as the Chinese reflexive taziji.

We now turn to the relation between a referential and its potential antecedent as defined in the standard Binding Theory. On the one hand, it is expected that there must be a nominal to which a referential in question can be bound. The basic requirement, as has been assumed in the literature, is that it does not conflict with the referential in phi-features. On the other hand, it is expected that the antecedent of an anaphor should c-command the referential and their co-indexing should not violate the so-called the \textit{i-within-i filter} (Chomsky 1981). The three atomic conditions are put together in (19) as Accessibility Conditions:

\begin{enumerate}
  \item Accessibility Conditions
  \begin{enumerate}
    \item $\phi$-feature agreement: \textit{*(αF, βF)*}. Two co-indexed nominals must have no conflicting $\phi$-features.
    \item i-within-i: \textit{*(β... α...)*}, where α and β also form $\langle β, α \rangle$.
    \item c-commanding: The antecedent must c-command the referential.
  \end{enumerate}
\end{enumerate}

With these atomic conditions, both Principle A and B can in fact be taken as a requirement on the distance between the antecedent and a referential in terms of minimal governing category. On the one hand, Principle A requires that the antecedent be the closest qualified nominal (to an referential) which satisfies the above atomic conditions. On the other hand, Principle B excludes this nominal from being a candidate, even though it does not specify where a legal candidate should appear.

With the above outline, we now turn to the discussion of the Chinese reflexive taziji to see how those ideas can be applied to account for its binding properties.

\textbf{5. Optimal binding of taziji}

In Section 2 and 3, I have shown that the reflexive can behave as a local anaphor. This kind of binding property can be captured with the first two atomic conditions of (19), namely, phi-feature agreement (19a) and \textit{i-within-i filter} (19b). (c-commanding condition is to be discussed later in this section.). These two constraints require that an optimal co-indexed pair such as $\langle \text{NP}_1, \text{NP}_2 \rangle$ should agree in their phi-features and that NP\textsubscript{2} is neither identical to nor embedded within NP\textsubscript{1}. A violation will be registered for any phi-feature disagreement or embedding relation between the two nominals. Since these two conditions are the basis for an meaningful binding relation, they are to be placed at the top of the constraint hierarchy we are now building, even though ranking between the two is not necessary.

Further, we need the constraint BIND as in (18a), since taziji is by nature a referential. This constraint requires that a nominal be co-indexed with another one as its antecedent. It will register a violation for a nominal outside an ordered pair (in our current notation). As for its ranking, it is to appear lower than the two
accessibility conditions, namely, Accessibility >> BIND. Such a ranking amounts to say that if a nominal is to be co-indexed with another one and their binding be well-formed, the minimal requirement is that their co-indexing be interpretable.

These three constraints can handle the selection of a A- or A'-free tajiji as seen in (9) in which no compatible antecedent for the reflexive is found in the whole sentence. Tableau 1 demonstrates this evaluation.

    Tableau 1

<table>
<thead>
<tr>
<th>Candidates</th>
<th>*&lt;αF, βF&gt;</th>
<th>*[β... α ...]</th>
<th>BIND</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. &lt;ni, taziji&gt;</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. &lt;taziji, taziji&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. taziji</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In the above tableau, even though the A-free taziji violates BIND, it is nonetheless more optimal than the other two co-indexed pairs. In comparison, <ni, taziji>, in which the two elements do not agree in person, violates the higher *<αF, βF>. <taziji, taziji>, on the other hand, creates a violation of the i-within-i condition. This is exact the common assumption in the literature that an anaphor need not be bound if there is no nominal which COULD serve as an accessible subject.

The same constraint hierarchy is also good for the selection of a local antecedent for taziji as seen in (1) in which there is only one qualified potential antecedent in the sentence whose configuration satisfies the standard Principle A. Tableau 2 demonstrates the relevant evaluation:

    Tableau 2

<table>
<thead>
<tr>
<th>Candidates</th>
<th>*&lt;αF, βF&gt;</th>
<th>BIND</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. taziji</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. &lt;John, taziji&gt;</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As the above evaluation indicates, even though both candidates do not violate the higher *<αF, βF>, <John, taziji> outperforms the unbound taziji which is outside an ordered pair, since the latter constitutes a violation of the lower BIND. Compared with Tableau 1, this evaluation demonstrates how an optimal output in one utterance can become less optimal in another.

The competition for the best may also lead to the long-distance reading of taziji, as we have seen in cases such as (7a) in which John, the subject in the higher clause, is the antecedent to which the reflexive has to be bound. Tableau 3 demonstrates this competition:

    Tableau 3

<table>
<thead>
<tr>
<th>Candidates</th>
<th>*&lt;αF, βF&gt;</th>
<th>BIND</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. taziji</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. &lt;John, taziji&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. &lt;Mary, taziji&gt;</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In Tableau 3, <John, taziji> emerges as the winner, since it violates none of the two constraints. In comparison, <Mary, taziji> loses because of the two elements' disagreement in gender. The defeat of the unbound taziji comes from its failure of having an binder and hence constituting a violation of BIND.

The above set of constraints, however, is not sufficient for the selection of a locally bound taziji when there are other potential compatible antecedents available in a sentence. In (3), for example, both John and Bill can serve as an antecedent for the reflexive, though taziji will choose the latter as its actual antecedent. In Battistella & Xu (1990) and Pan (1995), this has been referred to as the closeness
effect. The same insight will also be adopted here as a constraint, though it is worded in a slightly different way:

(20) Minimal Distance Condition (MinDist)
Assign a * to each intervening dominating maximal projection between \( \alpha \) and \( \beta \).

Technically, the number of maximal projections is counted in the way as defined in Pan's Closeness Condition (cf. (14)). This constraint says that the distance between a referential and its antecedent, measured by the number of maximal projections, should be kept minimal if at all possible. It will register a violation for each intervening maximal projection. As for its ranking, it is to appear after BIND. Otherwise, an unbound *taziji* will always emerge as the most optimal candidate since the constraint applies to it vacuously. The following tableau demonstrates how this constraint interacts with previous ones such as BIND in the selection of an antecedent for *taziji* in (3):

<table>
<thead>
<tr>
<th>Tableau 4</th>
<th>Candidates</th>
<th>BIND</th>
<th>MinDist</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. <em>taziji</em></td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. <em>&lt;John, taziji&gt;</em></td>
<td></td>
<td>****!</td>
<td></td>
</tr>
<tr>
<td><em>c. &lt;Bill, taziji&gt;</em></td>
<td></td>
<td>**</td>
<td></td>
</tr>
</tbody>
</table>

In Tableau 4, *<Bill, taziji>* is the most optimal, since the path between the two elements is \([IP_2, VP_2]\) which creates only two violations of the Minimal Distance Constraint. In comparison, the path between the two elements in *<John, taziji>* is \([IP_1, VP_1, IP_2, VP_2]\) which constitutes four violations of the same constraint. Further, the unbound *taziji* violates BIND and hence is ruled out as an optimal output.

The necessity of the constraint is also evident in the account of (2) in which, if we ignore the c-commanding condition for a moment, Pan's Closeness Condition produces the same effect as Chomsky's (1981) Principle A. We note that in (2) both *John* and *pengyou (friend)* can function as a possible antecedent for *taziji*. However, the latter is closer to the former, since the path from *pengyou (friend)* to *taziji* is \(\{IP, VP\}\) which is a subset of the path from *John* to *taziji* \(\{NP_1, IP, VP\}\). The following tableau demonstrates how the sentential subject *pengyou (friend)* instead of *John* is selected as its antecedent in (2):

<table>
<thead>
<tr>
<th>Tableau 5</th>
<th>Candidates</th>
<th>BIND</th>
<th>MinDist</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. <em>taziji</em></td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. <em>&lt;John, taziji&gt;</em></td>
<td></td>
<td>***!</td>
<td></td>
</tr>
<tr>
<td><em>c. &lt;pengyou, taziji&gt;</em></td>
<td></td>
<td>**</td>
<td></td>
</tr>
</tbody>
</table>

As can be seen in the above tableau, the evaluation works just like that in Tableau 4 in which anaphoric resolution is worked out through the Minimal Distance Condition.

Note that in the above sentence, the best choice happens to c-command the reflexive. Even though Minimal Distance Condition can account for the grammaticality of the sentence without imposing a c-commanding condition on the antecedent, it does not imply that c-commanding condition is never needed, as we have already seen in (16). (In (16) we have noted that it is a c-commanding nominal *Bill* which functions as the antecedent for *taziji*, even though in terms of Pan's
Closeness Condition, it is as distant from the reflexive as David in the PP and Sam is even closer to the reflexive.) That being the case, the c-commanding condition as in (19c) is to be included in our hierarchy. As for its ranking, the case of Sam vs Bill in (16) suggests that it is to be placed higher than the Minimal Distance Condition, whereas the case of an unbound taziji indicates that it should be placed lower than BIND. Otherwise, an unbound taziji will always be a winner. The following tableau demonstrates how it interacts with other constraints in the evaluation of (16):

<table>
<thead>
<tr>
<th>Candidates</th>
<th>BIND</th>
<th>c-commanding</th>
<th>MinDist</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. taziji</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. &lt;John, taziji&gt;</td>
<td>*</td>
<td>******!</td>
<td></td>
</tr>
<tr>
<td>c. &lt;David, taziji&gt;</td>
<td>*</td>
<td>***</td>
<td></td>
</tr>
<tr>
<td>d. &lt;Sam, taziji&gt;</td>
<td>*</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>e. &lt;Bill, taziji&gt;</td>
<td></td>
<td>***</td>
<td></td>
</tr>
</tbody>
</table>

In the above tableau, we find that <Bill, taziji> emerges as the most optimal. In comparison, <John, taziji> is ruled out as an optimal output since the pair has four violations of the Minimal Distance Condition. The other two ordered pairs, <David, taziji> and <Sam, taziji>, lose their chances simply because they constitute a violation of the c-commanding condition, even though David is as distant from the reflexive as Bill and Sam even closer. The story of the unbound taziji is a familiar one.

To conclude this section, the constraint hierarchy we have proposed is listed in (21):

\[(21) \quad \star \alpha F, \beta F, \star [\beta \ldots \alpha \ldots] \gg \text{BIND} \gg \text{c-commanding} \gg \text{MinDist}\]

6. Concluding remarks

In this paper, I started by pointing out two embedded problems in the conventional approach to the Binding Theory, namely, the violation of binding principles and the multiple binding properties of a nominal. I illustrated the two problems with the Chinese reflexive taziji in which I showed that its binding is decided not by any single rule but the interaction of a set of conditions. To solve these problems, an Optimality-theoretic approach to the Binding Theory was proposed. In particular, I argued that the assignment of binding relations can be related to GEN and the binding principles can be interpreted as constraints used by EVAL. With this new perspective, I showed that a properly ranked constraint hierarchy as in (21) can account for the complex binding properties of the Chinese reflexive when it appears in a grammatical input sentence.

There are remarks to be made here before we conclude this paper: First, note that I have assumed that for evaluation, comparison is made among surface representations of grammatical input only. I did not discuss why and how an ungrammatical sentence such as John says that I like himself is excluded from consideration. I leave this question open for further research. Secondly, the Minimal Distance Constraint is proposed here with a gradient nature (rather than discrete). The validity of this proposal needs further empirical verification. Finally, it should be pointed out that even though the two issues addressed in this study are common across many languages, I have not had any chances, due to the size of the paper, to examine similar reflexives in other systems. Future research is needed to verify the validity of this approach by applying it to a larger body of empirical data.
References


Ethnolinguistic Loyalties among Barcelona’s Teens
Hope N. Doyle
University of Colorado at Denver

Introduction
1. Owing to years of repression of the Catalan language and culture in Catalonia, Spain, and the demographic shift in the autonomous region, wherein, at the end of the Franco dictatorship, nearly half of the population of Barcelona was born elsewhere in Spain, a Catalan normalization campaign was formally initiated in 1983 with the Law of Linguistic Normalization. Students have since been required to take one Catalan course a year and content courses in the language. Cultural activities in the Catalan language (TV, theater, film, radio, etc.) have been subsidized, and with this official promotion has come the resurgence of national holidays which heighten and elaborate Catalan national loyalties.

The group most concretely affected by the normalization campaign has been teens, who have studied Catalan throughout all of their school careers. According to the 1991 Census, they have the highest proficiency level of all age groups. A significant number of youths are first generation Catalan, having one or both parents born elsewhere in Spain. The ‘identity planning’ experienced by these youths has often juxtaposed Catalan and Spanish cultural and linguistic loyalties. Where do their linguistic loyalties lie? Are Catalan and Spanish identities inherently conflictive? What is the configuration of this bicultural and bilingual identity?

This paper discusses some of the results of a lengthy sociolinguistic survey of 420 public secondary school students carried out in the city of Barcelona during the spring of 1991. The questionnaire included four sections: socio-demographic, Catalan language proficiency, habitual language usage in a variety of domains and among a variety of interlocutors, attitudes about the two languages, cultures and identities (with both open and closed answer questions). The design of the survey was based on the work in ethnolinguistic vitality and identity of Fishman (various studies), Bourhis et al. (1981), Allard and Landry (1984), and Russinovich Sóle (1990).

The results of the survey indicate that, among first generation Catalan teens, Catalan is most frequently used in instrumental exchanges, among unknown or socially distant interlocutors. The language is accorded high instrumental and integrative value and its use is considered by all subjects as the clearest expression of Catalan identity. The subjects prefer bilingualism to Catalan (or Spanish) monolingualism and appear to integrate both Spanish and Catalan identities with less conflict than had been predicted by many sociolinguists, though this ethnolinguistic configuration is not a simple one. They also see the need for continued official promotion of the language and culture, since they continue to perceive a threat (posed by the significant presence of the Spanish language in Barcelona) to the life of the language.

Ethnolinguistic loyalties and identity: responses and analysis.
2. Subjects were asked: How would you characterize the Catalan identity? (The following representative responses, and all italicized script herein, represent my own translation from either Spanish or Catalan.)

*[Being Catalan means] speaking Catalan, following Catalan customs and being a fan of the Barça [soccer team].

*Well, among [the characteristics that define being Catalan], being Catalan means not being Spanish, this is something that a lot of people just don’t understand.
*People who like the language, the people, the Catalan world, and not [who don't like] the Castellano world.

*There isn't anything that defines [Catalan identity], one is [Catalan] because he was born here and feels Catalan.

*Being Catalan implies living in Catalonia, not only having been born here, but watching TV in Catalan as well as in Castellano, being able to speak in both languages, etc.

Discerning who is "Catalan" and who is "Castilian" is something that speakers in Barcelona must do every day, and, as several subjects in my sample report, mistakes are sometimes made. Determining cultural identity is crucial with respect to bilingualism in the Catalan capital, since it has been shown to be primarily interlocutor-specific with respect to choosing between Catalan and Castilian. (Casamigia/Tusón, 1980)

A common thread in the discussion of language in the city is the issue of identity. It is of particular interest in the case of Barcelona teens because they are in the process of solidifying their social identities. These adolescents are at the same time the principle target of Catalanization campaigns. A prominent Catalan sociolinguist, Miguel Strubell (1984:100) uses the somewhat orwellian expression "identity planning" in referring to a process he calls "tempting to promote," but pending "empirical data demonstrating a direct causal relationship between the two variables - language use and national identity." This idea was repugnant to Catalans during both the Franco and Primo de Rivera dictatorships, when the identity planned was the Spanish one, not to mention throughout the 19th century, as the concept of the nation-state spread throughout Europe.

There has been quite a bit of planning with respect to national identity in Catalonia; the linguistic normalization law of 1983 in itself addresses this issue, as it is intended as a framework for disseminating not only the Catalan language, but also culture, identity. The (autonomous Catalan) state has taken an active role in the "heightening" and "elaboration" of the beliefs, attitudes, and behaviors of particularly the young in Barcelona.

Fishman points out that for "societally organized goal activity" (incrementing the use of Catalan at more domains, in the case at hand) to occur, people must become convinced of their communality and ethnocultural uniqueness (1972:5).

Many subjects of each of the parentage groups in the present study are inspired by Catalan nationalism. They are aware of Catalonia's great achievements and achievers. It seems that Catalan history has been sufficiently authenticated for the respondents. The normalization campaign has been thorough in this regard, helping at once to unify the population and to link the past to the present.

A host of factors are at work with regard to language and nationalism; I explored the nature of the relationship between the two with special consideration of the palpable effects of the 'identity planning carried out so far.

One example of this kind of official activity was the Bishop of Barcelona's call for the immigrant population of Barcelona to "Catalanize" itself, repeated by Parish fathers in Sunday masses in August and September of 1990. This call spurred a debate throughout the city, at dinner tables, on the metro. A large number of letters to the editor on this subject appeared in the (Castilian language) Catalan daily Vanguardia in the ensuing weeks.

Another such activity was the (ultimately failed) proposal declared in the Fall of 1990 by Minister Guitart, Ministry of Culture for Catalonia, that all children
begin their school day with the song, Els Segadors (The Reapers). This hymn commemorates the War of Spanish Succession, in which Catalonia sided with the Austrian Archduke Charles III, who seemed to better assure its political and economic ambitions (Vilar I:354), against the Bourbon crown (centered in Madrid), and was defeated in 1714. (The Catalans had had to fight valiantly primarily because Austria had abandoned them when the Bourbon victory seemed imminent.)

Many Catalan national celebrations have successfully reemerged nonetheless. One thousand years of Catalan culture was recently celebrated, (marking Count Berenguer's alleged break with the crown of Charlemagne and, after a fashion, the independence of Catalonia.) It seems that the more remote in chronological time the event, the deeper the unifying effect.

Not all of the symbols which have been more recently formulated, or historical events reframed, have been immediately accepted by the population of Barcelona. The Catalan national identity emerges, formulated and guided by the autonomous government's active role in the process. Despite Fishman's assertion that, "there are integrative 'times and tides'" (1972:22), the integrative moments, unlike tides, are concretely related to specific promotional attempts and concurrent socio-economic factors.

The Catalanization campaign itself appears to have inadvertently deepened the Spanish-Catalan cultural divide for some subjects, highlighting precisely those who are not Catalan, in order to encourage them to be so. (Consider the Bishop's call for Catalanization.)

Respondents' feelings about national and cultural identity were elicited by questions that asked them to say how they referred to their place of origin when visiting other parts of Spain and abroad. Subjects were also asked if it was possible to consider oneself Catalan without speaking Catalan, and if it was possible to be Catalan without having been born in Catalonia. Then, in an open-ended format, they were asked to describe the characteristics of being Catalan.

A general, unleashing, open-ended question asked subjects to define the characteristics of Catalan identity. The coded results are seen in Table 1.

The most frequent response for all samples referred to language as the primary factor in Catalan identity:

*Speaking Catalan is the feature that most defines being Catalan, as one respondent said, though the degree of command of the language required (to be considered Catalan) varied in responses. The following respondent stressed understanding the language:

*Being able to understand the language perfectly and being able to speak it too. Other respondents expressed the need to practice the language frequently in defense of Catalan identity and as a way of averting extinction of the language:

*Speaking Catalan and doing everything possible to slow-down [the process of] its extinction.

Another said:

*Respecting the Catalan customs and language and practicing them frequently. Several respondents expressed feelings, sentiments with respect to the language, more than simply knowing it, but identifying with it:

*Feeling identified with the language, the culture and the Catalan landscape. Language is given primary importance in the following response, but it is coupled with other feelings of nationalism and cultural uniqueness:

*A specific language, a specific history, a specific sense of nationalism, being a unique people.
Catalan identity was also perceived in a more practical sense with respect to language:

*The feature that most characterizes being Catalan is the accent, since, even if they are speaking Castilian, you can tell.*

Accent is often the most difficult aspect of language to acquire, and many subjects expressed the feeling that their Catalan pronunciation was less than perfect. From this perspective, being considered Catalan could seem impossible for many.

This view with regard to language may be contrasted to another infinitely more inclusive one:

*Not being bothered by hearing Catalan.*

A fruitful comparison may be made between tables 2 and 3. While only one in ten second-generation immigrants said that it was impossible to be Catalan without speaking the Catalan language (table 2), nearly a third found it to be the primary determining factor when asked what the defining characteristics are for being Catalan. With respect to the mixed parentage group, the same comparison may be made: one in five felt that not speaking the language would disqualify one from Catalan identity (table 2), but more than a third gave the language primary importance with respect to identity. Many others saw the language as one of a number of defining characteristics including customs, residency, birth, etc.

The Catalan inclusiveness campaigns of the last decade may have contributed to the students' formulation of what defines Catalan identity, since respondents are less likely to hypothetically disqualify one's Catalan identity because of lack of language use.

*All those who live and work in Barcelona are Catalan* - a slogan I had seen around the city during my 1990-91 stay, and had previously seen cited in Woolard's (1988) commentaries on the subject, is an apt summation of these campaigns. As a cognitive response to more open-ended questions, the subjects were more likely to cite language as a defining quality of Catalan identity.

A third of the entire sample mentioned language as the primary determining feature of Catalan identity. Attitudes with regard to the normalization of the Catalan language would elucidate the relationship between language and identity in Catalonia. In other words, what subjects would consider a 'normal' situation for the language determines their relative inclusion in the Catalan identity if they regard language as the primary feature. For example, those that understand a 'normal' linguistic situation as one that would be bilingual co-equally may feel Catalan, though they do not often speak the language. Conversely, those that conceive of 'normal' as a situation in which Barcelona would be first and foremost monolingually Catalan, would not conceive of themselves as Catalan if they did not speak Catalan most of the time. The variety of responses with respect to language promotion and normalization shows that there is clearly no consensus on this issue. Arriving at a consensus will determine the future direction of language legislation (if not of the linguistic complexion for the city itself).

Many subjects pointed to sentiment in response to the question 'What are the characteristics that define being Catalan?':

*A person can be Catalan without speaking Catalan, since there are many reasons why a person might not have been able to learn the language. But I think that if a person doesn't make an attempt to internalize another culture, he can't identify with it.*

Another responded:
*A profound feeling of love for the land, knowing that you have your own identity, knowing that you have something: a feeling very difficult to understand if you are not from here, its your most valued possession.*

This feeling is expressed as a connection to Catalonia as a land, as love for Catalonia, as something only barely definable. Sentiment highlights the affective component of identity. For one in four respondents, affection for Catalonia was deemed more a defining characteristic of the Catalan identity than residence or particular traits. Expressed as such, this feature of identity is a rather inclusive one; regardless of background, one may be Catalan simply by feeling that one is such. But some respondents further point out that to be recognized as Catalan, it is necessary to give voice to this sentiment, to express it, and to thereby "defend Catalonia." Fishman points out:

[H Nationalist symbols] heighten awareness that are only latent, so that not only will [people] come to feel that they constitute a nationality but that they will also be willing to act upon the basis of that feeling. (1972: 15) (Underline own)

To many subjects who highlight sentiment as an important factor in Catalan identity, the language inspires a desire to defend Catalonia, since it was so brutally repressed in the post-Civil War period. They also mention this idea in discussing bilingualism in general; because the language was another of Franco's victims, it literally requires 'affirmative action' and begs reassertion. "Nationalism is a phoenix that is repeatedly capable of arising not only out of its own ashes but out of whatever other injustices modern societies perpetrate." (Fishman, 1972: 29)

Several subjects advised me to carefully consider the huge influx of immigrants into Barcelona during the post-war period in my analysis of bilingualism there. They expressed the sense that bilingualism in Barcelona was the result primarily of Franco's policies and of immigration, as if before the war the city had been Catalan-monolingual. However, the desire for language normalization is almost an integral element of Catalan culture, beginning in the 18th century, heightened in the 19th with romantic fervor, before the war during the Second Republic, etc. Castilian had permeated Catalonia in the 15th century and has played a decisive role there ever since. One respondent characterized the Catalan identity in the following way:

*We would like to once again become the great nation we once were.*

Another response expressed this same sentiment:

*[Being Catalan means] having a consciousness of our unique historical past.*

Several subjects illustrated a juxtaposition of loyalties and sentiments with respect to Catalan and Castilian identity. Some of the statements that open this chapter reflect this: to be Catalan is to be a fan of the Barça professional soccer team (not Real Madrid!); another expressed the idea that a Catalan loves his land and his language and not Spain.

This view was also widely contradicted by a more harmonious idea of being both Catalan and Spanish, being of both cultures, an outlook also expressed at the opening of the paper. The subject felt that being Catalan did not mean *not* being Spanish, but that this was something that many people "could not get into their heads." Indeed, some of his classmates clearly stated that being Catalan meant not being Spanish, or being Catalan first, then Spanish.

It seems that, to some degree, this sentiment has been influenced by a reframing of history. The immigrants are a more palpable target of blame for the dominant status of Castilian in Catalonia, yet the nation's own past indicates that the language had a threatening hold on Barcelona long before the newcomers arrived en masse.
The Catalans have an expression, 'fer pais,' (literally, to make country), in referring to the acts of spending time with family and following Catalan traditions. The expression highlights the sense that Catalonia is in need of making, in the process of being made- that there is a component of the national identity which relies on a continual, conscious effort to make or maintain itself. 'Fer pais' illustrates the component of action that sentiment must infuse in order to maintain (or acquire) the Catalan identity, and speaking Catalan is a large part of that action. *Catalan, aside from the stereotypes that tend to deceive, like in the case of the Spanish, it is a culture with historical, linguistic and social differences, and defining Catalan-ness would require using stereotypes.*

One in five subjects (one in four Catalan ones) used specific traits, stereotypes, to describe being Catalan, as this subject points out. I initially coded the answers into a single 'personality traits' category. Given that the subsample was comprised of a significant number, 67 respondents, further analysis of the stereotypes themselves was due.

Adam Schaff, in an essay on the pragmatic function of stereotypes, affirms that the analysis of the cognitive and emotional aspects of stereotypes is "dangerous to many seemingly noble and proud clichés." It is for this reason that: defensive mechanisms are set in motion to prevent such researches; such mechanisms work as psychological smoke-screens which conceal unpleasant realities. (1984: 94)

According to Schaff, this paradigm primarily applies to the mechanism of 'cognitive dissonance': a cognitive and emotional function in the individual that allows us to hold firmly to deeply felt beliefs, even when confronted with evidence contrary to those beliefs, a 'knowing and not knowing.' Schaff further points out that while cognitive dissonance "serves to protect stereotypes against inquisitive and trouble-making analyses," stereotypes at the same time have "an immense role in making that mechanism effective" (1984: 94).

The pragmatic function of stereotypes can be seen clearly in Barcelona, as many subjects in the study (regardless of parentage) used such stereotypes to define being Catalan. This is so because not all stereotypes are negative.

It is useful to determine the symbols that underlie the stereotypes discussed, and their role in shaping and reflecting sociolinguistic relationships. Thus, stereotypes are important to analyze inasmuch as they help reveal the nature of the sociolinguistic relationship between groups. For stereotypes not only reveal historically conditioned feelings, but they themselves help link the past to the present and are an integral part of social identity.

The subjects freely responded to the open-ended question: what are the characteristics that define being Catalan? As has been shown, four in five respondents did not use stereotypes, personal traits, descriptive adjectives, to define being Catalan, but rather language, residence, and sentiments. The stereotype response was somewhat unexpected since this type of answer did not appear in the pilot survey. The pattern of responses encompassed the traits listed in table 4.

These responses bear some relation to those found in the Ros (1988) and Woolard (1985) studies and to well-known stereotypes. The most frequent of all responses (32.8%) were those that described competence, refinement, education, and elegance. Drawing on Brown and Gilman's (1960) distinction between power and solidarity (explored in their study with respect to pronouns of address), Woolard (1985) extended the notion with regard to attitudes toward both Castilian and Catalan speakers among teens in Barcelona. Using the matched-guise technique with a semantic differential scale in the measurement of status or power
on one scale, and solidarity on another, she finds a somewhat non-reciprocal power relationship (Catalan vis-a-vis Castilian). Woolard (1985) confirms: "there is no experimental evidence that Catalan is a low prestige language; in fact, there is noteworthy indication that Catalan gives higher status to speakers than does Castilian" (1985:104). In light of Woolard's findings, these responses (competent and refined) may be considered symbols of power. The data from my sample suggests the existence of this power dynamic, since one in three of those who used personality traits to describe Catalans judged them as competent and educated.

The trait labeled 'unsharing' was expressed by one in five who used stereotypes to express Catalan identity. In some cases, subjects related this trait to the success of the Catalans in their endeavors, and to account for their financial success.

As previously discussed, a fundamental component of Catalan identity is cultural struggle. One respondent remarked with respect to the Catalan language: *Aside from being my own language, it represents a struggle, an ideal of distinction, a unique identity. Something that I have in my heart which will never die and about which I'm very proud.*

There is official promotion of the sense that Catalan-ness is intertwined with such struggles for survival. (The most important national celebration remembers a monumental defeat: La Diada.) To some degree, this notion propagates the idea that in order to survive, the Catalan people have had to cling to what is most dear. Thus, the perception of some that they have clung too tightly. José Ferrater Mora, in a book on the Catalan personality, gives voice to this stereotype:

*In their great majority, the Catalans don't wait for things to fall from the sky, like a manna or a gift from the gods; according to the well known, and often poorly interpreted adage, the Catalans can get blood (bread) from a stone.* (1987: 123)

This quality also expresses the instrumentality of Catalan - suggesting an integrative bond based on economics (Fishman 1972:7). The status and power of Catalan was also expressed by a respondent who said that to him the Catalan language meant:

*Knowing another language, Being able to get jobs in Catalonia, and perhaps that people treat you better than if you speak Spanish. The superior social status of Catalan was expressed by several respondents: *I consider Catalan to be more elegant and modern than Castilian.*

*Being Catalan means being more cultivated and refined with respect to people from the south of Spain, because its proximity to Africa and not Europe.* These sentiments are not uncommon in Barcelona. Indeed, they bear some relation to some of the cultural promotion campaigns of the Generalitat, the autonomous government that I witnessed during my 1990-91 stay, reminding drivers on highways and streets throughout Barcelona that Catalonia is a part of Europe, with slogans on billboards such as:

"Europe Doesn't End at the Pirennies" and
"Catalonia, a European Country".

Another sign was placed in subway stations throughout the city during my stay there:

Work well done has no boundaries.
Work poorly done has no future.

This saying tends to highlight that which is "well done," "in good taste" as constitutive elements in Catalan identity. This well-known stereotype is also highlighted by Ferrater Mora:
To use a very picturesque Catalan expression, Catalans can spend many hours 'dreaming about tortillas' (in vague, yet innocuous things), but the least one may ask is that the "tortillas" are well made. (1987: 124)

Of course, striving for this image does not make Catalonia unique; the rest of Spain (and other poorer nations of the European Economic Community) are aware of their need to compete with the rest of Europe in the pursuit of 'quality', greater productivity, etc. The autonomous government has made pragmatic use of this stereotype in order to attain standards that are being explicitly demanded of all members of the European Community.

The stereotypes that express such traits as openness, friendliness, and exclusiveness may be distinguished from those that are related to power, though both features share a component of status. More than one in four respondents used 'open' and 'friendly' to define Catalans, as opposed to one in ten who expressed the opposite: 'closed'. These are the two poles of the continuum which could be used to measure solidarity, as in Woolard (1985). Since these traits were not placed on a continuum, there was no middle ground expressed. That is, only subjects with strong feelings one way or another responded to the question in this way. The remaining 80% of the subjects are not included in the present analysis and may feel solidarity to varying degrees. There is a substantial degree of solidarity with Catalans among the subjects of the subsample since they mentioned qualities such as openness and friendliness, expressing an integrative bond with the in-group which may or may not be born of ethnocultural similarity.

Another trait of Catalan identity mentioned in the subsample (20.8%) was nationalism, pro-independence, patriotism, what Ros (1988) categorized as expressions of national and cultural identity. According to these respondents, those who voice nationalist sentiments give voice to their Catalan identity. Schaff points out:

[T]he less people realize that stereotypes exist and affect human actions, the greater the role of such stereotypes. People say, for instance, 'love of one's country', 'patriotism', [...] but they not only use such formulations, but much more important still- they act accordingly. (1984: 99)

This is not to suggest that this subsample is unique in its reference to stereotypes. Indeed, as Festinger (1957) and Schaff (1984) contend, stereotypes and the cognitive dissonance involved in their creation are natural mechanisms which defend profoundly felt beliefs and ideas from inconvenient information (that which provides evidence to the contrary) or unpleasant realities. In other words, none of us may escape them. A priori knowledge about an outgroup makes that group more familiar. This study has focused on Catalan identity, and does not include stereotypes of other communities, but many commonly felt stereotypes about the people of the differing regions of Spain have a great deal of content. In a sense, they evoke historical experience, one which involves "historically conditioned feelings of hostility" (Schaff, 1984: 96). These feelings resurge on national holidays, such as the Diada. Stereotypes are transmitted just as other historical information is transmitted and elaborated from generation to generation. Note one subject's feelings on the Catalan language:

*Catalan is my language, I've always spoken it and its something that I care a lot about (although that may sound corny). I care about it even more when I study what the Castilians did with our language (prohibit it) that infuriates me and it makes me be more against Castilian.*
Conclusion

3. It is clear that Catalonia seeks to define itself (as all nations continually do) vis-a-vis two important larger elements: namely, Spain (seen in the symbol of Madrid) and Europe. The campaigns of the autonomous government and the views of many subjects expressed the notion that while Spain is seen as a subtractive element in Catalan identity, having always posed a threat to it (the Diada symbolizes that feeling), Europe is seen as the larger element which might legitimize Catalonia, not subtract from it. The emerging European Community is a symbol of hope for Catalonia, since it is supposed to represent unity in variety, not to mention business opportunities. Spain has long represented a repressive State in which there was no place for cultural distinctiveness.

The stereotypes expressed by the subjects, as we have seen, are both positive and negative. They often serve a practical (or pragmatic) function in identity building, a phenomenon clearly indicated by the fact that a larger proportion of Catalan subjects (one in four) used stereotypes to define Catalan identity than those of immigrant and 'mixed' families (one in five).

In conclusion, Catalan identity is based on knowing the Catalan language, having a deep sentiment and love for the region, birth in the region or length of residence, and stereotypes. The autonomous government has worked to (successfully) provide knowledge of Catalan to the general population, and to inspire a sense of common history.
Tables

Table 1: Characteristics of Catalan Identity

<table>
<thead>
<tr>
<th></th>
<th>Immigrant parents</th>
<th>'Mixed'</th>
<th>Catalonia-parents</th>
<th>born</th>
<th>% of N=340</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language, customs</td>
<td>31.3 (47)</td>
<td>37.1 (29)</td>
<td>29.2 (26)</td>
<td>32.3 (102)</td>
<td></td>
</tr>
<tr>
<td>Sentiment, feeling of history</td>
<td>27.3 (41)</td>
<td>29 (23)</td>
<td>25.8 (23)</td>
<td>27.5 (87)</td>
<td></td>
</tr>
<tr>
<td>Stereotypes</td>
<td>16.7 (25)</td>
<td>19.2 (15)</td>
<td>24.7 (22)</td>
<td>19.6 (62)</td>
<td></td>
</tr>
<tr>
<td>Birth, residence</td>
<td>13.3 (20)</td>
<td>6.4 (5)</td>
<td>7.8 (7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nothing in particular</td>
<td>4 (6)</td>
<td>5 (4)</td>
<td>5.6 (5)</td>
<td>4.7 (15)</td>
<td></td>
</tr>
<tr>
<td>Things Catalan</td>
<td>4.7 (7)</td>
<td>1.2 (1)</td>
<td>3.3 (3)</td>
<td>3.5 (11)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>2.4 (4)</td>
<td>-----</td>
<td>3.3 (3)</td>
<td>2.1 (7)</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Can one consider oneself Catalan if s/he does not speak the language?

<table>
<thead>
<tr>
<th></th>
<th>Immigrant Parents</th>
<th>Mixed Parents</th>
<th>Cat-born Parents</th>
<th>All Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>87.1 (142)</td>
<td>80.7 (63)</td>
<td>61.6 (61)</td>
<td>78.2% (266)</td>
</tr>
<tr>
<td>No</td>
<td>11 (18)</td>
<td>19.2 (15)</td>
<td>37.3 (37)</td>
<td>20.5% (70)</td>
</tr>
<tr>
<td>Total</td>
<td>(N=163)</td>
<td>(N=78)</td>
<td>(N=99)</td>
<td>(N=340)</td>
</tr>
</tbody>
</table>

Table 3: Can one consider oneself Catalan if s/he were not born in Catalonia?

<table>
<thead>
<tr>
<th></th>
<th>Immigrant Parents</th>
<th>Mixed Parents</th>
<th>Cat-born Parents</th>
<th>All Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>74.2 (121)</td>
<td>82 (64)</td>
<td>82.8 (82)</td>
<td>78.5% (267)</td>
</tr>
<tr>
<td>No</td>
<td>25.2 (41)</td>
<td>17.9 (14)</td>
<td>15.1 (15)</td>
<td>20.5% (70)</td>
</tr>
<tr>
<td>Total</td>
<td>(N=163)</td>
<td>(N=78)</td>
<td>(N=99)</td>
<td>(N=340)</td>
</tr>
</tbody>
</table>

Table 4: Stereotypes: What are the defining characteristics of Catalan Identity?

<table>
<thead>
<tr>
<th>Stereotypes</th>
<th>Examples</th>
<th>% of N=67</th>
</tr>
</thead>
<tbody>
<tr>
<td>Competent, Intelligent</td>
<td>&quot;puede llegar muy lejos;&quot; &quot;culto&quot;</td>
<td>32.8% (22)</td>
</tr>
<tr>
<td>Open, friendly</td>
<td>&quot;simpático;&quot; &quot;alegre&quot;</td>
<td>26.8% (18)</td>
</tr>
<tr>
<td>Nationalist, pro-independence</td>
<td>&quot;mayor sentido de nacionalidad&quot;</td>
<td>20.8% (14)</td>
</tr>
<tr>
<td>Unsharing</td>
<td>&quot;tacaño&quot;</td>
<td>19.4% (13)</td>
</tr>
<tr>
<td>Closed, to themselves</td>
<td>&quot;muy exclusivista&quot;</td>
<td>11.9% (8)</td>
</tr>
</tbody>
</table>
References


The Lexical Representation of Light Verb Constructions

Martin Everaert
Utrecht University

Bart Hollebrandse
University of Massachusetts, Amherst

0 Introduction
In much recent work on the lexical representation of syntactic relations it is argued that a distinction should be made between a predicate's lexical semantic representation - often called Lexical Conceptual Structure (LCS), which represents its lexical meaning, and its lexical syntactic representation - the (Predicate) Argument Structure (AS) - which is derived from the former through some mapping principle(s). A crucial assumption in these approaches is that the rules of grammar have no direct access to LCS but only to AS. We will present evidence from Dutch light verb constructions which, at first indication, seem to indicate that the rules of grammar should have access to information that is generally thought of as encoded at LCS. This would thus present evidence against a disjunctive representation of the lexical information of a predicate.

After a preliminary introduction in section 1 of the notion light verb construction (LVC) and the Argument Transfer (AT) analysis proposed by Grimshaw & Mester (1988) - an analysis in which the merger of the light verb (LV) and light verb nominal (LVN) is treated as a transfer of the AS of the LVN to the LV - section 2 will present some of the main features of LVCs in Dutch. In section 3 we present two problems for an AT-analysis of LVCs and in section 4 we discuss how LVs may be analyzed in a framework in which an LCS/AS distinction is made. Our analysis of the merging process of LV and LVN is discussed in section 5.

1. Light Verb Constructions: an introduction
Jespersen (1965) coined the term LV for the verbs in expressions like to take a walk, to have a drink, to give a kiss, etc.: The verb plus 'direct object noun' forms a complex predicate in which the thematic content of the argument positions seems to be determined by the noun:¹

(1)  a.  Ik geef hem een zoan
      I give him a kiss 'I kiss him'
    a.'  Ik zoan hem
      'I kiss him'

   b.  Ik maakte een val
      I made a fall 'I fell'
   b.'  Ik viel
      'I fell'

One of the striking facts in the discussion of light verb constructions is that the notion LV is hardly ever clearly defined. A good example is presented by Kearns' (1989:123) description: "The defining characteristic of those expressions is that the main semantic content of the predicate is provided not by the verb, but by the action nominal complement". In sections 2.1 and 4 we will try to be more precise about what it actually means to say that a verb is a LV (cf. also Hollebrandse & Van Hout (to appear)). The object nouns zoan (kiss), val (fall) in (1) which determine the semantic content of the open argument positions, we will call LVNs.

Central in the discussion of LVCs is the question how the LV and the LVN are combined, fused. To exemplify this issue, we will briefly review Grimshaw & Mester's (1988) AT-analysis of LVCs in Japanese. They claim that the Japanese
verb *suru* in sentences like (2a) are LVs. They argue that the verb has no semantic content and is only specified for case-marking properties (cf. (2b)). They assume that LVs select theta-transparent NPs, i.e. NPs which have a theta-grid (cf. (2c)). Through the mechanism of AT the verb *suru* takes over the argument structure properties from the noun with which it forms a complex predicate. In essence, this means that the LV starts out as an 'empty' verb and becomes a full verb (FV) via AT (cf. (2d)). Consequently, the LV is able to project its lexical properties like any other verb. The theta-roles of the noun are thus assigned to elements outside the maximal projection of the NP.

(2)  
\[\text{a. } \text{John-wa murabito-ni ookami-ga kuru to keikoku-o shita} \]
\[\text{John-TOP villager-D wolf-N come warning-A do} \]
\[\text{‘John warned the villagers that a wolf would come’} \]

\[\text{b. LV: } \text{suru } (<acc>) \]

\[\text{c. LVN: } \text{keikoku } (\theta_a, \theta_g, \theta_t) \]

\[\text{d. LVC: } \text{suru } (<acc>) + \text{keikoku } (\theta_a, \theta_g, \theta_t) \rightarrow \text{suru } (\theta_a, \theta_g, \theta_t)<acc> + \text{keikoku } () \]

The merger of LV and LVN is a manipulation of ASs. Crucial in the AT-analysis is that (i) the LVs have no theta-role specification and that (ii) the LVNs have a theta-role specification. We will challenge both assumptions for Dutch LVCs. The Dutch LVs in (1) are not as empty as Japanese *suru* is; some information is still present in the argument structure of the verb.2 The essential observation is that for the majority of LVs in Dutch (and maybe even all) the LV also has a full variant. For instance, Dutch *maken* in *maakte een val* (cf. (1b)) is a LV but a FV in *maakte een tafel* (made a table). We will show that LV *maken* retains essential lexical properties of its related FV. In other words, LVs are not empty place holders but verbs which have lexical properties of their own, although not complete. In our analysis LVs are partly specified and partly inherit their lexical properties from the LVN.

2. Light Verb Constructions: some observations
A first approximation of the properties we attribute to LVCs is given in (3); we will elaborate (3a,c) in section 2.1 and (3b) in section 2.2.

(3)  
\[\begin{align*}
\text{a. A LV is lexically related to its full variant.} \\
\text{b. The LVN is an event nominal and, as such, has AS-specification.} \\
\text{c. The LV and the LVN form a complex predicate, i.e. the lexical properties of the LVC as a whole are determined on the basis of fusion of the lexical properties of the LV and LVN.}
\end{align*} \]

2.1 (Light) verbs
If the LV and the LVN form a complex predicate, the LVN cannot be an argument but has to have predicative status. We will assume that this means that the phrase headed by the LVN is an NP and not a DP. In other words the complement of the LVN does not head an extended projection (Grimshaw 1991). Given this hypothesis we predict that the LVN behaves different from a regular object, syntactically. It follows, for instance, from the NP/DP status of LVNs that WH-movement is excluded, that pronominalization is forbidden and that modification of the LVN is predicate modification. In Hollebrandse (1993) this is shown to be true (cf. also Kearns (1989)).
A LVC contains two predicative elements but it acts as one. One could thus describe it as a discontinuous predicate. An important question is how to describe this discontinuity. In the AT-analysis this is clear: it is the LVN that contains all the information that defines a head as a predicate (theta-specification), and this information is transferred to the LV. In section 3.1 we will present evidence that such an analysis cannot be upheld.

Because the LV and the LVN behave as a complex predicate, the syntactic properties of the LV (cf. (3a)) and the semantic properties of the LVN (cf. (3b)) must be compatible. In (4) violations of this requirement are illustrated:

(4)  a. *Jan geeft Marie een gil
    Jan gives Mary a yell
    b. Jan geeft een gil
       Jan gives a yell ‘Jan yells’

(5)  a. *Jan maakt mijn jurk een opmerking
    Jan makes my dress a remark
    b. Jan maakt een opmerking over mijn jurk
       ‘Jan makes a remark about my dress’

LV geven in (4) is syntactically a tryadic predicate, i.e. the verb is subcategorized for two object NP-positions. The LVN gil is semantically a monadic predicate. With regard to its syntactic arity the LVC in (4a) is grammatical, but the semantic arity is wrong. There are two open argument positions - the subject position and the secondary object position - but given the LVN gil there is only one theta role available. Hence the theta criterion is violated in (4a). In (4b) the syntactic arity and the semantic arity of the LV and the LVN are compatible, hence the grammaticality of (4b). The examples in (4) thus illustrates what happens if the LV in the LVC defines more argument positions than the LVN can accommodate. (5) illustrates the other possibility: the LVN in the LVC defines more argument positions than the LV can supply for, syntactically. The LVN opmerking is semantically a dyadic predicate. As a result, both the subject and the secondary object in (5a) could receive a theta-role, but syntactically the latter position is not licensed by the LV. In (5b) the object theta-role is projected as a complement of the LVN, not of the LV, and hence its grammaticality.

2.2 Light Verb Nominals
The nominals that are used as LVNs seem to have a common semantic property: Kearns (1989) describes them as action nominals. Whatever the right description might be, it should incorporate the nouns in (6a,b), but not (6c):

(6)  a. Jan maakte een beweging, een analyse, een afspraak, ...
       ‘Jan made a move, an analysis, an appointment, ...’
    b. Jan maakt een doelpunt, een grimas (tegen), een bedevaart, ...
       Jan makes a goal, a grimace (at), a pilgrimage, ...
       ‘Jan scores a goal, makes faces (at), goes on a pilgrimage, ...
    c. Jan maakte een stoel, een auto, een boek, ...
       ‘Jan made a chair, a car, a book, ...’

If we look at the type of nouns in (6a,b) we can observe that it are all NPs denoting actions or events. This means that if an NP can only be used as denoting objects it will not be found in LVCs. Nouns like stoel, auto or boek (cf. (6c)) are strictly
object denoting nouns, and can, therefore, not be used as LVNs. Of course many nouns are ambiguous between an object and an event denoting reading. In these cases only the latter (cf. (7b)) is found in the LVC.

(7)  
  a. Marie gaf Jan een schop
      ‘Marie gave Jan a shovel’
  b. Marie gaf Jan een schop
      Marie gave Jan a kick ‘Mary kicked Jan’

Although all nouns in (6a,b) could be classified as event-denoting, there is difference between the nouns in (6a) and those in (6b): only the nouns in (6a) are morphologically derived from verbs. In section 3.2 we will elaborate on the distinction between object and event denoting nouns.

3. Two arguments against an Argument Transfer analysis

3.1 Light Verbs are not ‘empty’

An analysis of all LVCs by means of AT would have the implication that all LVs are essentially identical. So there would be no intrinsic difference between the LV _geven_ (to give) in (8a) or the LV _krijgen_ (to get) in (8b):

(8)  
  a. Zij geeft hem een zoen
      ‘She gives him a kiss’
  b. Zij krijgt van hem een zoen
      She gets from him a kiss ‘She gets a kiss from him’

However, the subject of _geven_ in (8a) is the one who kisses while the subject of _krijgen_ in (8b) is the one who is being kissed, and no other interpretation is possible. The LVN _zoen_ has two thematic roles available but it is the LV that determines which of the thematic roles is realized as the subject of the LVC. The examples in (8) further illustrate that the LV also determines how the other available thematic role is syntactically projected. In (8a) the remaining thematic role of _een zoen_ is realized as a secondary object of _geven_ while it is realized as a prepositional object in the case of _krijgen_. The fact that the internal role is realized in this particular syntactic position is no surprise given the FV equivalents of these LVs, compare (8a,b) with (9a,b), respectively.

(9)  
  a. Hij geeft haar een boek
      ‘He gives her a book’
  b. Zij krijgt van hem een boek
      She gets from him a book ‘She gets a book from him’

A similar problem arises in the case of impersonal passives. Observe the examples in (10-13):

(10)  
  a. Het vliegtuig landt
       ‘The airplane lands’
  b. *Er wordt geland door het vliegtuig
     There is landed by the airplane

(11)  
  a. Het vliegtuig maakt een landing
       ‘The airplane makes a landing’
b. Er wordt een landing gemaakt
   There is a smooth landing made ‘The airplane lands’

(12) a. Hij zoent
   ‘He kisses’
   There is kissed ‘People are kissing’

b. Er wordt gezoend
   There is kissed ‘People are kissing’

(13) a. Hij krijgt een zoen
   He gets a kiss ‘He is being kissed’

b. *Er wordt een zoen gekregen
   There is a kiss got

In an AT-analysis LVs become FVs, completely specified for argument structure. This means that, after AT, the lexical specifications of the LVs *maken* in (11a) and *krijgen* in (13a) are identical to the lexical specifications of the FVs *landen* in (10a) and *zoenen* in (12a). Still, the FVs behave differently from the (former) LVs with respect to passivization, as (10b-13b) illustrate. What seems to be relevant are the properties of the LV, contrary to what the AT-analysis seems to predict. Since the FV *maken* allows passivization (cf. (14a)) the LV *maken* allows passivization, and because the FV *krijgen* does not allow passivization (cf. (14b)), the LV *krijgen* does not allow passivization:

(14) a. Er werd een stoel gemaakt
   There is a chair made ‘A chair is made’

b. *Er werd een boek gekregen
   There is a book got

3.2 Do light verb nominals have an argument structure?
In 2.2 we have stated that LVNs can be defined as event denoting. We will now make explicit what we mean by that. Following Grimshaw (1990:59) we will classify nouns as in (15):

(15) a. dog
    in: the dog
    is a result noun

b. observation
    in: the unexpected observation
    of a black hole
    is a simple event noun

c. examination
    in: the examination of the
    patients took a long time
    is a complex event noun

In Grimshaw’s view all nouns have an LCS, or lexical meaning, but not all nouns have an AS. She argues that it is not possible for a head to be specified for thematic information without having an event structure. Simple event nominals do refer to events at the level of their LCS but do not have AS properties. In other words, both classes could be described as having LCSs which refer to an event. Result nouns - in our view an instantiation of the wider class of object denoting nouns - are defined as nouns having LCSs which do not refer to an event, and, therefore, not having ASs. From this perspective LVNs are either simple event nominals or complex event nominals. There are, however, several indications that the LVNs in the cases under discussion are simple event nouns, or behave as such.

Firstly, Grimshaw (1990:67) claims that zero-derivation (conversion) derives either result nouns or simple event nominals, and crucially not complex event nouns. Given this it is quite surprising that LVCs so often make use of zero-derived nouns (Kearns 1989, Hollebrandse 1993).
Secondly, Grimshaw observes that the numeral one is not compatible with complex event nouns (16a), but is compatible with simple event nouns (16b). The examples in (17) show that the Dutch LVNs are simple event nominals

(16) a. Een/*één landing door een vliegtuig is altijd een fascinerend gezicht
   A/one landing by an airplane is always a fascinating sight 'It is always a
   fascinating sight to see an airplane land'
   b. Een/één bedevaart is altijd verstandig
   A/one pilgrimage is always wise 'It is always wise to go on a/one
   pilgrimage'

(17) a. Het vliegtuig maakte een/éér. lading
   'The airplane made a/one landing'
   b. Hij gaf een/één schreeuw
   He gave a/one yell 'He yelled (once)'

Thirdly, Grimshaw argues that if a non-plural noun disallows aspectual modifiers like regelmatig (regular, frequent) or voortdurend (constant) it will be a simple event nominal (or a result nominal) (18a). Complex event nouns do allow these types of modifiers (18b). Given the ungrammaticality of (19b,20b) we have to conclude that the LVN is a simple event noun:

(18) a. *De dokter adviseerde een regelmatige bedevaart naar Lourdes
   The doctor advised a regular pilgrimage to Lourdes
   b. Regelmatiche controle van het apparaat door een onderhoudsmonteur is
   belangrijk
   Regular control of the appliance by a mechanic is important 'It is important
   that the appliance is regularly controlled by a mechanic'

(19) a. Hij gaf regelmatig een demonstratie
   He gave regularly a demonstration 'He regularly gave a demonstration'
   b. *Hij gaf een regelmatige demonstratie
   He gave a regular demonstration

(20) a. Hij maakte voortdurend een beweging
   He made constantly a move 'He moved constantly'
   b. *Hij maakte een voortdurende beweging
   He made a constant move'

Van Hout (1991) adds a fourth test, for Dutch. She observes that in Dutch nouns the external argument can be either expressed by a van (of)-PP or a door (by)-PP. She observes that complex event nouns can only use the door-PP to express the agent:

(21) a. De verwaarlozing *van/door dat stel van hun kind is schokkend
   The neglect of/bye that couple of their first child is shocking 'The neglect of
   their child by that couple is shocking'
   b. De vertalingen van/door Karel zijn over het algemeen slecht
   The translations of/bye Karel are in the general bad 'Karel's translations are
   generally bad'

Again, the LVN does not seem to behave as a complex event noun. (22b) shows that passivization of the LVC (22a) is possible. However, (22c) shows that it is impossible to interpret the door-phrase as a complement of the LVN.
(22) a. Hij maakte een afspraak met Karel
   'He made an appointment with Karel’
   b. Een afspraak met Karel kon niet worden gemaakt door mij
      An appointment with Karel could not be made by me ‘I was not able to
      make an appointment with Karel’
   c. *Een afspraak met Karel door mij kon niet worden gemaakt
      An appointment with Karel by me could not be made

On the basis of (17-22) there seems to be reason to conclude that the LVNs in
the LVCs under discussion do not seem to behave as complex event nouns but
more as simple event nouns.

3.3 Summary
On the basis of the preceding discussion we are forced to two conclusions:
(i) The LVN has less thematic content than we need in an AT analysis. In AT the
   LVN must have argument structure properties that it can transfer to the LV.
   However if LVNs are simple event nouns they don’t have argument structure
   properties, and, therefore, have nothing to transfer.
(ii) The LV has more information than we expect in an AT analysis. It seems as if
   the LV has retained some of the properties of its FV equivalent.

In the next section we will outline our view on the lexicon-syntax interface. After
that we are in a position to discuss what the consequences are of (i-ii).

4. The lexicon-syntax interface
The LCS of a predicate is the ‘deep’ semantic description which is probably unique
for any particular predicate, or class of predicates. Such a semantic description is
mapped onto a more syntax-like representation, AS. AS represents how many
arguments a verb requires and to which syntactic argument positions they are
linked. The AS representation is not unique for individual predicates or classes of
predicates. Two different predicates like walk and swim will have the same AS.
Although essentially different, LCS and AS are part of the lexical representation
of a predicate and thus part of the lexicon, to be distinguished from syntax, as
sketched in (23):

```
(23) LCS --linking--> AS --linking--> SS
    rules                      rules
    LEXICON                   SYNTAX
```

What is represented in (23) is a conflation of different positions in the literature (cf.
Carrier & Randall (1993), Grimshaw (1990), Hale & Keyser (1986), Rappaport &
Levin (1988), Zubizarreta (1987)). Two questions arise:

A. Should we really make a distinction between LCS and AS, or is it possible to
have one, although internally structured, lexical representation (cf. Jackendoff
(1990), Levin & Rappaport (1994), Pinker (1989), Van Valin 1990)? In the former
position there are two mapping relations. In the latter position there is only one full-
fledged LCS which directly maps onto syntax.

A good example is given in Žaenen (1993). She argues that auxiliary selection is
sensitive to intrinsic argument classification of a predicate (Bresnan & Kanerva
1989, Bresnan & Zaenen 1992), which to some extent could be equated with AS. However, Zaenen describes the availability of passivization in terms of the notion controllability. Verbs which have a volitional dimension can be passivized. Although she doesn’t make explicit statements about it, one might hypothesize that the notion controllability is not part of AS. In other words, auxiliary selection seems to be sensitive to AS, passivization to LCS. If the latter conclusion is on the right track, one might argue that the distinction between LCS and AS is superfluous.

B. If we do make a distinction between LCS and AS, what is the status of these two levels of representations in the grammatical system? If AS is all that syntax sees and LCS is inaccessible, what role does LCS play?

The fact that LCS is not directly accessible for syntax does not necessarily mean that LCS plays no role whatsoever in grammar. One might envisage that LCS is relevant for the lexicon. Lexical rules should, perhaps, be formulated in terms of LCS (cf. Pinker (1989)). Speas (1990), for instance, explains dative shift as derived by a lexical rule relating two different LCSs, resulting in two different ASs and Carrier & Randall 1993 claim that resultative verbs are derived by a lexical rule, and that this lexical rule should be formulated in terms of LCS-representations, and not AS-representations.

For argument sake we will adhere to the position that some grammatical phenomena are best described in terms of LCS, and that, as such, this level of representation is motivated. If we would not make a distinction between LCS and AS, all relevant information of LCS could, in principle be relevant for syntax. In other words, syntactic phenomena could be sensitive to any kind of semantic information that is made available at LCS. Since this is clearly not the case, it makes sense to make an LCS/AS distinction and restrict accessibility of syntax to AS. In summary, (23) must be interpreted as in (24):

(24) a. Syntax has only access to AS
   b. AS is a derivative of LCS, in the unmarked case
   c. LCS contains only information that is semantic in nature, i.e. notions like agent/goal/theme, volitional, CAUSE, etc.
   d. AS contains only information that is ‘syntactic’ in nature, i.e. (i) the number of argument positions and their relative prominence, and (ii) the case-marking properties of the verb.

The position outlined in (24) enables us to be more specific about shifts in meaning, as in the case in of LVs, idioms or, in general, polysemy. That is to say, we will assume that the following statement is true (cf. also Everaert (to appear)):

(25) (Idiosyncratic) meaning shifts affect only the LCS of a verb and leave the AS of a verb unchanged

On the basis of (25) LVs are defined as verbs having the same AS as their FV equivalent, but without the same corresponding LCS:

(26) a. literal make: LCS\textsubscript{X}, AS\textsubscript{X}  b. light verb make: LCS\textsubscript{Y}, AS\textsubscript{X}

5. Analysis
How can we interpret the conclusions of section 3.3 in view of what is stated in the preceding section?
We have seen that for LVNs the only distinction that seems to be relevant is the difference between nouns having an event denoting LCS and nouns having an object denoting LCS. In LVCs the distinction between simple event nouns and complex event nouns seems to disappear. Given (24a) we are forced to the conclusion that simple event nominals must have argument structure properties, distinguishing them from result nouns. One way of executing this is to refute Grimshaw’s claim that it is not possible for a head to be specified for thematic information without having an event structure. One might envisage that all event nouns have ASs but that complex event nouns in addition have event structure properties. This allows us to explain why LVNs behave as they do, and, at the same, account for the fact that there still is a difference in behaviour between simple event nouns and complex event nouns in other configurations (cf. Van Hout (1991)). However, the situation is, in fact, more complicated. Even straightforward complex event nouns behave as if they were simple event nouns when they are LVNs. It is not clear why this is so. We do not have the space to discuss this issue fully but the following approach seems to suggest itself. Hollebrandse (1993) presents evidence that LVNs are NPs and not DPs. This could explain why the LVN does not show any evidence of having an event structure representation. It has been argued that the relation of D w.r.t. N is similar to that of INFL (or Tense) w.r.t. V (cf. Zwarts (1992)). If we follow Higginbotham (1985) in his assumption that the event-specification of a verb is discharged by an inflectional element, then D could be taken as the element which discharges the event-specification of a noun. If there is no D, the event-specification cannot be discharged. If we, furthermore, assume that in the case of nouns event-specification is not mapped from LCS if it cannot be discharged, then we account for the observed behaviour of LVNs.

From the principle (25) it follows that if a FV becomes a LV it must retain its AS-specification while its LCS-specification changes. Since merging of the LV and the LVN does not seem to be lexical in nature - the LVC does not behave as a single lexical unit - it should not be described at the level of LCS (cf. Ramchand (1990) and Neeleman (1994) for relevant discussion). We are, thus, forced to the conclusion that the ASs of the LV and the LVN fuse, merge. In the literature several mechanisms have been suggested to this effect (cf. Ackema (1995), Alsina (1992), Isoda (1991), Neeleman (1994), Ramchand (1990), Rosen (1990), a.o.). We do not have the space here to discuss the differences between these proposals and how they differ from the AS merging process we ourselves want to advocate. We will simply give a brief outline of our analysis (cf. also Hollebrandse (1993) and Everaert (to appear)).

We assume that the argument structures of the LV and LVN are merged through the process of theta-identification Higginbotham (1985). An important feature of our analysis is that the LV contains specific information about the identification procedure. That is, the LV specifies which of its argument positions is identified with which argument position from the LVN (cf. Williams 1985). The remaining argument position of the LV will be used up by the LVN itself. In the case of a LVC as in (1a) (een zoent geven) the lexical specification of the LV and LVN can be described as in (27):

(27) LV: gewen ‘give’ (Q_3(θ_y(θ_LVN))(θ_x) identifies with θ_p, θ_y identifies with θ_q)

LVN: zoen ‘kiss’ (θ_y(θ_q))

Under the standard assumptions about percolation this will result in a syntactic
structure like (28):

(28) \[ \begin{array}{c}
   \text{NP} \\
   \text{een zoen} \\
   (\theta_p(\theta_q)) \\
   \vdots \\
   \text{geven} \\
   (\theta_x(\theta_y(\theta_{LVN}))) \\
\end{array} \] 

Now LVs will differ in the formulation of the identification procedure. The LV *krijgen* ‘to get’, for example, forms a minimal pair with the LV *geven*. The only difference between *geven* en *krijgen* is the information about identification. For *krijgen* this information is specified as in (29).

(29) LV: *krijgen* ‘to get’ (\(\theta_x(\theta_y(\theta_{LV}))\))

LVN: *zoen* ‘kiss’ (\(\theta_p(\theta_q)\))

It is thus clear that the *geven/krijgen* alternation observed in (8) is readily explainable in terms of theta-identification. We don’t have to refer to the thematic content of argument positions, only to their relative position in the ASs. This analysis will also account for the thematic mismatches in (4,5).

The passivization facts are less straightforwardly explained in terms of theta-identification. It strongly depends on your analysis of why certain predicates resist or allow passivization. If we would follow Zaenen (1993), Levin & Rappaport (1994) a.o. that controllability or volitionality is the crucial semantic factor, this would create a problem for theta-identification. It is not straightforwardly clear that such a notion is encoded at the level of AS, although it not, in principle, excluded (Cf. Jackendoff (1990:125-130) who encodes this feature on the action tier).

However, another line of reasoning is possible. Grimshaw (1990) argues that passivization is not possible if there is no external argument. She defines an external argument as an argument that has highest prominence both on the level of thematic structure and on the level of event structure. This explains why unaccusative predicates like in (10) don’t allow passivization. It also explains the grammaticality of (11b) under the assumption that because the LV *maken* retains its AS-specification it will be marked for an external thematic role. From this it follows that passivization is possible despite the fact that the LVC as a whole has an unaccusative semantics (cf. Everaert to appear). The question is whether our analysis can also give an explanation for the ungrammaticality of (13b)? Or in other words, why is (14b) *Er wordt een boek gekregen* ungrammatical? If we can prove that the verb *krijgen* has no external argument in the above sense, then there is no problem. We leave this problem open here.

References
Carrier, Jill, and Janet Randall. 1993. Lexical Mapping. Knowledge and Language II, Lexical and


1 Observe that Jespersen also used the term LV for do in the constructions John did see him. In recent publications the terminology has also been used for causative verbs, raising verbs and serial verbs (cf. Rosen (1989), (1990)). In Everaert (to appear) the relation between participles and aspectual auxiliaries hebben/zijn are analyzed as LVCs (cf. also Ramchand (1990)).

2 Isoda (1991) discusses evidence that indicates that even for Japanese suru (i) is untenable.

3 We assume that the LCS of a LV is incomplete compared to the FV and not completely different, or empty, as might be the case of idioms.
Synchronic and Diachronic Typology: The Case of Ejective Voicing*

Paul D. Fallon
Ohio State University

1. Introduction

Although ejectives are the fourth most common type of stop (Henton, Ladefoged, and Maddieson (1992)), and occur in approximately 20% of the world’s languages (Catford 1992), there is relatively little known about their phonological behavior in both synchronic rules and diachronic sound changes. And yet ejectives have now become an important area of research in both phonological theory and in historical linguistics. In phonology, ejectives are important in evaluating the alleged privativity of laryngeal features (Lombardi 1991, 1995) and in testing the possible rule types such as spreading and delinking within autosegmental phonology (McCarthy 1988, Clements and Hume 1995). In historical linguistics, the behavior of ejectives has become central to Indo-European studies in light of the Glottalic Theory. In this paper I will focus on putative cases of ejective voicing by examining phonological data from both synchronic and diachronic sources. The remainder of this section will discuss the background and issues at stake in phonological theory (1.1) and in historical linguistics (1.2). In section 2, I examine putative cases of ejective voicing and show that the direct laryngeal feature changing is not necessary. I then turn to look at several diachronic cases of ejective voicing in section 3. Loanwords with ejectives provide several cases of ejective voicing and these are discussed in section 4. Conclusions are drawn in section 5.

1.1 Phonological background

In the generative phonology of The sound pattern of English (Chomksy and Halle 1968), the evaluation metric of phonological rules could not formally distinguish between common rules of assimilation (1a) and unattested rules of random changes (1b), which both had the same formalism:

(1a). [+syllabic] → [+nasal] / ___ [+nasal]
(1b). [+anterior] → [+high] / ___ [+round]

In Clements (1985) the formalism of feature geometry organized phonological features into hierarchical class nodes, grouping, for example, laryngeal features under one node, and place features under another. This model allowed common sound changes such as assimilation to be expressed naturally, while uncommon sound changes were either impossible to express or more complicated (see McCarthy 1988). Thus in this model, assimilation is viewed as the spreading of one feature by creating an association line from one segment to another, for example. Dissimilation is viewed as delinking of a feature with subsequent default fill-in. In addition, in Clements and Hume (1995), feature-changing rules are admittedly only reluctantly. They suggest that feature-changing rules involving such features as [sonorant], [vocoid], and [continuant] ‘may be required to express processes of strengthening and weakening’. No suggestion is made that laryngeal features may be changed directly. This is a falsifiable hypothesis, and may be disproven empirically if there are indisputable cases of, say, laryngeal features which change from one feature to another directly. In this paper, I will look at
potential cases which illustrate alternations between ejective and voiced, but will conclude that other explanations are possible. Thus phonological theory does not need to express such laryngeal alternations through feature changing.

1.2. The Glottalic Theory and typology

The Glottalic Theory of Proto-Indo-European (PIE) posits that traditionally reconstructed PIE voiced stops should be reinterpreted as glottalic consonants (usually ejectives). It was formulated initially by Gamkrelidze and Ivanov (1972; see also 1973, 1984, 1995) and somewhat independently by Hopper (1973), and since then has generated a host of literature (see Salmons 1993 for an accessible overview). A comparison of the traditional reconstruction and the Ejective Model of the Glottalic Theory (Job 1989) is illustrated by the dental series in (2).

<table>
<thead>
<tr>
<th>Series</th>
<th>Traditional</th>
<th>Glottalic (Ejective Model)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>d</td>
<td>t'</td>
</tr>
<tr>
<td>II</td>
<td>dh</td>
<td>d/dh</td>
</tr>
<tr>
<td>III</td>
<td>t</td>
<td>t/th</td>
</tr>
</tbody>
</table>

Advocates of the Glottalic Theory argue that historical reconstructions should be accountable to typological data. The following quote by Gamkrelidze and Ivanov is illustrative:

‘Thus reconstructions can be considered real if they are consistent with two basic typological criteria: they must agree with synchronic typological universals and they must agree with diachronic typological universals (general schemas for change and transformation of languages).’ (1995:xcv)

Despite this viewpoint, the vast majority of literature surrounding the Glottalic Theory has dealt only with synchronic typology. For example, arguments to support the theory are based on the patterning and gaps of inventories and on root structure constraints.

Regarding diachronic typology, Gamkrelidze and Ivanov note that ‘a necessary condition for reality of reconstructions is that they must be consistent with diachronic typological data, with schemas for the change of particular linguistic structures over time, as established by the study of historical facts from individual languages’ (1995:xciv-v). This approach implies that the proposed trajectories of sound change should be well documented from attested historical changes. In the Ejective Model, as Job (1989) points out, roughly 75% of the reflexes of the PIE ejectives are voiced stops (in at least seven IE branches, according to Garret 1991), so this model predicts that such a shift should be relatively common. Compare this to the traditional model, in which 75% of the reflexes do not change, while 25% devoice. In his review of Venneman (1989), which deals largely with the Glottalic Theory, Garrett reminds us that:

‘while assessments of linguistic plausibility are undeniably necessary in reconstruction, a model of relationship includes not only a proto-language but the set of linguistic events which resulted in the attested daughter languages, and each of these events must also be plausible’ (1991:796).

The question, then, is, how plausible is the change from ejective to voiced?
The problem with ejectives is simply that there is not yet a firm schema for change. Although Gamkrelidze and Ivanov discuss the importance of diachronic typology, they provide only one detailed, independent example of such a change (Chechen-Ingush, discussed in Gamkrelidze and Ivanov 1995:44-6, and below). They also briefly mention some cases of voicing of ejective affricates in Northeast Caucasian, as well as dissilvative voicing of ejectives from loanwords. Finally, they mention voicing of *Proto-Semitic *q (*k*), in Arabic gāl dialects. This is not a firm empirical basis on which to base a proposed diachronic typological ‘schema’.

To my knowledge, only Job 1984, 1989 has examined diachronic shifts involving ejectives, and most of his examples are from the languages of the Caucasus. This study, then, will expand our database of sound change by investigating putative cases of ejective voicing, drawing first upon attested synchronic alternations, then diachronic changes, and finally loanword phonology.

2. Synchronic alternations

In this section, I will review potential cases of synchronic ejective voicing. The question for phonological theory is whether these cases need to be described in one step involving laryngeal feature changing, or whether there are other explanations at hand. I will argue that all of these cases are explainable through feature delinking and default, or from assimilation, and thus do not require reassessment of the prohibition against direct laryngeal feature changing.

Donnelly (1992) reports that in Xhosa, a Nguni language which has full aspirated, ejective, and breathy-voiced series, in addition to its many clicks, ‘/*k*/ is the only voiceless ejective which tends to be unejected and frequently voiced’ (1992:2). In the first syllable of any root, ejection is maintained; however, in non-initial position, the velar ejective is never fully ejected and can always be voiced:

\[
\begin{align*}
\text{-p\textsuperscript{b}ek'-a/} & \quad [\text{-p\textsuperscript{b}ega}] & \text{‘cook’} \\
\text{/bek'-a/} & \quad [\text{\textsuperscript{b}ega}] & \text{‘put’} \\
\text{-t\textsuperscript{b}ak'at\textsuperscript{b}-a/} & \quad [\text{t\textsuperscript{b}agat\textsuperscript{b}a}] & \text{‘bewitch’} \\
\text{/dak'aq-a/} & \quad [\text{\textsuperscript{d}aq\textsuperscript{a}d\textsuperscript{a}}] & \text{‘tear to pieces, mangle’} \\
\text{-k\textsuperscript{b}ok'-a/} & \quad [\text{\textsuperscript{k}o\textsuperscript{a}g\textsuperscript{a}}] & \text{‘draw out’}
\end{align*}
\]

This phenomenon of ejective voicing appears to be postlexical and structure-creating, since plain voiced stops are not found underlingly. (Doke (1967:92) notes a similar process for Zulu, another Nguni language). The question is whether the process takes place in two steps, or in one step. In the orthodox view, first the ejective feature [constricted glottis] is delinked from the laryngeal node of a velar and then there is fill-in of the feature [voice], perhaps by default. (Kiparsky’s (1995:646) treatment of intervocalic voicing is incompatible with a privative view of laryngeal features.) On the other hand, if we allow laryngeal feature changing, we can accomplish this rule in one step, in this case [c.g.] $\Rightarrow$ [voice]. However, I believe that we can avoid laryngeal feature changing if we view ejective voicing as the result of the spread of [voice] from an adjacent vowel onto the velar ejective. Although [voice] on vowels is typically underspecified, this rule does appear to be post-lexical, at which point in the grammar many theories assume full specification of segments.

Another possible case of ejective voicing is found in Doke (1967:41), who describes the ‘somewhat irregular’ process of ‘vocalization’ in Venda, a Southeastern Bantu language\(^2\). In the formation of nouns of class 5, the ‘unvoiced
explosives’ change regularly such that /p’ t’ t’ k’/ → /b d d g/. The voiced consonants remain unchanged, e.g. /gonal/ ‘knee’, /magonal/ ‘knees’ (159). (The situation with the fricatives is more complicated and need not concern us here). Examples are given in (4a), in which there is no alternation and the class prefix remains, and in (4b) in which there is alternation between the initial voiced consonants in the singular (without an overt prefix) and their cognate ejective consonants in the plural:

(4a)  
<table>
<thead>
<tr>
<th>sg. (cl. 5)</th>
<th>pl. (cl. 6)</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>liño</td>
<td>mano</td>
<td>tooth</td>
</tr>
<tr>
<td>lijo</td>
<td>maño</td>
<td>eye</td>
</tr>
<tr>
<td>lila</td>
<td>mala</td>
<td>intestine</td>
</tr>
</tbody>
</table>

(4b)  
<table>
<thead>
<tr>
<th>bako</th>
<th>map’ako</th>
<th>cave</th>
</tr>
</thead>
<tbody>
<tr>
<td>dope</td>
<td>mat’ope</td>
<td>mud</td>
</tr>
<tr>
<td>ḷaḳa</td>
<td>maṭ’aka</td>
<td>bush</td>
</tr>
<tr>
<td>gumbu</td>
<td>mak’umwu</td>
<td>big calabash</td>
</tr>
</tbody>
</table>

(Doke 1967: 157)

According to Doke, ‘in each case the unaffected stem consonant is shown in the plural, the voiced consonant of the singular indicating the action of the suppressed prefix ji-’ (157). That is, the noun class 5 prefix disappears and there is concomitant mutation (‘vocalization’) of the initial root consonant. The plural prefix, /ma-/ in class 6 nouns, shows the initial stem without such voicing.

It is possible to interpret this as an instance of (morphologically restricted) ejective voicing due to feature changing, but this view is otherwise unmotivated. Another possibility would view this as the assimilatory spreading of [voice] from the prefix marker, which is subsequently deleted. At any rate, although there is clearly alternation between ejective and voiced consonants, it is not obvious that we must adopt a feature-changing approach, so I propose that this is another case of the assimilatory spreading of [voice].

Another possible case of ejective voicing is found in the Daghestanian language Lezgian, which has four underlying series of stops: ejective, voiced, voiceless aspirated, and voiceless unaspirated, and thus it is an ideal language on which to test theories of laryngeal feature-changing, especially since it has many unusual laryngeal alternations. I will ultimately argue that Lezgian alternations between ejective and voiced sounds are due to the role [voice] plays as a default laryngeal feature.

Following Lombardi’s (1991) approach, I propose that Lezgian stops are represented as follows:

(5) a. voiceless unaspirated  
| root  | (Lar) |

b. voiceless aspirated  
| root  |

  | Lar  |

c. voiced  
| root  |

  | Lar  |

d. ejective  
| root  |

  | [s.g.] (=asp) | [voice] |

  | [c.g.] (=gl) |
All stop series except the voiceless unaspirated bear a single, privative laryngeal feature. The voiceless unaspirated stops do not bear any laryngeal feature and therefore the laryngeal node itself is not present. Normal cases of laryngeal neutralization involve the delinking of a laryngeal feature, typically in syllable- or word-final position. The neutralized consonant, then is a plain voiceless unaspirated consonant. In Lezgian, however, there is a constraint against word-final unaspirated stops (Haskel 1993). In current theories, this would have to be expressed as a positive licensing constraint that all word-final stops must have a laryngeal node, since we cannot ban nodes which are not present. This constraint thus accounts for the presence of word-final voiced, aspirated, and ejective stops, and the absence of voiceless unaspirated stops.

Lombardi notes that languages may have marked default consonants. Lombardi (1995:39) states that:

> It would be possible to analyze neutralization to voiced (or other laryngeally marked) obstruents in this system, if an authentic case were found. It would require [a rule delinking the laryngeal node], and then a specific rule of fill-in of [voice] (or other feature) on unmarked sounds. This is obviously more complex than linking alone, which explains why it never occurs (or if such a case were found, it would account for its extreme rarity).

I propose that Lezgian could be a language in which the marked feature [voice] is such a default. I should note here that this is not always the result of default, and occurs only in certain morphological environments. Let us first examine the rule involving alternations between ejectives and voiced consonants, and then the evidence for [voice] as a default.

Lezgian shows alternations between word-final ejectives and voiced stops when there is a preceding ejective in a closed class of lexical items. Compare the following forms from Haskel (1993:61):

<table>
<thead>
<tr>
<th>Ergative Singular</th>
<th>Absolutive Singular</th>
</tr>
</thead>
<tbody>
<tr>
<td>q'ep'ini</td>
<td>q'ebo</td>
</tr>
<tr>
<td>t'ap'unii</td>
<td>t'abo</td>
</tr>
</tbody>
</table>

‘cradle’ ‘block’

Another rule of [e.g.] Dissimilation (Haskel’s Pre-ejective Ejective Aspiration) deglottalizes the preceding consonant when there is an intervening high vowel. In these cases, the first root ejective in the singular is deglottalized, and after a syncope rule due to stress shift in the absolutive singular, the stop is subsequently aspirated phonetically. The final consonant, which is an underlying ejective, undergoes Deglottalization in the absolutive singular, as it did in (6) above. The following are thus additional examples of the same alternations between ejectives and voiced stops.

<table>
<thead>
<tr>
<th>UR (pl.)</th>
<th>SR (Plural)</th>
<th>Abs. Sg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>t'up'ar</td>
<td>thw'p'er</td>
<td>t'ub</td>
</tr>
<tr>
<td>tf'ip'er</td>
<td>tf'ib'p'er</td>
<td>tf'ib</td>
</tr>
<tr>
<td>ts'ip'er</td>
<td>ts'ib'p'er</td>
<td>ts'ib</td>
</tr>
<tr>
<td>t'ip'er</td>
<td>t'ib'p'er</td>
<td>t'ib</td>
</tr>
<tr>
<td>ts'ik'er</td>
<td>ts'ib'k'er</td>
<td>ts'ig</td>
</tr>
<tr>
<td>q'yt'er</td>
<td>q'ib'ut'er</td>
<td>q'yd</td>
</tr>
</tbody>
</table>

‘finger’ ‘span’ ‘pot’ ‘owl’ ‘middle’ ‘winter’
I formalize the rule as follows:

(8) Lezgian Deglottalization and Default voicing

<table>
<thead>
<tr>
<th>Delinking</th>
<th>Default</th>
</tr>
</thead>
<tbody>
<tr>
<td>root V root_ω</td>
<td>root V root_ω</td>
</tr>
<tr>
<td>Lar</td>
<td>Lar</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>[c.g.]</td>
<td>[c.g.] →</td>
</tr>
<tr>
<td></td>
<td>[c.g.] [voice]</td>
</tr>
</tbody>
</table>

(Morphologically restricted to some nouns in the absolutive singular)

This rule states that in the absolutive singular, the second of two ejectives dissimilates to become voiced, first by deglottalization and then by default fill-in of [voice]. As noted above, these alternations are restricted to a closed class of lexical items. There are many instances of ejectives in word-final position, even in the absolutive singular, e.g. /k'uk'/ ‘peak’ and /q yotf'/ ‘armpit’.

Next we will examine some other aspects of Lezgian laryngeal phonology to motivate the view that [voice] is the default consonant. Haspelmath notes that there is a ban on word-final unaspirated stops. One rule which is relevant to determining the status of [voice] as a default feature is found in the rule Haspelmath calls Word-final Unaspirated Voicing. There is a regular alternation in all monosyllabic nouns between root-final voiceless unaspirated stop and the corresponding voiced stop word-finally. Compare the following:

(9)  |  Plural  | Abs. sg. | Gloss |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>jeper</td>
<td>jeb</td>
<td>string</td>
<td></td>
</tr>
<tr>
<td>gatar</td>
<td>gad</td>
<td>summer</td>
<td></td>
</tr>
<tr>
<td>pak*ar</td>
<td>pag*</td>
<td>side, rib</td>
<td></td>
</tr>
</tbody>
</table>

The voiced equivalents of /tʃ, ts, q/ are /ʒ, ʃ, ɸ/, respectively. (Most dialects of Lezgian have lost the historical *dʒ, *dz, *ɡ). Because all consonants must have a laryngeal node word-finally, there appears to be a feature default of [voice], which accounts for the alternation between voiceless unaspirated and voiced in word-final position.

Additional evidence for this is found in certain reduplicated imperatives. The final consonant of the stem is reduplicated as a voiceless unaspirated plosive after the morphemic /-u-/, as shown in (10):

(10)  | Masdar | Imperative | Gloss     |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>k-u-n</td>
<td>k-u-g</td>
<td>burn</td>
<td></td>
</tr>
<tr>
<td>kʰutu-n</td>
<td>kʰutu-z</td>
<td>pour</td>
<td></td>
</tr>
<tr>
<td>q-u-n</td>
<td>q-u-ʃ</td>
<td>fall (precipitation)</td>
<td></td>
</tr>
</tbody>
</table>

We have thus seen some evidence to suggest that the feature [voice] may be the default consonant in Lezgian. This may be due to the typologically unusual requirement that word-final consonants bear a laryngeal node. Because [voice] is the default consonant, laryngeal feature changing is not required. Instead, through delinking and default, we may preserve the more highly constrained system which prohibits direct feature changing (at least within laryngeal features).
Next, we will look at a case in which ejectives vary, apparently freely, between ejective and voiced. Coast Tsimshian (Dunn 1979), a Penutian language, has voiced, voiceless, and ejective plosives. The language apparently has free variation between ejectives and voiced stops. Regarding the ejective series, Dunn notes that 'glottalized segments often simplify by losing the glottalization and then becoming voiced' (1979:12). Dunn does not, however, provide evidence that this process is accomplished in two steps. He gives the following examples of the free variation:

(11)  

<table>
<thead>
<tr>
<th>Word</th>
<th>Gloss</th>
<th>Diminutive</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>k'wili</td>
<td>gwylii</td>
<td>'three (general number)'</td>
<td></td>
</tr>
<tr>
<td>q'asq'ooos</td>
<td>gasgoos</td>
<td>'crane, stork'</td>
<td></td>
</tr>
<tr>
<td>q'asq'adzn</td>
<td>gascadzn</td>
<td>'ants'</td>
<td></td>
</tr>
<tr>
<td>ts'awes</td>
<td>dzα?west</td>
<td>'salal'</td>
<td></td>
</tr>
<tr>
<td>gals'tap</td>
<td>caldzap</td>
<td>'town'</td>
<td></td>
</tr>
</tbody>
</table>

These examples are thus in free variation, which shows additional synchronic evidence for the alternation between ejectives and voiced stops.

Finally, an alleged case of ejective voicing is found in the Salishan language Tillamook, as reported by Edel 1939 and Reichard 1958-1960. These authors report that Tillamook has voiceless aspirated, “intermediate” (i.e., it alternates between voiceless unaspirated and voiced), and ejective plosives. In certain reduplicated forms such as the diminutive and verbal frequentative, the initial consonant of the reduplicated form shows a voiced member, while the root maintains its ejection. (Voiceless stops also show this alternation). For example,

(12)  

<table>
<thead>
<tr>
<th>Word</th>
<th>Gloss</th>
<th>Diminutive</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>t'iįǰ̂hù</td>
<td>man</td>
<td>dut'iįǰ̂hù</td>
<td>boy</td>
</tr>
<tr>
<td>ts'.q</td>
<td>to split</td>
<td>s-dzats'.q-æn</td>
<td>he split them</td>
</tr>
<tr>
<td>t'.tf</td>
<td>to shoot</td>
<td>daf-dít'tf-ën</td>
<td>he shot it till he killed it</td>
</tr>
</tbody>
</table>

Thompson and Thompson (1966) reanalyzed the language on more contemporary phonemic terms and in light of other Salishan languages. They proposed only two phonemic series, voiceless unaspirated and ejectives. Aspirates are treated as clusters of consonant plus /h/. The case of reduplicative voicing described above would currently be handled as delinking the laryngeal node of the initial ejective reduplicative consonant under pressure from the Obligatory Contour Principle, which resolves the violation through dissimilation. Thompson and Thompson also posit what would currently be called a postlexical rule in which plain obstruents (and presumably laryngeal-less segments, in my re-analysis) are ‘regularly partially or fully voiced in position directly before vowels’. Thus phonologically there is feature delinking, not feature changing. And for those ejectives that do get (partially) voiced, this is accomplished in two steps, delinking plus the postlexical rule which spreads [voice] regressively from a following vowel.

3. Diachronic ejective voicing

As mentioned above, in Gamkrelidze and Ivanov (1984, 1995) the only detailed evidence given for a diachronic change from ejective to voiced stop is found in the Nakh languages Chechen and Ingush. Gamkrelidze and Ivanov posit that Chechen and Ingush non-initial ejectives changed to voiced stops, while Bats (Tsvoa-Tush) maintains the Proto-Nakh ejectives in this position. I should note that this interpretation has not gone unchallenged by some Soviet scholars, and
most recently by Job (1989). However, the foremost Western scholar on the Nakh languages, Johanna Nichols (1993; see also her note in Gamkrelidze and Ivanov 1995:47), provides additional convincing evidence regarding loanwords and cognate Daghestican languages.

In initial position, all three stop series (ejective, aspirated, and voiced) correspond. In non-initial position, the Nakh voiceless stops all correspond. The voiced stops in Bats correspond to zero or a glide in Chechen-Ingush. The ejectives in Bats correspond to voiced stops in Chechen and Ingush, with the exception of uvular ejectives, which remain unchanged. Here follow some examples of this correspondence between voiced consonants in Chechen and Ingush, and ejectives in Bats:

<table>
<thead>
<tr>
<th>Chechen</th>
<th>Ingush</th>
<th>Bats</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>nab</td>
<td>nab</td>
<td>n'sap'</td>
<td>sleep</td>
</tr>
<tr>
<td>jad</td>
<td>jod</td>
<td>f'at'</td>
<td>knot</td>
</tr>
<tr>
<td>phagal</td>
<td>phagal</td>
<td>phak'al</td>
<td>hare</td>
</tr>
<tr>
<td>k'ezi</td>
<td>k'aza</td>
<td>k'ats'</td>
<td>puppy</td>
</tr>
<tr>
<td>ha3ar</td>
<td>ha3ar</td>
<td>hatf'ar</td>
<td>to look</td>
</tr>
<tr>
<td>juq'</td>
<td>juq'</td>
<td>juq'</td>
<td>middle</td>
</tr>
</tbody>
</table>

It is highly likely that the intervocalic deletion of the voiced stops (not shown here) and the voicing of ejectives could be conceived of as part of the same lenition process. It is possible, as David Odden (p.c.) has suggested, that the ejectives were deglottalized and that the resulting voiceless unaspirated stops were reinterpreted as voiced stops. Perhaps this is a more common phonetic path. Or perhaps there was simply direct spread of voicing from an adjacent vowel. But what is important for the Glottalic Theory is to establish that the change from ejective to voiced is both possible and plausible.

The implications of a change from ejective to voiced for the Ejective Model have also been noted by Colorusso (1981), who notes that in the Northwest Caucasian language Abaza, the Anatolian dialect has developed voiced variants of two morphemes in final position. (See also Lomtatidze and Klychev 1989:110).

<table>
<thead>
<tr>
<th>Standard Abaza</th>
<th>Anatolian Abaza</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. /s-tʃ'wa-p'/</td>
<td>/s-tʃ'wa-b/</td>
<td>'I sit, I am sitting'</td>
</tr>
<tr>
<td>b. /s-ʃ'w-j-t'/</td>
<td>/s-ʃ'w-j-d/</td>
<td>'I write, I am writing'</td>
</tr>
</tbody>
</table>

Apparently, however, this process occurred only in these two morphemes and was not a regular process.

So far, the examples we have seen do not provide evidence for the exact type of change needed by the Ejective Model; namely, a wholesale change of all ejectives, especially in initial position. There is one possible example of initial ejective voicing from the under-documented Hokan languages Jicaque-Subtiaba (Oltrogge 1977). Oltrogge notes that Proto-Jicaque-Subtiaba *t*' remains ejective in Jicaque while it voices initially in Subtiaba, but deglottalizes elsewhere. The only example provided by Oltrogge is given in (15):

(15a) *t'o?o[n]  
1'o?on 'to shut'  
Jicaque

(15b) *t'i  
1'i  
Jicaque

(15b) *t'i  
(-spa:)tu 'to chop'  
Subtiaba

(15b) *t'i  
Jicaque

(15b) *t'i  
Subtiaba
Thus, although this example looks promising for the Ejective Model, there must be more comparative work done before the sound change can be securely accepted. There are other sparsely documented changes from ejective to voiced in other languages discussed in Fallon (forthcoming). Other putative changes from ejective to voiced have also been reported, but these include cases such as Proto-Semitic (and Proto-Indo-European), mentioned above, in which the reconstructed forms are hotly contested. It is therefore important to try to seek independent cases to assess the plausibility of ejective voicing. Next we shall examine how ejectives became voiced in the course of loanword phonology.

4. Loanwords

Loanword adaptation can also shed light on phonological processes involving ejectives. Kartvelian languages provide interesting examples of several types of loanword adaptation. All Kartvelian languages have voiceless (aspirated), voiced, and ejective series of stops. Georgian loanwords (some themselves adapted from Greek) which contained two ejectives within a root were often changed by dissimilation to voiced consonants in the borrowing language. In the following examples, the initial ejective became voiced in Svan.

Svan (Schmidt 1991)

(16) Georgian Svan Gloss
k'ak'-ali gak' 'walnut'
p'ap' (Ming.) bap' 'priest'
p'et're bet're 'Peter'
kat'o gat' 'bran'

The process is presumably the result of an OCP-driven constraint against two ejectives within a root. In some cases, however, perhaps from a different period of loans, the second consonant voiced, as in the following example:

(17) Georgian Svan Gloss
tf'ak'i tf'âg 'mare' (Schmidt)

In still other examples, Svan ejectives apparently became voiced after preceding sibilants, as the following examples illustrate.

(18) Georgian Svan Gloss
laf'k'ari lafgâr army
skeli sgeli thick
škeri šger Black Sea rhododendron (Rogava 1982)

Georgian loanwords into the Indo-Iranian language Ossetic, spoken in the Caucasus, often dissimilated the first of two ejectives in a root (Rogava 1982), as we saw in Svan above.

(19) Georgian Ossetic Gloss
p'it'na bit'na mint
t'ik'i dik'î/dîf'î wineskin
p'ark'i bark'î/bartf'î small bag, little sack
k'ot'ofî got'ofî banočka (for bloodletting)
kotosora crested
In sum, loanword phonology provides rich examples of ejective voicing. Many of these are the result of dissimilation. It is unusual in these cases that the default consonant appears to be voiced.

5. Conclusion

In this paper, I have examined putative cases of ejective voicing, which is an important testing ground on two fronts. Phonologically, direct ejective voicing could have forced a re-examination of the assumption that laryngeal features cannot change automatically. This assumption falls out of restrictions imposed by phonological theory to try to constrain the different possible rule types. We have seen that all putative cases of ejective voicing may be explained by other means: by the spread of [voice] from vowels (Nguni languages) and from a particular morpheme (Venda), or by the treatment of voice as a default (Lezgian). This does not, however, deny the plausibility of the Ejective Model of the Glottalic Theory, since there were several examples of diachronic ejective voicing. The example from the Nakh language is good testimony, though perhaps it is a moot point whether it was the result of deglottalization and then voicing or simply the spread of voice from adjacent vowels. We saw additional evidence from Anatolian Abaza, though this was limited to two morphemes. Additional possibilities were found in Proto-Jicaque-Subtiaba, which is unfortunately not as robust an example as one would like. Finally, we have seen evidence from loanword phonology, often involving dissimilatory voicing of ejectives.

In sum, through an informed interplay between synchronic and diachronic phonology, scholars interested in the behavior of sounds can provide complementary views of the same phenomenon. In this case, examination of ejective voicing provides confirmation of current phonological theories which constrain types of phonological processes, and illuminates the diachronic typology of sound changes for the historical linguist.

Notes

*. I am grateful to David Odden for valuable discussion and detailed comments, though I have not always followed his suggestions. Funding for travel to the BLS conference was provided by the Ohio State University Department of Linguistics Language Files funds.

1. A fuller treatment of ejective voicing is found in Fallon (forthcoming), upon which this study draws. The phonetic evidence discussed at the conference presentation are not given here due to space limitations but may be found in Fallon (forthcoming).

2. David Odden (p.c.) notes that in most Southern Bantu languages (except Shona), there is no contrast between voiceless unaspirated stops and ejectives. He claims that the fact that nonaspirate voiceless stops are pronounced as ejectives in these languages is a phonetic fact and that there is no phonological evidence that these are true ejectives. Although I respect Odden’s expertise on Bantu matters, I prefer to posit underlying ejectives without convincing evidence to the contrary.
References


The Phonological Composition of Personal Pronouns: Implications for Genetic Hypotheses
Matthew J. Gordon
University of Michigan

1. INTRODUCTION. Personal pronouns\(^1\) can be found in virtually every language in the world. In many languages they are used with remarkable frequency and serve a variety of functions. Although linguists from almost every subfield have discussed the unique role that pronouns play, very little work has been done to determine what a typical pronoun or pronoun system looks like. This is particularly true with regard to their phonological make-up. While it may not be immediately apparent why we should be concerned with this gap in the research on pronouns, this issue has recently attracted the attention of historical linguists, in particular those involved in attempts to establish or refute claims of language relatedness.

Personal pronouns are, in many languages at least, more stable than other elements. They are reported to be more resistant to semantic change and less likely to be replaced through borrowing.\(^2\) Because of their putative stability, they, along with other 'core' vocabulary items, often figure as evidence in proposals for the genetic affiliation of languages. Recently, however, some questions have been raised about the use of pronoun data in such proposals, particularly when these proposals are arrived at through the method of 'mass lexical comparison' associated with Joseph Greenberg. This method draws its genetic conclusions primarily on the basis of shared similarities of form and meaning across languages. Unlike traditional approaches, Greenberg's method does not rely on systematic sound correspondences or comparative reconstruction. Because of this, many historical linguists believe mass comparison to be incapable of distinguishing similarities due to genuine inheritance from those due to chance or other non-genetic explanations. Opponents of Greenberg's classification of American Indian languages (1987) have been especially critical on this methodological point (see Campbell 1993, Thomason 1990). Although almost every aspect of Greenberg's 'Amerind' proposal has been questioned, some of the strongest opposition has been focussed on the pronominal evidence. Criticism of these data relies in part on some generally accepted but rarely tested notions regarding the phonological composition of pronouns. Specifically, it has been claimed that (1) pronouns (and other grammatical morphemes) tend to involve a limited number of segments from a language's phonemic inventory, (2) the same sounds reappear in pronouns from apparently unrelated languages, and (3) there is a non-arbitrary relationship between certain sounds and certain pronominal meanings. These claims seem to suggest that special factors operate in determining the phonological shape of pronouns and that one must be cautious, therefore, in offering pronoun data to support hypotheses of language relatedness.

This paper uses data collected from a cross-linguistic survey of pronouns to explore these claims. Following a brief methodological overview, each of the three claims is examined and tested. In discussing the results of these tests, the focus is on their implications for the problem of establishing genetic relationships.
2. METHODOLOGY. The present research was conducted as part of a broad study which investigated phonological as well as morphological, semantic and syntactic aspects of pronouns and pronoun systems. The goal of this project was not to establish a typology of pronouns (i.e. to define the limits of variability by outlining what is and is not possible), but rather to provide some guidelines as to what is typical of pronouns. This objective necessitated that many languages be considered and that the data from these languages be as independent as possible. Therefore, my sample included 62 languages selected to be both genetically and areally diverse.3 (see Appendix for a list of languages with their genetic affiliations and locations).

The pronominal data collected included both independent and bound morphemes. A form was taken to be pronominal if it functioned to mark the grammatical category of person.4 In addition to the actual pronominal forms, an inventory of phonemes was taken for each language in the sample. These inventories were used not only to compare the range of sounds available to each language but also to establish a uniform transcription system.

3. RESULTS AND DISCUSSION

3.1. LIMITED USE OF PHONEMIC INVENTORY. Pronouns are grammatical ('function'), as opposed to lexical ('content'), morphemes. They serve the grammatical function of marking *inter alia* person. Certain phonological properties are said to be characteristic of grammatical morphemes. We are most concerned here with the claim made by several authors that languages tend to use only a limited number of sounds from their total segment inventory in the formation of grammatical morphemes. For example, Floyd (1981) finds only seven (m n t th k r s) of Classical Greek's fifteen consonants occurring in inflectional forms and mentions similar constraints for Hebrew, German, Latin and English.5 If such restrictiveness is, in fact, a widespread phenomenon, historical linguists must use greater caution in judging similarities in grammatical forms because, with fewer sounds to choose from, there is a higher probability of random (i.e. non-genetic) resemblances across languages.6

I tested this hypothesis as it relates to pronouns by comparing the phonemic inventory of each language in the sample to the sounds appearing in the database of pronominal forms from that language; that is, for each language I counted the number of phonemes available in inventory and the number that appeared in the language's pronouns.7 A summary of the results of these comparisons is found in Table 1, which shows the range of variation and averages across the 62 language sample for inventory size, the number of phonemes used in pronouns, and the ratio of the number of phonemes used to the number in inventory.

<table>
<thead>
<tr>
<th></th>
<th>Inventory Range</th>
<th>Avg.</th>
<th>Used Range</th>
<th>Avg.</th>
<th>% Used Range</th>
<th>Avg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consonants:</td>
<td>7-88</td>
<td>22.2</td>
<td>5-15</td>
<td>9.4</td>
<td>10.7-100%</td>
<td>49.5%</td>
</tr>
<tr>
<td>Vowels:</td>
<td>3-17</td>
<td>7.1</td>
<td>3-10</td>
<td>5.1</td>
<td>23.5-100%</td>
<td>78.4%</td>
</tr>
<tr>
<td>All Sounds:</td>
<td>10-94</td>
<td>29.3</td>
<td>8-22</td>
<td>14.5</td>
<td>17.2-100%</td>
<td>54.4%</td>
</tr>
</tbody>
</table>

Table 1: Summary comparison of phonemic inventory size and number of phonemes appearing in pronominal forms from a sample of 62 languages. Figures indicate number of segments or percentages of phonemes used as a ratio of total inventory.
As Table 1 shows, languages vary greatly with regard to their phonemic inventories. The largest inventory was found in Tsaxur (NE Caucasian) with its 88 consonants and 6 vowels and the smallest in Piraha (Paezan) which has only 7 consonants and 3 vowels. Actually, these extremes suggest greater variation than there really is, as the standard deviation was 11.9 segments and over 77% of the languages have inventories in the range of 20-37 segments. The overall average inventory size was 29, which is in line with the average of 31 obtained by Maddieson (1984) with a much larger sample. Turning to the data on sounds used in pronoun forms, we see much less variation, with an average of 14.5 segments and a standard deviation of 3.3. Khmer made use of the greatest number of sounds (22), while both Mandarin and Piraha used only 8. In some instances there seems to be a direct connection between the number of sounds used and the size of the language's inventory; however, there are many exceptions. A linear regression analysis found inventory size to be an unreliable predictor of the number of sounds used (r=.193, f(1,60)=2.312, p=.134). It seems that many languages use approximately the same number of sounds in pronouns despite variation in their inventory size. This is confirmed, in part, by the wide range of the ratios in Table 1. Languages with large inventories tend to employ only a small percentage of their sounds (e.g. White Hmong 11/64 = 17.2%), while those with small inventories employ a much greater percentage (e.g. Seneca 16/16 = 100%). In either case, most languages (over 75% in this sample) use between 10 and 17 sounds in pronoun forms.

These results are intriguing and indicate that there may in fact be some type of limitation on the number of sounds appearing in pronouns. However, the consistency in these numbers suggests a more pragmatic explanation than the mysterious restriction implicit in the claim that languages make only limited use of their phonemic inventories for grammatical forms. First of all, pronouns, and grammatical morphemes in general, tend to be shorter (i.e. contain fewer segments) than other elements in a language. As a practical explanation for this phenomenon, one might suggest that such morphemes need to be short because they appear so frequently. Whatever the explanation, this tendency has the net effect of limiting the range of sounds used in pronouns by simply reducing the number of available slots. Similarly, for any given language, the number of sounds used is clearly influenced by the number of forms collected. In the present sample there was tremendous variation in the number of pronominal forms available in the languages, from as few as 8 to as many as 232 forms. Obviously, having fewer forms reduces a language's capacity to display its sounds. Not surprisingly, then, the correlation between the number of forms and the number of sounds used proved to be quite strong when calculated through linear regression (r=.43, f(1,60)=13.58, p<.0005).

It must be pointed out that these results do not necessarily disprove the notion that some special property of pronouns restricts the number of sounds used in their composition. I have merely suggested that the length and number of pronominal forms may provide another, less interesting, explanation. This explanation could be tested more thoroughly by comparing the phonemes used in pronouns with those used in non-grammatical morphemes of the same length and frequency. If languages can be shown to use a comparable repertoire of sounds in both contexts, it would be more difficult to maintain that pronouns are uniquely selective.
In sum, we have found that many languages do employ only a partial set of their phonemic inventories in pronoun forms. The original claim does not, however, hold universally, since there are languages (2 in this sample) that use all of the sounds available to them and others that use nearly all their sounds (e.g. Tiwi 17/18; N. Sierra Miwok 18/20). More importantly, however, even in those cases where a limited number of sounds appear in pronominal forms this fact appears to be a predictable consequence of relatively straightforward aspects of the data, such as the number of sounds in inventory, the length of the forms, or the number of forms available.

3.2. THE PHONEMES USED IN PRONOUN FORMS. In addition to pronouns using only a limited number of sounds, it is claimed that the same types of sounds tend to reappear in pronouns of unrelated languages (see e.g. Campbell 1993). This tendency, if true, is important for historical linguists to note because it provides a non-genetic explanation of phonetic similarities among comparable morphemes of different languages. The validity of this claim was tested with this set of 62 languages by simply counting the number of languages using each of the various sounds found in the database. In the present sample 95 different sounds appeared in pronominals: 26 vowels and 69 consonants. As expected, most (54%) of these sounds appeared in only one or two languages, while a smaller set of sounds was found to recur in many languages. The most frequently used sounds, those appearing in at least 10% of the languages, are listed in Table 2.

<table>
<thead>
<tr>
<th>Consonants</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>58</td>
<td>93.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>m</td>
<td>47</td>
<td>75.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k</td>
<td>44</td>
<td>71.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>42</td>
<td>67.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>j</td>
<td>33</td>
<td>53.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>w</td>
<td>27</td>
<td>43.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>h</td>
<td>25</td>
<td>40.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ñ</td>
<td>24</td>
<td>38.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>s</td>
<td>23</td>
<td>37.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>23</td>
<td>37.1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Vowels</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>61</td>
<td>98.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>i</td>
<td>56</td>
<td>90.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>u</td>
<td>43</td>
<td>69.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>o</td>
<td>35</td>
<td>56.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e</td>
<td>32</td>
<td>51.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: The most frequently used sounds in pronominal forms. Indicated are the number of languages (n) using each sound and the percentages they represent of the total sample of 62 languages.

The results in Table 2 support the claim that certain sounds can be expected to appear in pronoun forms, but again we must be careful in assessing the significance of this fact. When we examine the data, we find that there is nothing very
surprising about the sounds listed in Table 2. The vowels are all quite frequent in the languages of the world: 7 of the 10 (a, i, u, o, e, e, e) constitute Maddieson's list of most common vowel qualities (1984:125). Similarly, the consonants are also among the most commonly occurring in the world. In fact, they are all included by Maddieson in his 'modal' inventory, a set comprising the 20 most frequent consonants found in his 317 language sample (Maddieson 1984:12). These facts beg the question of whether the sounds found in Table 2 are the most frequent in the data simply because they are the most frequent cross-linguistically or whether their frequency is somehow relatable to their appearance in pronouns.

Although the simplest solution may be to explain the data in Table 2 as a reflection of general markedness patterns (as measured by cross-linguistic frequency), other factors may have some explanatory value and should not be overlooked. One such factor that deserves brief mention is the notion of perceptual salience. The basic idea is that since grammatical morphemes typically contain only a few segments, a language must get as much out of each segment as possible. One way of achieving this goal is to choose the most perceptually salient sounds in these contexts. Campbell (1993:4) offers this as an explanation for the common appearance of nasals in grammatical markers and cites Maddieson's (1984:70) observation that nasals are rarely subject to confusion with other types of consonants. There is a problem with this argument, however, which Campbell does not consider. While nasals as a class are very easy to distinguish from other sounds, and hence more salient, there is often confusion within the class differentiating one nasal from another (see studies cited in Maddieson 1984:70). If salience were the key factor, then we would expect languages to use only one nasal, but in fact this does not seem to be the case. Of the 60 languages in my sample that used nasals only 5 (8.3%) chose to employ a single nasal. The limitations of the salience argument are evident in the following pronominal paradigm from Katla, a Kordofanian language:

<table>
<thead>
<tr>
<th></th>
<th>Sing.</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st per.</td>
<td>ṇọŋ</td>
<td>ńẹŋ</td>
</tr>
<tr>
<td>2nd per.</td>
<td>ṇẹŋ</td>
<td>ńọŋ</td>
</tr>
<tr>
<td>3rd per.</td>
<td>ṇẹŋ</td>
<td>ńẹŋ</td>
</tr>
</tbody>
</table>

While these data may represent an extreme in terms of both nasal usage and overall perceptual similarity, it is not uncommon to find only minor phonetic differences distinguishing two or more forms within the same paradigm. This suggests that the need for grammatical forms to be clearly distinguished may not be as great as has been assumed. Perceptual salience may play some role in determining the phonological composition of pronouns, but there seems to be no reason to believe that its influence is greater in pronoun systems than in other areas of the language.

Despite this negative assessment, we should not rule out completely the possibility that certain other phonetic and phonological properties make some sounds particularly well suited to function in pronouns and other grammatical morphemes. Various types of explanations have been suggested, such as the claim that dentals and sonorants are preferred in bound morphology because they readily participate in clusters (see Callaghan 1991:53). These claims cannot be explored in this paper, but certainly merit further consideration.
3.3. CONNECTIONS OF SOUND AND MEANING. The final issue to be addressed involves the possibility of a non-arbitrary relationship between the meanings denoted by pronouns and the sounds used to compose them. Evidence for such a relationship would be of interest to historical linguists as another non-genetic way of accounting for similarities across languages. Thomason (1990:9) raises this issue in her criticism of Greenberg's use of pronoun data. Among her objections she lists the widespread 'affective and onomatopoetic uses of nasals' and mentions the examples of *mama* and *nana* as kin terms, but she gives no indication of how pronouns are affected by such sound symbolic patterns. Another of Greenberg's critics, Campbell, is much more explicit in his formulation of this claim. Borrowing an idea originally presented by Goddard (1986), Campbell suggests that the pattern of *n* marking first person (which is posited for Greenberg's Amerind family) may have a basis in child language. Since Campbell claims this argument has been misrepresented by Greenberg and his supporters, I quote it in full.\(^1\)

> Child-language expressions around the world abound in self-directed and other-directed words containing nasal consonants. The ultimate reason for this is the universal physical fact that a gesture equivalent to that used to articulate the sound *n* is the single most important voluntary muscular activity of the nursing infant. As Goddard (1986:202) points out, possibly this factor and the tendency for primary grammatical morphemes to consist of a single, unmarked (phonetically commonplace) segment account for the widespread appearance of *n*- in 'first-person' pronouns. Incidentally, in many societies, particularly among hunting and gathering groups, infants may continue to nurse until the age of five, sometimes longer, well into and beyond the age of language-acquisition. (Campbell 1993:6)

This statement suggests that Greenberg's Amerind evidence is merely a reflection of a universal preference for first person markers using *n*. If this preference really exists, we should expect to find *n* unusually frequent in the first person as opposed to the other two persons.

This claim can be investigated in the present sample by counting the number of languages employing *n* in each of the three persons. This approach presents a complication, however, due to the broad nature of the database. Since the database contains all the pronominal forms from the sample languages, counting all the data increases the amount of phonological material having nothing to do with person marking. For example, suppose a language has a different set of pronouns for each of several grammatical cases, and suppose that one of these cases is formed by adding the suffix *-n* to the pronominal stems for each person/number combination. If all the data are considered, that *n* would be counted for each person. In some instances this would simply add unnecessarily to the number of languages with a certain sound in each of the three persons; however, in many instances the net effect would be to obscure any real correlations that might otherwise be evident. Returning to our example, suppose the first person was the only one of the three in which an *n* appeared in the stem. This potentially interesting fact would be lost because *n* would be recorded for each person, due to its appearance as a case marker. For this reason a single set of pronouns was selected for analysis from each language. The goal in choosing these sets was to eliminate as much unnecessary phonemic overlap as possible, while preserving cross linguistic comparability. For most languages this meant choosing the least marked set of independent pronouns (e.g. the nominative or absolutive set); however, in a few instances bound person markers were selected, particularly when they were obviously segmentable from a common pronominal stem. Using these 'basic' sets
of pronouns, I measured the frequency of occurrence of all sounds across the three grammatical persons, with the idea that even if I found no significant pattern for \( n \), I might for other sounds. The results for the most frequent sounds (those appearing in over 50% of the languages (see Table 2)) were distributed as follows:

<table>
<thead>
<tr>
<th>Person</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>42</td>
<td>37</td>
<td>42</td>
</tr>
<tr>
<td>i</td>
<td>38</td>
<td>34</td>
<td>31</td>
</tr>
<tr>
<td>u</td>
<td>16</td>
<td>16</td>
<td>22</td>
</tr>
<tr>
<td>o</td>
<td>11</td>
<td>9</td>
<td>17</td>
</tr>
<tr>
<td>e</td>
<td>14</td>
<td>15</td>
<td>22</td>
</tr>
<tr>
<td>n</td>
<td>34</td>
<td>30</td>
<td>32</td>
</tr>
<tr>
<td>m</td>
<td>25</td>
<td>16</td>
<td>11</td>
</tr>
<tr>
<td>k</td>
<td>19</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td>t</td>
<td>10</td>
<td>14</td>
<td>24</td>
</tr>
<tr>
<td>j</td>
<td>11</td>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 3: Frequency of sounds occurring in each grammatical person for the basic sets of pronouns. Figures represent number of languages out of 62 total.

It is quite evident from the consistency of these figures that most sounds seem to be distributed independently of the category of person. The data in Table 3 were subjected to chi-square testing and with two exceptions none of the patterns was found to be significant beyond the level of .05. The exceptional cases are \( m \), which shows a preference for first person (chi-square = 8.06, df = 2, \( p = .018 \)), and \( t \), which correlates with third person (chi-square = 8.76, df = 2, \( p = .013 \)).

The implications of these results for the specific claim made by Campbell and Goddard are quite apparent: there is simply no indication of an overall preference for \( n \) in first person markers. More importantly, my data give some indications that the distribution of \( n \) within the members of the proposed Amerind group is somewhat unusual. This is clear when the data for \( n \) are separated to compare the 17 Amerind languages with the others:

<table>
<thead>
<tr>
<th></th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>( \chi^2 )</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amerind</td>
<td>10</td>
<td>4</td>
<td>8</td>
<td>4.47</td>
<td>.107</td>
</tr>
<tr>
<td>Others</td>
<td>24</td>
<td>26</td>
<td>24</td>
<td>.239</td>
<td>.887</td>
</tr>
</tbody>
</table>

Table 4: Frequency of \( n \) occurring in each grammatical person for the basic sets of pronouns from 17 Amerind and 45 non-Amerind languages.

Although neither of these patterns is statistically significant, it is obvious that the data are much more evenly distributed outside of the Amerind group. This observation should not, however, be taken as strong support for the validity of Amerind as a genetic group. We are dealing with the results of a superficial phonological analysis of a limited amount of data from a small set of languages. Even if we are willing to overlook these limitations, we still have no reason to claim that the patterns demonstrated are due to common genetic inheritance. Nevertheless, if further studies should confirm a significant connection of first person forms with \( n \) in the Amerind languages, then this fact would need to be
explained and it is clear that the situation cannot be attributed to some universal preference.

The fact that a significant correlation was found linking m and 'first person' raises some interesting questions. Perhaps Campbell and Goddard were essentially correct in their claim but simply chose the wrong nasal. After all, it seems plausible that m could be incorporated into their child language explanation. There is a complication, however, involving another of Greenberg's genetic proposals, namely his Eurasian group. For this group Greenberg has claimed that m is the characteristic marker of first person (see Greenberg 1991). Given the results discussed here we may wonder whether Greenberg has simply observed a general linguistic tendency that has nothing to do with genetic inheritance. To test this, we may simply compare the distribution of m across the three persons with and without the Eurasian languages in the sample. The original sample contains 9 members of the Eurasian group, five of which have m in the first person. With so few languages we cannot reliably determine the significance of the distribution within the proposed family, but a chi-square test was performed for the other languages:

<table>
<thead>
<tr>
<th></th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>X²</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eurasian</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Others</td>
<td>20</td>
<td>15</td>
<td>10</td>
<td>4.65</td>
<td>.098</td>
</tr>
</tbody>
</table>

Table 5: Frequency of m occurring in each grammatical person for the basic sets of pronouns from 9 Eurasian and 53 non-Eurasian languages

The original distribution of m was significantly weighted in favor of the first person, but when the Eurasian languages are not included, this correlation is no longer statistically significant. Still, the general pattern remains: m is more frequent in first person than in the other two. If more languages were sampled it seems likely that a significant pattern could be confirmed. This suggests that there is indeed some general preference for m to appear in first person forms. Whether this preference is ultimately connected to child language phenomena or to some other explanation remains an open question. Certainly, we will need to investigate new avenues to explain the other significant pattern, the one linking t and third person, although, at this point, I know of none that has been offered.

4. CONCLUSIONS. Various claims have been made to suggest that pronouns merit special consideration by linguists attempting to establish genetic relations among languages. Their phonological composition is said to derive from a restricted subset of the language's inventory of sounds, a small set of unmarked sounds is thought to be common to pronouns of diverse languages, and certain pronominal meanings are claimed to have a predilection for certain sounds. Each of these assertions has been investigated in this paper using data collected from a controlled sample of languages. Although the results have been mixed, they offer clear methodological implications for genetic linguistics.

The data considered here confirm that there is a tendency, though certainly not universal, for languages to employ a limited number of their phonemes in pronoun forms. This really should come as no surprise, however, since any subset of data from a language is likely to display only a partial amount of the complete phonemic inventory. Similarly, we are not surprised to find that the sounds which recur in the pronouns of language after language are also among the least marked, most
cross-linguistically frequent sounds in general. While these observations suggest that there are no mysterious factors operating to determine the phonological composition of pronouns, the uncovering of two significant patterns correlating sound and meaning (m and t with first and third person respectively) raises new questions. Further research is needed to confirm and seek explanations for these patterns, but at this point it is valuable to be aware of such connections and consider the possibility that they may result in some non-etymological similarities across languages.

Actually, the same can be said of the other findings as well. If, for whatever reason, the number of sounds used in pronouns is limited and certain sounds are generally more frequent, we can certainly expect a greater incidence of chance matchings that do not reflect any common ancestry. The methodological lesson to be drawn here is that the burden of proof in such cases is necessarily greater (cf. Ringe 1992). When considering cross-linguistic resemblances of form and meaning, we can never prove that the genetic hypothesis is the only explanation, the best we can hope for is to establish that it is the most reasonable explanation. In working with pronouns, it seems this job is all the more difficult.

**APPENDIX: Languages used in this study**

<table>
<thead>
<tr>
<th>Language (Location)</th>
<th>Genetic Affiliation*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alamblak (Papua-New Guinea)</td>
<td>Sepik-Ramu: Indo-Pacific</td>
</tr>
<tr>
<td>Amele (Papua-New Guinea)</td>
<td>Madang: Trans New Guinea; Indo-Pacific</td>
</tr>
<tr>
<td>Apalai (N Brazil)</td>
<td>Carib; Amerind</td>
</tr>
<tr>
<td>Arabic (Persian Gulf)</td>
<td>Semitic; Afro-Asiatic</td>
</tr>
<tr>
<td>Barasano (Colombia)</td>
<td>Tucanoan; Amerind</td>
</tr>
<tr>
<td>Bashkir (Bashkir Rep.)</td>
<td>Turkic; Altaic; Eurasian</td>
</tr>
<tr>
<td>Basque (N Spain)</td>
<td>Isolate (possibly Dene-Caucasian)</td>
</tr>
<tr>
<td>Bobo (Burkina Faso)</td>
<td>Mande; Niger-Kordofanian</td>
</tr>
<tr>
<td>Bukiyip (N Papua-New Guinea)</td>
<td>Torricelli; Indo-Pacific</td>
</tr>
<tr>
<td>Cahuilla (S California)</td>
<td>Uto-Aztec; Amerind</td>
</tr>
<tr>
<td>Canela-Kraho (Central Brazil)</td>
<td>Je; Amerind</td>
</tr>
<tr>
<td>Delaware (NE Coast America)</td>
<td>Algonquian; Almosan; Amerind</td>
</tr>
<tr>
<td>English (USA)</td>
<td>Germanic; Indo-European; Eurasian</td>
</tr>
<tr>
<td>Ewe (Togo)</td>
<td>South Central Niger-Congo; Niger-Kordofanian</td>
</tr>
<tr>
<td>Fula (Senegal)</td>
<td>W. Atlantic; Niger-Congo; Niger-Kordofanian</td>
</tr>
<tr>
<td>Georgian (Georgia)</td>
<td>Kartvelian</td>
</tr>
<tr>
<td>Greenlandic (W Greenland)</td>
<td>Eskimo-Aleut; Eurasian</td>
</tr>
<tr>
<td>Gujarati (W India)</td>
<td>Indic; Indo-European; Eurasian</td>
</tr>
<tr>
<td>Hmong (White) (N Thailand)</td>
<td>Miao-Yao; Austric</td>
</tr>
<tr>
<td>Hua (Papua-New Guinea)</td>
<td>East New Guinea Highlands; Indo-Pacific</td>
</tr>
<tr>
<td>Hualapai (Arizona)</td>
<td>Yuman; Amerind</td>
</tr>
<tr>
<td>Ika (Colombia)</td>
<td>Chibchan; Amerind</td>
</tr>
<tr>
<td>Itelmeny (Kamchatka Penn.)</td>
<td>Chukchi-Kamchatkan; Eurasian</td>
</tr>
<tr>
<td>Jacaltec (Guatemala)</td>
<td>Mayan; Amerind</td>
</tr>
<tr>
<td>Kanuri (Nigeria)</td>
<td>Western Saharan; Nilo-Saharan</td>
</tr>
<tr>
<td>Katla (S Sudan)</td>
<td>Kordofanian; Niger-Kordofanian</td>
</tr>
<tr>
<td>Khmer (Cambodia)</td>
<td>Mon-Khmer; Austroasiant; Austric</td>
</tr>
<tr>
<td>Kikuyu (Kenya)</td>
<td>Central Bantu; Niger-Kordofanian</td>
</tr>
<tr>
<td>Koasati (Louisiana)</td>
<td>Muskogean; Amerind</td>
</tr>
<tr>
<td>Korean (Korea)</td>
<td>Altaic (?); Eurasian</td>
</tr>
<tr>
<td>Kunama (Eritrea)</td>
<td>Nilo-Saharan</td>
</tr>
<tr>
<td>Lahu (China/Myanmar)</td>
<td>Lolo-Burmese; Tibeto-Burman; Sino-Tibetan</td>
</tr>
<tr>
<td>Language</td>
<td>Geographic Region</td>
</tr>
<tr>
<td>----------</td>
<td>------------------</td>
</tr>
<tr>
<td>Maasai</td>
<td>Kenya</td>
</tr>
<tr>
<td>Mandarin</td>
<td>China</td>
</tr>
<tr>
<td>Margi</td>
<td>Nigeria</td>
</tr>
<tr>
<td>Mari</td>
<td>(S. Russia)</td>
</tr>
<tr>
<td>Mba</td>
<td>Zaire</td>
</tr>
<tr>
<td>Nama</td>
<td>SW Africa</td>
</tr>
<tr>
<td>Nasiol</td>
<td>Bougainville</td>
</tr>
<tr>
<td>Ngandi</td>
<td>N Australia</td>
</tr>
<tr>
<td>Ngarinj</td>
<td>(NW Australia)</td>
</tr>
<tr>
<td>Olcha</td>
<td>(E Siberia)</td>
</tr>
<tr>
<td>Ometo</td>
<td>(W Ethiopia)</td>
</tr>
<tr>
<td>Paiman</td>
<td>(Taiwan)</td>
</tr>
<tr>
<td>Paumari</td>
<td>(Brazil)</td>
</tr>
<tr>
<td>Piraha</td>
<td>(Brazil)</td>
</tr>
<tr>
<td>Pomo</td>
<td>(SE) (California)</td>
</tr>
<tr>
<td>Quechua</td>
<td>(Huallaga) (Peru)</td>
</tr>
<tr>
<td>Rawa</td>
<td>(Papua-New Guinea)</td>
</tr>
<tr>
<td>Samoan</td>
<td>Samoa</td>
</tr>
<tr>
<td>Sarcee</td>
<td>(Alberta)</td>
</tr>
<tr>
<td>Seneca</td>
<td>(New York State)</td>
</tr>
<tr>
<td>SierraMiwok</td>
<td>(N California)</td>
</tr>
<tr>
<td>Squamish</td>
<td>(NW Coast America)</td>
</tr>
<tr>
<td>Tamil</td>
<td>(S India)</td>
</tr>
<tr>
<td>Tiwi</td>
<td>(N Australia)</td>
</tr>
<tr>
<td>Tsaxur</td>
<td>(N Azerbaijan)</td>
</tr>
<tr>
<td>Urubu-Kaapor</td>
<td>(NE Brazil)</td>
</tr>
<tr>
<td>Watjari</td>
<td>(W Australia)</td>
</tr>
<tr>
<td>Yay</td>
<td>(N Vietnam)</td>
</tr>
<tr>
<td>Yukaghir</td>
<td>(NE Siberia)</td>
</tr>
<tr>
<td>Zulu</td>
<td>(South Africa)</td>
</tr>
<tr>
<td>Nilotic</td>
<td>Nilo-Saharan</td>
</tr>
<tr>
<td>Sinitic</td>
<td>Sino-Tibetan</td>
</tr>
<tr>
<td>Chadic</td>
<td>Afro-Asiatic</td>
</tr>
<tr>
<td>Finnic</td>
<td>Finno-Ugric; Uralic; Eurasian</td>
</tr>
<tr>
<td>North Central Niger-Congo; Niger-Kordofanian</td>
<td></td>
</tr>
<tr>
<td>Khoisan</td>
<td></td>
</tr>
<tr>
<td>East Papuan; Indo-Pacific</td>
<td></td>
</tr>
<tr>
<td>Gunwinyguan; Australian</td>
<td></td>
</tr>
<tr>
<td>Woror; Australian</td>
<td></td>
</tr>
<tr>
<td>Tungusic; Altaic; Eurasian</td>
<td></td>
</tr>
<tr>
<td>Omotic; Afro-Asiatic</td>
<td></td>
</tr>
<tr>
<td>Paimanic; Austro-Asiatic</td>
<td></td>
</tr>
<tr>
<td>Arawakan; Amerind</td>
<td></td>
</tr>
<tr>
<td>Paezan; Amerind</td>
<td></td>
</tr>
<tr>
<td>Northern Hokin; Amerind</td>
<td></td>
</tr>
<tr>
<td>Andean; Amerind</td>
<td></td>
</tr>
<tr>
<td>Finisterre-Huon; Trans-NG; Indo-Pacific</td>
<td></td>
</tr>
<tr>
<td>E.Malayo-Polynesian; Austro-Asiatic</td>
<td></td>
</tr>
<tr>
<td>Athabaskan; Na-Dene</td>
<td></td>
</tr>
<tr>
<td>Iroquoian; Amerind</td>
<td></td>
</tr>
<tr>
<td>Penutian; Amerind</td>
<td></td>
</tr>
<tr>
<td>Salishan; Almosan; Amerind</td>
<td></td>
</tr>
<tr>
<td>Dravidian</td>
<td></td>
</tr>
<tr>
<td>Australian</td>
<td></td>
</tr>
<tr>
<td>Lezgian; NE.Caucasian; North Caucasian</td>
<td></td>
</tr>
<tr>
<td>Tupi; Amerind</td>
<td></td>
</tr>
<tr>
<td>Pama-Nyungan; Australian</td>
<td></td>
</tr>
<tr>
<td>Tai; Daic; Austric</td>
<td></td>
</tr>
<tr>
<td>Uralic; Eurasian</td>
<td></td>
</tr>
<tr>
<td>Central Bantu; Niger-Kordofanian</td>
<td></td>
</tr>
</tbody>
</table>

* For consistency, information on genetic groupings was taken from Ruhlen (1991), although many groups listed therein are controversial.

**NOTES**

1. The term "pronoun" as used throughout this paper refers only to personal pronouns, not to other types (e.g. demonstrative, interrogative), and actually denotes various types of person markers as is explained below.

2. In Dolgopolsky's list of the fifteen semantic values "most impervious to change" the first and second person markers are ranked first and third respectively (Dolgopolsky 1964).

3. Although care was taken to ensure genetic diversity, it would be impossible to select 62 languages that all linguists would agree are not demonstrably related. Besides including several members of such long-range proposals as Nostratic and Amerind, in a few instances the sample does contain two members of well-established language families; e.g. Indo-European, Bantu, Austro-Asiatic. However, even in these cases, an effort was made to select languages as diverse as possible within the family; hence, for example, English and Gujarati were used from Indo-European.

4. This broad definition presents problems of comparability in some languages, particularly in East and Southeast Asia, where "pronouns" are clearly derived from nouns. In such cases I had to rely on the author's judgement as to which forms function comparably to pronouns in Western languages.

5. Although Floyd clearly refers only to inflectional elements, Goddard and Campbell (1994) suggest this phenomenon obtains more generally in "primary grammatical morphemes". As this database contains both bound and free forms, my results will apply to this broader formulation.
This assumes that languages are drawing from roughly the same set of sounds to begin with, a claim taken up below.

Although our focus here is the number of sounds appearing in pronoun forms, actually, for many languages these pronominal data contain a variety of grammatical morphemes (e.g. case endings, tense markers etc.).

Everett (1986) claims, apparently correctly, that Piraha has the smallest recorded phonemic inventory in the world.

Maddieson prefers to talk about a range of segments, 20-37, as being typical, because his mean value of 31 is somewhat positively skewed (1984:7).

The relationship between the number of pronominal forms in a language and the number of sounds appearing in those forms is complicated by the fact that in some cases the forms are clearly related across or even within different paradigms (e.g. *me/my/mine*) and in others their is no relationship. In the former case we would expect much greater overlap in terms of the sounds used.

In a limited number of cases phonetically distinct sounds were counted together for the purposes of comparison. This was only done for three classes of sounds and only with languages in which the particular phonetic difference was noncontrastive. Thus, a language's voiceless aspirated stops were included in a count of voiceless unaspirated stops, if the language did not have two separate voiceless series. Similarly, dental sounds were counted together with alveolars, unless a language distinguished these two places of articulation. Finally, various "r-sounds" (including */r/, */s/ and */l/) were counted with alveolar flaps, unless a language had more than one.

35 appeared only once. 17 appeared twice.

The only difference between this list and Maddieson's is that the latter includes */l/, which was quite infrequent in my sample, appearing in the pronouns of only 2 languages (3.2%).

In fairness to Campbell it should be pointed out that this explanation appears at the end of a long discussion of a variety of objections and he himself describes this claim as "a minor case, certainly not one I favor most" (1993:footnote 11).

Since their argument was raised in the context of the Amerind discussion, this may have led Campbell and Goddard to focus only on *n*.

As few as two more languages using *m* in the first person would produce a significant result beyond the level of .05.

In his comments following the oral presentation of this paper, John Ohala suggested that the sucking reflex of infants most resembles a velar articulation and, therefore, that if any nasal is to be especially frequent in pronouns due to the influence of child language, we should expect that nasal to be an *[ŋ]*.

REFERENCES

[Due to the limited space available, the data sources for the 62 languages on which this study is based are not listed here.]


Vowel phonotactic positions in Australian Aboriginal languages
Philip Hamilton
University of Toronto

0. Introduction.
In the definitive overview of Australian phonotactics in Dixon (1980:159-178) the focus is on the phonotactic positions of consonants. Dixon uses the labels for the consonantal phonotactic positions within the word shown in (1).

(1) Descriptive word templates in Australian (Dixon 1980).

\[ \begin{align*}
C_{\text{init}} & \text{VC}_{\text{inter}} V(C_{\text{fin}}) \\
C_{\text{init}} & \text{VC}_1 \text{C}_2 V(C_{\text{fin}})
\end{align*} \]

Each of the five different consonantal positions show distinct phonotactic patterns. As discussed by Dixon, consonants in word-initial position \(C_{\text{init}}\) and consonants in the second position of a word-internal heterorganic cluster \(C_2\) show similar, though distinct, phonotactic constraints; similarly, word-final \(C_{\text{fin}}\) consonants and consonants in the first position of a heterorganic cluster \(C_1\) show similar phonotactics. Dixon observes that in the former two positions labial and dorsal segments have a preferential distribution over coronals, while the opposite pattern obtains in the latter two positions. Intervocalic position, \(C_{\text{inter}}\) stands out as the only structural position in the word where consonantal contrasts are unconstrained and all of the consonants in the inventory of the language in question are in contrast.

What is lacking in Dixon's discussion of Australian phonotactics is the fact that there are distinct vocalic phonotactic positions within the word. This is the topic of this paper. As I will show, there are two distinct positions of neutralisation. These are word-initial position, discussed in section 1, and word-final position, discussed in section 2. Metrical anomalies of word-edge vowels are discussed in section 3, and these are offered as further evidence of the special phonotactic status of word-initial and word-final vocalic positions. Finally, there is a unique vowel position in the word which stands out as the site of maximal contrast. This is the first post-consonantal vowel, a position which I label the Focus Vowel. In this position all of the vocalic contrasts present in the language are active, whether they are permitted in non-focus positions in the word or not. Discussion and motivation of the Focus Vowel position is in section 4, followed by concluding discussion in section 5.

1. Word-initial vowel phonotactics.
It is the majority pattern in Australian Aboriginal languages that words are required to begin with a consonant, as expressed in the word templates given in (1). There is, however, a significant minority of languages which allow vowel-initial words. In some of these languages the vowel-initial words
constitute a very small portion of the lexicon while in others the lexicon is either predominantly or exclusively vowel-initial. Aboriginal languages which permit vowel-initial words normally do not permit word-medial vowel-initial syllables.

Among languages which permit vowel-initial words, it is frequently the case that the set of vowels which contrast in initial position is only a subset of the vowel inventory of the language. This is a very common, almost typical, pattern, especially (though not exclusively) in the languages where only a minority of the lexicon is V-initial. It is certainly the case that no language has more vowel contrasts in word-initial than in other phonotactic positions. In cases of a disparity in the vowels in initial and non-initial positions, low vowels enjoy a favoured status. This is shown in the fact that they are more likely to be permitted than non-low vowels in initial position. The statistical predominance of the low vowel mentioned for Alyawarra and Atampaya Uradhi mentioned above reinforces this observation. A vowel length contrast is typically also not licensed in word-initial position.

The language Mbabaram (Dixon 1991) is an example. This language has six distinctive vowel qualities as well as contrastive length. The Mbabaram vowel phonemes are displayed in (2.a) along with some data illustrating vocalic contrasts (2.b). In spite of the rich inventory of vowels which Mbabaram enjoys, only one of these vowels, with no length contrast, is permitted in word-initial position (2.c). This is the vowel /a/.

(2a. Mbabaram

\[ \begin{align*}
\text{i, i:} & \quad \text{u, u:} \\
\text{e, e:} & \quad \text{a, a:} \\
\text{b. no:mbi} & \quad \text{big red wallaroo} \\
\text{gey} & \quad \text{sand goanna} \\
\text{mura:l} & \quad \text{cold} \\
\text{nambu:r} & \quad \text{big brown snake} \\
\text{guridal} & \quad \text{eagle hawk} \\
\text{gi} & \quad \text{tree} \\
\text{c. albo} & \quad \text{egg} \\
\text{alngida} & \quad \text{downhill} \\
\text{ang*ayir} & \quad \text{hungry} \\
\text{araman} & \quad \text{woomera} \\
\text{*a:CVC, *i:CVC, *o:CVC etc.}
\end{align*} \]

The bias in favour of the low vowel is typical. The general pattern in languages that allow V-initial words is that it is only the low vowel or vowels in the inventory which may occur initially. In languages with either the three-vowel system, as in Anindilyakwa (3.a), and the five-vowel system, such as Marra and others (3.b), the single low vowel is often the only one permitted.
In four vowel inventories, word-initial position is often restricted to either only one (3.c) or both (3.d) of the low vowels.

(3)a. Anindilyakwa (Leeding 1989:62)
Vowels which contrast C__C  Vowels which contrast #__C
i    u                      a
    a

b. Marra (Heath 1981); Ngalakan (Merlan 1983:11); Ngandi (Heath 1978); Warndarang (Heath 1980b).
Vowels which contrast C__C  Vowels which contrast #__C
i    u                      a
e    o                      a
    a
c. Tiwi (Lee 1987:30)
Vowels which contrast C__C  Vowels which contrast #__C
i    u                      a
    o                      a
d. Alawa (Sharpe 1972:22); Ngan'gityemneri (Reid 1990:69).
Vowels which contrast C__C  Vowels which contrast #__C
i    u                      a
    a                      a
    e

Another example of the special status of the low vowel is shown in Nunggubuyu. In this language the length contrast is licensed for only the low vowel in word-initial position, although the two non-low vowels show contrastive length in other positions (3.e). A combination of preference for low vowels as well as the lack of a length contrasts is shown in the word-initial vowel phonotactics of Kitja (3.f). Kitja has four contrastive vowel qualities, as well as a length contrast active for only the low vowel /a/. But initially in Kitja only the low vowel occurs, and its length contrast is neutralised.

(3)e. Nunggubuyu (Heath 1984:17)
Vowels which contrast C__C  Vowels which contrast #__C
i, i: u, u: i    u
    a, a:
    a, a:
f. Kitja (Taylor & Taylor 1971:107)
Vowels which contrast C__C  Vowels which contrast #__C
i    i    u
    a, a:
    a

An extreme case of positional neutralisation of phonemic contrasts for vowels in initial position is the Mpaikwithi dialect of Anguthimri. This language
contrasts seventeen vowels, constituting the largest vowel inventory in Australia. In V-initial words, however, only the cardinal vowels /i, a, u/ contrast in initial position (3.g).

(3g) Mpakwithi Anguthimri (Crowley 1983)

Vowels which contrast C_C
i, i:, i, ü

Vowels which contrast #_C
u, u: i

To summarise to this point, in many Australian Aboriginal languages which permit vowel-initial words only a subset of the phonemic vowel contrasts exploited in the language are licensed initially. This is the case for both vocalic quality and quantity contrasts. In the case of quality neutralisation, there is strong preference for low vowel(s) in this position.

2. Word-final vowel phonotactics.
The other word-edge, word-final position, is also unable to license the full range of vowel contrasts that are permitted medially. A comparison of Australian languages reveals that in this position the quantity contrast is more robustly impoverished than the quality contrasts. For example, Kuuku Yalpu has the vowel inventory shown in (4.a), which includes contrastive length as demonstrated by the (near-) minimal set in (4.b). The length contrast is permitted in all syllables (4.c), including the word-final syllable in consonant-final words (4.d), but not for vowels in word-final position (4.e).

(4a) Kuuku Yalpu

Vowels which contrast C_C
i, i: u, u: i u

Vowels which contrast C_#
a, a: a

b. ṇaṭina find
naṭina laugh
yaṭina call out

c. tulu grass tree
wiːmumu large number of ants
wiyaːna another
tawuraːlu with a knife
muːmaːna rub

d. kulaːn possum

e. tulu, *tuluː

Both Atampaya Uradhi (5.a) and Mpakwithi Anguthimri (5.b), which have relatively rich vowel inventories, fail to license vowel length in word-final
position, similar to Kuuku Ya?u. These languages also do not contrast the full range of lexical vowel qualities word-finally.

(5)a. Atampaya Uradhi (Crowley 1983)
   Vowels which contrast C_C  Vowels which contrast C_#
   i, i:  u, u:  i  u
   e, e:  a, a:  a

b. Mpakwithi Anguthimri (Crowley 1981)
   Vowels which contrast C_C  Vowels which contrast C_#
   i, i:, i, ü  u, u:  i, i  u
   e, e:, ē, ö  o  e, ē
   æ, æ:, æ  a, a:, ā  a, ā

From the discussion of vowel phonotactics in these two sections, and the discussion of consonant phonotactic positions by Dixon 1980, it is clear that the word-edge positions, for both consonants and vowels, are unable to license the full set of phonemic contrasts permitted in the language.

3. Metrical considerations in the special status of word-edge vowels.
   Another symptom of the fact that word-edge positions are defective for vowels is that in many languages word-initial and -final vowels are metrically anomalous. In these languages word-edge vowels are unable to bear stress. The majority pattern in Australian languages is that primary stress falls on initial syllables, but in many languages which allow vowel-initial words, stress falls on the first post-consonantal vowel. Put another way, stress fall on the first syllable with an onset; word-initial vowels are unable to carry stress. Most of the languages of the Arandic group are like this, including Alyawarra (Yallop 1977) and Andegerebenha (Breen 1977:374). Stress falls on the first syllable in C-initial words (6.a) and on the second syllable in V-initial words (6.b). Data here are from Alyawarra.

(6)a. kwáta  water, rain, urine
      riña  3.sg.ACC
b. ațâ  1.sg.ERG
      ampá  child
      ilipa  traditional stone axe
      arákiṭa  mouth

The same pattern obtains in Mbabaram (Dixon 1991:360). Stress in Mbabaram is quantity-sensitive—long vowels necessarily attract primary stress—and, as discussed above, contrastive vowel length is not licensed in word-initial position. However, even in words with no long vowels stress never falls on a word-initial vowel (7.a).
Word-initial vowels are doubly-defective in Mbabaram: in this position all of the contrasts for vowel quality are neutralised as well as the length contrast, in addition to the fact that the position is unable to carry stress.

The parallel pattern for word-final vowels being unable to bear stress is attested in Wemba-Wemba (Hercus 1969:23) and Wergaia (Hercus 1969: 118). In these languages primary stress falls on the initial syllable and secondary stress falls on the third syllable in three- and four-syllable words. Therefore word-final syllables receive secondary stress in trisyllabic words, but only when the word is consonant-final (8.a). In trisyllables which end in a vowel, the final vowel is unable to bear secondary stress (8.b). Data here are from Wemba-Wemba.

(8)a. dînəŋək his foot
b. bûrganda I am sighing
   *bûrgandə

A similar pattern is attested in Djabugay (Patz 1991). In Djabugay stress assignment is sensitive to the quantity of the nucleus. This is shown by the fact that stress is normally on the initial syllable in words with only short vowels (9.a) but will be on a non-initial syllable in words with an internal long vowel (9.b). In consonant-final words a long vowel in the final syllable will attract stress as well (9.c), but a word-final long vowel does not attract stress, and thus is treated as if it were short (9.d). This pattern gives rise to the stress alternations in suffixed forms, where adding a suffix to a form with a final long vowel makes it visible to the quantity sensitive stress rule and thus allowing it to attract stress (9.e).

(9)a. bîna ear
    búluṟu name of creator figure
gúruŋga kookaburra
b. wûruːril to take out
    bundá:ra cassowary
c. gúlaj: these
guqá: these
d. dîna: foot
    wûru: river
e. dînaj:-la foot-LOC/INST
    wûrû:-la river-LOC
This pattern of word-edge vowels being metrically defective reinforces the segmentally defective nature of vowel positions at word-edges that I demonstrated earlier in the previous two sections.

4. The Focus Vowel.
There is one more structural position which is singled out as having special phonotactic status in Australian languages, and this is the vowel following the first consonant (or consonant cluster) of the word. In languages with only C-initial words this always corresponds to the nucleus of the first syllable. In languages which allow vowel-initial words, the nucleus of the first onsetted syllable has this status. I will refer to this special vocalic position as the Focus Vowel. This position has already been singled out as being the position where word-stress is assigned in the majority of Australian languages. (I will discuss the relationship between stress and the Focus Vowel in the following section.)

Ritharngu (Heath 1980a) is a C-initial language and therefore the Focus Vowel corresponds to the nucleus of the initial syllable. Ritharngu contrasts long and short vowels, but only in initial syllables (Heath 1980a:11). There is no prefixal morphology in this language, and so word-initial syllables are also root-initial. There are instances of roots with an underlying long vowel ending up in a non-initial syllable in compounding and reduplication. In such cases the second long vowel is shortened (Heath 1980a, 12-13) (10.b).

(10a). Ritharngu
Focus Vowel position Non-Focus Vowel position
i, iː u, uː i u
a, aː a
b. daːra- to stand
duːdaːm?-dara- to bend over
daːra?-dara- all to stand

Just as a vowel length contrast is often only attested in the Focus Vowel, some languages also show vowel quality contrasts only in the Focus Vowel. This is attested in the Cape York initial-dropping language Kuku-Thaypan (Rigsby 1976). The vowel inventory for this language is shown in (11). In C-initial words these vowels contrast in the first syllable and in V-initial words they contrast in the second syllable. Vowels in other structural positions in words, including word-initial vowels, are non-contrastive. These are written as /a/.

(11a). Kuku-Thaypan
Focus Vowel position Non-Focus Vowel position
i i u
e o
æ a o a
It will be recalled from discussion in the preceding section that the Focus Vowel is the invariant location of primary stress in many Australian languages. This correlation is likely not completely coincidental, and I return to the topic of the relationship between the Focus Vowel and stress in this section. The question is what exactly is the nature of the relationship between stress and the Focus Vowel.

It could be argued that the special status of the Focus Vowel, as the position of maximum vocalic contrast, derives trivially from the fact that this is the position of primary stress in the majority of Australian languages. This could be accomplished by the neutralisation rule expressed in (12), acting as a morpheme structure rule:

(12) Neutralise specific vowel contrasts in unstressed syllables.

In Ritharngu stress is invariant, always on the Focus Vowel. Therefore the location of the vowel length contrast in this language could be accounted for with reference to the location of word-stress by neutralising the vowel length contrast in unstressed syllables. Note that this type of rule is independently required in Ritharngu as witnessed in the alternations in (10.b). Likewise the Kuku-Thaypan vowel phonotactics could be accounted for by a rule such as (12) neutralising all vowel quality contrasts in unstressed syllables.

One potential problem with such a rule would be that rules should not be able to refer to the absense of a feature such as stress. To make allowance for this it could be assumed that a well-formedness condition exists requiring specific vowel contrasts to be licensed by stress (see Dresher & van der Hulst 1993). This condition could then act as a morpheme structure constraint as well as trigger rule (12) in morphologically complex forms. Either way, the empirical content of this rule in Ritharngu would be the attested location of the length contrast only in the stressed syllable. In this type of approach there is a direct relationship between the Focus Vowel and the location of stress, with the special phonotactic status of the Focus Vowel deriving trivially from metrical structure.

The advantage of this type of approach is that it collapses two positions, the Focus Vowel and the location of word-stress, into one, and thus avoids resorting to coincidence to explain the intersection of the Focus Vowel and the location of stress. But the special status of the Focus Vowel is independently motivated in Australian, and I will demonstrate this in Warrgamay (Dixon 1981). The same pattern discussed here for Warrgamay is also attested in Nyawaygi (Dixon 1979). Warrgamay is a C-initial language, and therefore the Focus Vowel always corresponds to the nucleus of the word-initial syllable. Warrgamay is like Ritharngu in that it has a length contrast for vowels (13.b) which is only attested in the Focus Vowel position (Dixon 1981). In contrast with Ritharngu, however, which has an invariant pattern of word-
stress on the Focus Vowel, Warrgamay stress varies predictably between the first and second syllables. The facts are as follows. In words with a long vowel (which can only occur in the initial syllable), this receives stress (13.c). Stress is also initial in words with an even number of syllables (13.d), whether the first vowel is long or short. The variation from the strict initial stress pattern comes in words with no long vowel and an odd number of syllables, in which case stress falls on the second syllable (13.e). This produces alternations where stems with an even number of syllables take a mono-syllabic suffix triggering shift of stress from the first to the second syllable (13.f).

(13)a. Warrgamay  

<table>
<thead>
<tr>
<th>Focus Vowel position</th>
<th>Non-Focus Vowel position</th>
</tr>
</thead>
<tbody>
<tr>
<td>i, iː</td>
<td>u, uː</td>
</tr>
<tr>
<td>a, aː</td>
<td>a</td>
</tr>
<tr>
<td>b. ganda-</td>
<td>to burn, cook</td>
</tr>
<tr>
<td>gaznda-</td>
<td>to crawl</td>
</tr>
<tr>
<td>c. múːba</td>
<td>stone fish</td>
</tr>
<tr>
<td>gíːbaːra</td>
<td>fig tree</td>
</tr>
<tr>
<td>d. báda</td>
<td>dog</td>
</tr>
<tr>
<td>gídawuːlu</td>
<td>freshwater jewfish</td>
</tr>
<tr>
<td>e. gagára</td>
<td>dillybag</td>
</tr>
<tr>
<td>d̀urágay-miri</td>
<td>Niagara vale-FROM</td>
</tr>
<tr>
<td>f. múŋan</td>
<td>mountain</td>
</tr>
<tr>
<td>múŋan-da</td>
<td>mountain-LOC</td>
</tr>
</tbody>
</table>

A vowel length neutralisation rule conditioned by stress, as in (12), as an account of the distribution of the vowel length contrast is impossible for Warrgamay. First, since the second syllable in trisyllabic roots receives stress, and the neutralisation rule (12) by definition cannot apply until after the stress rule, a hypothetically possible length contrast would be licensed by stress in this position. Therefore this approach leaves unexplained why there is no length contrast for the second syllable in trisyllabic roots. Second, since stress is obviously quantity-sensitive in Warrgamay, hypothetical long vowels in non-initial syllables should attract stress and then be protected from the neutralisation rule. It is clear that the phonological pattern in Warrgamay is that there are no long vowels in non-initial syllables before the application of the stress rule. Therefore the special phonotactic status of the Focus Vowel position is motivated in Australian Aboriginal languages independently of information derived solely from metrical structure.

5. Concluding discussion.
There are interesting differences and similarities between the phonotactics of consonants and vowels in Australian. First, for consonants the position of
maximum contrast is intervocalic position, \( C_{\text{inter}} \) while the position of maximum contrast for vowels is the Focus Vowel. The difference here is that a word may contain more than one or even several \( C_{\text{inter}} \) slots, but a word can only have one Focus Vowel. This difference derives from the fact that \( C_{\text{inter}} \) is defined strictly in segmental terms—the vocalic environment—while the Focus Vowel is defined in relation to a unique morphological boundary.

An important similarity is that for both consonants and vowels word edges are positions of neutralisation. Trubetzkoy observes that both word-edges as positions of neutralisation, and adduces evidence from both consonant and vowel phonotactics to demonstrate this (Trubetzkoy 1939/1969:235). Acoustic phonetics may provide some insight into why consonants are neutralised at word edges, particularly in terms of their place contrasts. In acoustic terms, the place of articulation of a consonant is signalled by the contour in the formant patterns in the adjacent vowel(s). The fact that vowels play a crucial role in the saliency of consonantal contrasts provides an acoustic account of why consonants are restricted at word-edges (as well as in consonant clusters) while consonants in intervocalic position do not. The full range of possible phonological contrasts occur in the phonotactic position where the contrasts are most salient, i.e., between vowels. In environments where phonological contrasts may be less salient, predictability is encoded into the system constraining the set of permitted contrasts. All of this relates to recent work by Donca Steriade (1993, 1994) indicating that positions where phonological contrasts are not exploited are positions where the acoustic cues for those contrasts are relatively weak. Complementary work by Ohala and others (see Ohala 1990 and references therein) indicates that phonological contrasts undergo historical reduction in environments where they are not acoustically salient. Similarly, recent work in the perception of vocalic contrasts suggests that the consonantal context play an important role. For example, vowels in a /CVC/ context are more successfully identified than in a /\#V\#/ context. (See the overview in Rosner & Pickering 1994:319-331). Therefore a perceptual deficit inherent in word-edges may be responsible for the neutralisation of both consonantal and vocalic contrasts in edge positions.

One final point is that there is a fair amount of overlap between the vowel phonotactic positions illustrated here for vowels in Australian languages and those attested in languages outside of Australia. Telugu contrasts five vowel qualities plus contrastive length, /i, i, e, e, a, a, o, o, u, u/, except in word-final position where only four of the qualities, with no length contrast, /i, e, a, u/, are permitted (Sastry 1972). Trubetzkoy (1939/1969:235) observes that the length contrast for vowels is neutralised in final position in German, Dutch, English, Norwegian, and Swedish; in Czech it is neutralised in initial position. Also, the word-initial syllable has a special phonotactic status in many languages as a position of maximum vocalic contrast, a fact which partially parallels the Focus Vowel position in Australian. Several Altaic languages
license vowel rounding contrasts in initial syllables that are not permitted elsewhere in words: Vogul, Bashkir and Ostyak allow round vowels only in initial syllables. In Turkish, another Altaic language, both the high and low vowels contrast for rounding in initial syllables but only the high vowels do subsequently in the word. In Yokuts, an American Indian language of California, round vowels can only occur in non-initial syllables only when the vowel of the initial syllable is round. (See discussion and analysis of these facts in Steriade 1994, as well as references cited there.)

References


The function of F0-peak delay in Japanese

Yoko Hasegawa
University of California, Berkeley

Kazue Hata
Speech Technology Laboratory, Panasonic Technologies, Inc.

I. INTRODUCTION

There is a perceptual difference between male and female speech in fundamental frequency (F0) and, less significantly, in formant frequencies. Many, naturally, assume that these characteristics are physiologically determined to a great extent, and thus speakers exert little control over these characteristics in their normal utterances. F0 is certainly influenced by anatomical differences between male and female speakers. Nonetheless, observed F0 differences are much greater than those that can be attributed to anatomy: F0 is also manipulated by speakers to display their demeanor, for example in conforming to social norms.

The present study reports on one such F0 manipulation, the paralinguistic function of F0-peak delay in Japanese. We claim that F0 in Japanese is not only modified to indicate accentual distinctions but can also be used to convey femininity.

In this paper, we first provide a brief review of previous studies on male-female speech differences, and then explain how Japanese pitch accent is realized. Section II reports on our experiment to test the hypothesis that F0-peak delay and femininity are associated. Section III discusses the similarities and differences in functions of F0-peak delay between pitch-accent and intonation languages.

1.1 Comparisons of Male and Female Speech

Male speech and female speech are sufficiently different that people usually have no difficulties in identification of a speaker’s gender. Usually, female speakers have higher F0 (Peterson & Barney, 1952) and higher formant frequencies (Chiba & Kajiyama, 1941), with the F0 difference being more significant than the formant-frequency differences between male and female groups (Coleman, 1976). These acoustic differences are partly due to anatomical differences between males and females. For example, the length of adult male vocal folds (cords) is about 15 to 20 mm, whereas the length of adult female’s is 9 to 13 mm (Sundberg, 1987). Furthermore, the average vocal tract length of an adult male is 17.5 cm (Pickett, 1980), and that of an adult female is about 15% shorter (Nordström, 1977).
Although there is no dispute as to whether these anatomical differences are a cause of different F0 ranges observed between male and female speakers, the actual F0 differences cannot be explained solely in anatomical terms. F0 is also manipulated, whether consciously or unconsciously, by both gender groups (Sachs et al., 1973). Ohala (1983) contends from an ethological point of view that not only humans but also other species manipulate F0 to convey and clarify their intentions and attitudes. He claims that a confident or dominant individual utters low-pitched and often harsh sounds, while a submissive or subordinate individual utters high-pitched and tone-like sounds — he calls this phenomenon the frequency code, which evolved originally from the association between body size and emitted frequency for non-linguistic communication. This frequency code is used in linguistic communication as well. Human males frequently use low F0, which is associated not only with a large body size but also with character traits such as aggressiveness, assertiveness, self-confidence, and so forth. By contrast, high F0 suggests that the speaker has a small body, is non-threatening, submissive, subordinate, and in need of others' cooperation and good will (ibid.). In addition, when humans ask a question, a high pitch is a natural choice to signal their dependency on others. But when they make a statement, a low pitch would be more appropriate to show their confidence (ibid.).

Other researchers have postulated that F0 differences are more of a social phenomenon, reflecting different social norms laid down for men and women (Trudgill, 1974; Brend, 1975; Lakoff, 1975; Brown & Levinson, 1978; Edelsky, 1979; Jugaku, 1979; Loveday, 1981). Society assigns different social roles to men and women and expects different behavioral patterns from each group; language simply reflects this social fact (Trudgill, 1974). American females, for example, make great use of rising intonation (Brend, 1975), and, when using formal and polite register, Japanese female speakers use higher F0 than female speakers of other languages (Jugaku, 1979; Loveday, 1981). The frequent use of higher F0 and rising intonation are non-physiological aspects of female speech.

1.2 Ososagari: Delayed F0 Fall
Some Japanese speakers, especially female speakers, tend to delay the F0 peak that signals a lexical accent. This phenomenon is known as ososagari (delayed F0 fall), which is briefly explained in this section.

The Tokyo dialect of Japanese is a prototypical pitch-accent language in which accent is realized solely by a change in pitch, not by a change in loudness or duration such as found in English. Phonologically, it is widely assumed that the accented syllable in Japanese has a high tone, and the post-accent syllable a low tone; phonetically, the accentual high tone is realized by a higher F0 value on
the accented syllable than on surrounding syllables (Pierrehumbert & Beckman, 1988). This acccentual F0 peak, however, occasionally occurs on the post-accent syllable, without listeners detecting any change in accent placement. For example, in the word /námida/ ‘tears’, although the lexical accent falls on the first syllable /na/, the actual F0 peak may occur on the second syllable /mi/, as shown in Figure 1.

![Figure 1: F0 Contour of /ámí/ in the word námida ‘tears’](image)

However, the first syllable /a/ is still perceived as accented in tokens with F0-peak delay, even though the actual F0 peak does not occur on it. Listeners are not consciously aware of the delay and do not consider that the accent is shifted.

Investigating the ososagari phenomenon, Sugito (1968) discovered that the most significant acoustic correlate of the Japanese accent is a falling F0 contour of the post-accent syllable, rather than the F0-peak location itself; i.e., native speakers of Japanese perceive an accent on a syllable when it is followed by a falling F0 contour. In the /námida/ example above, if the post-accent syllable /mi/ contains an F0 fall, /na/ is perceived as accented, even when the F0 peak occurs on /mi/.

Although the frequency of ososagari varies between speakers as well as between words uttered by the same speaker, it is more commonly observed in words with an accent in certain positions or certain segmental environments. In Sugito’s data, about 36% of 3-, 4-, 5-syllable words with the accent on the initial syllable had delayed F0 peak. In Hasegawa and Hata (1988), we reported that 37% of the words beginning with /(C)VmV/ were uttered with delayed F0 peak. We also found in the same experiment that the phenomenon
occurs much more frequently in female speakers' utterances than those of male speakers (38% vs. 5% of the time). A subsequent experiment confirmed the same tendency (Hata & Hasegawa, 1992).

Based on the results of our previous experiments, we hypothesized that listeners associate F0-peak delay with a feminine speech style in Japanese and conducted a perceptual experiment to test this hypothesis. We presented synthesized sentences with and without peak delay to native speakers of Japanese, and the subjects judged which utterance in each pair sounded more female-like.

II. EXPERIMENT

2.1 Stimuli

The following sentences were prepared with a speech synthesizer. Each sentence contains a target word (shown in italics), which has the lexical accent on the first syllable.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A.</td>
<td>námida ga deru.</td>
<td>'Tears came into my eyes.'</td>
</tr>
<tr>
<td>B.</td>
<td>túmeri kore desu ne.</td>
<td>'You mean this, don't you?'</td>
</tr>
<tr>
<td>C.</td>
<td>kátta ga itai.</td>
<td>'I have sore shoulders.'</td>
</tr>
<tr>
<td>D.</td>
<td>kánari omosiroi.</td>
<td>'It's fairly interesting.'</td>
</tr>
</tbody>
</table>

The F0 contour of sentences A, C, and D was a rise-fall shape with the F0 starting at 230 Hz and ending at about 150 Hz. Sentence B contains a tag question, and thus had a slight rise at the end of the sentence. The global F0 peak always occurred within the target word.

Each sentence had two variations: in one the target word was made without an F0-peak delay (non-delayed token), and the other variation there was a delay (delayed token). Figures 2 and 3 represent these two types of F0 contour.

As shown in Figure 2, the non-delayed tokens had the peak (270 Hz) in the middle of the vowel of the lexically-accented first syllable, decreasing to 200 Hz at the onset of the third vowel. Once the fall reached 200 Hz, the F0 was sustained into the third syllable.

In the delayed tokens, as shown in Figure 3, the peak (300 Hz) occurred at 30 msec into the second vowel, followed by a 7-semitone F0 fall. As with the non-delayed tokens, once the fall reached 200 Hz, the F0 was leveled throughout the third syllable.⁴
Figure 2: F0 contour of a target word without a delayed F0 peak (non-delayed token)

Figure 3: F0 contour of a target word with a delayed F0 peak (delayed token)

Having obtained eight distinct sentence-tokens (4 sentences x 2 variations), we coupled the non-delayed token (ND token) and delayed token (D token) of each sentence in two orders: (1) the non-delayed token first and then the delayed token and (2) the delayed token first and then the non-delayed token, as shown below.

<table>
<thead>
<tr>
<th></th>
<th>A1</th>
<th>B1</th>
<th>C1</th>
<th>D1</th>
<th>A2</th>
<th>B2</th>
<th>C2</th>
<th>D2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ND-D</td>
<td>ND-D</td>
<td>ND-D</td>
<td>ND-D</td>
<td>D-ND</td>
<td>D-ND</td>
<td>D-ND</td>
<td>D-ND</td>
</tr>
<tr>
<td>námida ga deru.</td>
<td>tûmari kore desu ne.</td>
<td>kàta ga itai.</td>
<td>kânari omosiroi.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
These eight pairs were duplicated and randomized. The general instructions of the experiment were given to the subjects in written form as well as in spoken form using the synthetic voice. This precaution was taken in order to familiarize the subjects with the synthetic voice quality.

2.2 Procedure
Thirty-two subjects (19 males and 13 females, all native speakers of Japanese) participated in this experiment. They were told that one of the tokens in each pair was uttered by a male speaker and the other by a female speaker, and that the recorded voice was normalized with respect to pitch and length. The subjects then listened to the stimuli and decided which one in each pair sounded more female-like. Using answer sheets they circled 1 if they thought that the first token was likely to be uttered by a female speaker, and circled 2 if the second token sounded more female-like.

2.3 Results
Figure 4 summarizes the results. The abscissa shows how many delayed tokens each subject identified as uttered by a female speaker, and the ordinate shows how many subjects obtained the indicated score.

![Bar Chart]
Figure 4: Results of the experiment

There were 16 token-pairs, half of which had the ND-D order and the other half the D-ND order. Therefore, if there is no perceptual difference between the non-delayed tokens and delayed tokens, we expect the average of the subjects' judgments of delayed tokens as more feminine to be close to 8 (50%). The result, however, shows that the responses are skewed: the average is 9.47 (59.2%) with a standard deviation of 2.21. The difference is highly statistically significant ($t (31) = 3.7526$, two-tailed $p < 0.001$).
This result was surprising to us because all subjects stated after the experiment that they could not hear any difference between the two tokens in each pair, and that they circled numbers randomly. The statistics, however, strongly suggest that the difference was perceived, although the subjects were not consciously aware of it.

It has been reported in the literature that not all listeners utilize the same strategy in detecting the prominent syllable. For example, among four cues for English accent (F0, duration, amplitude, spectral patterns), F0 is reported to be most significant (Fry, 1958), and yet, other cues being equal, some native English listeners do not respond to F0 changes alone in determining pitch peaks (Hata & Hasegawa, 1991). Therefore, it can be useful to separate the subjects according to their strategies in F0-perception experiments (Bartels & Kingston, 1994).

As seen in Figure 4, there is a great variability in the between-speaker results in our experiment. For example, one subject associated delayed tokens with a female speaker 14 times out of 16. It is unlikely that the subject obtained this score by mere guessing. We thus investigated how consistent the subjects’ responses were. Because each order of tokens was duplicated, if subjects unconsciously perceived delayed tokens as more feminine, they would likely give the same responses for the identical pairs.

For this purpose, we counted only those responses that indicated (A) the delayed token as more feminine both times (consistent D-feminine responses) or (B) the non-delayed token as more feminine both times (consistent ND-feminine responses) for each identical pair of stimuli. We then divided (A) by the total consistent responses (A + B) and obtained the result shown in Figure 5.

![Figure 5: Consistent D-female Responses](image)

The average of consistent D-feminine responses was 67%; the mode was 80%; and five of the 32 subjects gave 100% consistent D-feminine responses. On the other hand, one subject each delivered 0, 20, and 30% consistent D-feminine responses. This result supports the hypothesis that delayed F0 peak is
a significant cue to femininity in Japanese, but not all native listeners are equally sensitive to it.

III. DISCUSSION

F0-peak delay with respect to the accented/stressed syllable has been discussed in the literature for decades to account for intonational meanings (Bolinger, 1958; O'Connell & Arnold, 1961; Ladd, 1983; Gussenhoven, 1984; Pierrehumbert & Steele, 1989; inter alia). For example, when an adult speaker says to a sobbing child ‘Would you rather have your Mommy take you to the hospital?’, the speaker may utter Mommy with a rise-fall contour (i.e. with a peak delay), rather than with an unmodified simple fall on Mommy, to indicate his/her concern for being understood by the inexperienced and possibly inattentive child (Gussenhoven, 1984). To give another example, when the speaker is certain that the person in question is a millionaire, millionaire is pronounced with a peak on the lexically stressed first syllable, but when the speaker is in doubt, the F0 peak shifts to the second syllable (Pierrehumbert & Steele 1989).

These previous works have demonstrated that delay in the alignment between the F0 peaks and accented syllables can have a significant function in human languages. However, what this significance actually designates may differ according to language types. In intonation languages, e.g. English, it is possible for the F0-peak delay to stretch over many syllables, and the listener can easily hear the delay and interpret it to be conveying some intonational meaning, e.g. the speaker’s attentiveness, incredulity or uncertainty.

In pitch-accent languages, on the other hand, if the F0 peak is detected on a post-accent syllable, the perceived accent will inevitably shift to that syllable, resulting in an anomalous pronunciation. Therefore, the magnitude of delay in pitch-accent languages is much more limited than that in intonation languages (cf. Hata & Hasegawa, 1988). As a consequence, the conveyed meaning in pitch-accent languages is likely to be subtler than the meanings found in intonation languages. In the case of Japanese, as the present experiment has demonstrated, the paralinguistic function of the delay is to convey femininity.

The meanings of F0-peak delay in both types of languages are not arbitrarily chosen, however. It appears that what is underlying those meanings is indecisiveness (or hesitancy), which normally invites the participation by the hearer in the conversation. Female speakers are reported to prefer more collaborative conversations than male speakers do (Lakoff, 1990; Tannen, 1990). We conjecture that peak delay signals a degree of indecisiveness, which is frequently associated with females, and in turn, is further conventionalized to convey a sense of femininity in Japanese.
NOTES

* We would like to thank the following individuals for their comments and criticism: Mary Beckman, Gregory De Haan, Carlos Gussenhoven, Larry Hyman, Anita Liang, John Ohala, Manjari Ohala, Natasha Warner, and Raymond Weitzman.. This project was supported in part by a grant from the Center for Japanese Studies at the University of California, Berkeley.

1 Mattingly (1966) reports that male-female differences in formant frequencies are also chiefly due to social conventions, not to physiological differences.

2 The phenomenon of ososagari was first reported by Neustupný (1966).

3 We suspect that nasals have some influence on the F0-peak delay.

4 We have chosen two different F0-peak frequencies for the non-delayed and delayed tokens in order to balance the perceived overall pitch ranges. Because the non-delayed tokens have more gradual F0 change, if the peak frequency were 300 Hz, as is the case for the delayed tokens, each non-delayed token as a whole would sound much higher than the corresponding delayed token. Comparing several non-delayed tokens with varying F0 peak frequencies, we have determined that a 270-Hz peak renders a voice range most compatible with the delayed tokens. Other characteristics (e.g. formant frequencies, amplitude, speech rate) are identical in both types of tokens.

REFERENCES


A non-derivation approach to Winnebago stress

Andrea Heiberg
University of Arizona

1. Introduction

The interaction of stress and epenthesis in Winnebago has long posed a challenge to metrical theory. From a derivational perspective, it seems that Winnebago epenthetic vowels (called 'Dorsey's Law vowels') in certain positions are inserted before stress rules apply, but in other positions, they are inserted after stress rules apply.

Derivational theories have proposed extra processes and conditions to handle the Winnebago paradox, including rightward accent shift (Miner 1979, 1981, 1990, Hayes 1995), repair of ill-formed metrical structure (Hale and White Eagle 1980, Hale 1985), the Domino Condition (Halle and Vergnaud 1987), insertion of foot boundaries for epenthetic vowels (Idsardi 1992), and relativized extrametricality (Hammond 1995).

The Optimality Theory (McCarthy and Prince 1993a,b, Prince and Smolensky 1993) account offered here avoids the derivational paradox by considering only surface generalizations about stress and Dorsey's Law vowels. The odd behavior of Dorsey's Law vowels is explained through a violable constraint which prohibits a phonetically prominent Dorsey's Law vowel from occupying the weak position of a foot. Other researchers have simply noted the prominence of Dorsey's Law vowels but have not incorporated it into their accounts (Susman 1943, Miner 1979), or have overlooked it entirely (Halle and Vergnaud 1987, Idsardi 1992, Hammond 1995). Hence, this study makes both theoretical and empirical contributions to an emerging non-derivation theory of language.

2. General stress pattern

A stressed syllable in Winnebago has higher tone and greater loudness than surrounding syllables (Susman 1943, Miner 1979, White Eagle 1988). There are no monomoraic words (Susman 1943); bimoraic monosyllabic words are realized with a falling tone contour (Susman 1943, Hale and White Eagle 1980 (henceforth HWE)). Stress falls on the second mora of other bimoraic words.

| (1) | číí  | 'house' HWE, Susman 1943 |
|     | wáak | 'man' HWE, Miner 1979, Susman 1943 |

| (2) | (μ, ū) | hapé | 'he waited for him' Susman 1943 |
|     |       | híník | 'son' Susman 1943 |
|     |       | rugás | 'to tear' Miner 1990 |
|     |       | wažók | 'mash (potatoes)' HWE |
In trimoraic and longer words, primary stress falls on the third mora and secondary stresses fall on every other mora thereafter. The following assumptions are made about the general stress pattern: (i) Right-headed binary feet are built on moras; (ii) The leftmost mora is not footed in words of three or more moras; (iii) A foot may 'split' a syllable (i.e. the two moras of a long syllable may belong to different feet, or in the case of an initial long syllable, the leftmost mora may be unfooted while the rightmost mora belongs to a foot).

(3) \( \mu (\mu \dot{\mu}) \) harapé 'you waited for him' Susman 1943
    hinigrá 'the son' Susman 1943
    hokit'é 'he speaks to' HWE
    booshíp 'he shot it down' Susman 1943
    číirá 'the house' Susman 1943
    mañí 'to walk' Susman 1943

(4) \( \mu (\mu \dot{\mu}) \mu \) harapége 'because you waited for him' Susman 1943
    horakít'é 'you speak to' HWE
    booshípke 'he shot it down often' Susman 1943
    mañíke 'he often walks' Susman 1943

(5) \( \mu (\mu \dot{\mu}) (\mu \dot{\mu}) \) hirawáhazrā 'the license' Miner 1979
    waborášípke 'you shot them down often' Susman 1943
    haakitujík 'I pull it taut' HWE

(6) \( \mu (\mu \dot{\mu}) (\mu \dot{\mu}) \mu \) hakirújíkgàjá 'after he pulls taut' Miner 1990

(7) \( \mu (\mu \dot{\mu}) (\mu \dot{\mu}) (\mu \dot{\mu}) \) haakitujíkgajá 'after I pull taut' Miner 1990

The foot binarity requirement is formally captured with the constraint \( FTBIN \) (Prince and Smolensky 1993). \( PARSE \) (Prince and Smolensky 1993) insures that as many moras as possible are footed.

(8) \( FTBIN: \) Feet must be binary under syllabic or moraic analysis.

(9) \( PARSE: \) All moras must be parsed by feet.

\( FTBIN \) is ranked over \( PARSE \), disallowing degenerate feet. In tableau (14) below, compare the optimal candidate (14d), which has no degenerate feet, to the non-optimal candidates (14b) and (14c), each of which has a degenerate foot.

\( NONINITIALITY \) (\( NONI \)), the logical counterpart of \( NONFINALITY \) (Prince and Smolensky 1993), prevents a word-initial mora from being footed. A member of the \( ALIGN \) family of constraints (McCarthy and Prince 1993a) requires that all feet be left-aligned to the prosodic word as much as possible (compare optimal (14d) to (14e)).
(10) NONFINALITY: No head of PrWd is final in PrWd.

(11) NONI (NONINITIALITY): No head of PrWd is initial in PrWd.

(12) ALIGN(Ft, L, PrWd, , L): The left edge of every foot is aligned with the left edge of a prosodic word.

NONI is ranked above PARSE: candidate (14d), since it does not violate NONI, is chosen over the more completely footed (14a). PARSE is ranked above ALIGN: even though candidate (14f) has fewer ALIGN violations than (14d), the optimal (14d) has fewer PARSE violations. The ranking of NONI and FtBIN is indeterminate. If NONI is ranked above FtBIN, then bimoraic forms have an unfooted initial mora followed by a single degenerate foot (see (15)). If FtBIN is ranked above NONI, then bimoraic forms have a single binary foot (see (16)). The total ranking of the four constraints is given in (13).

(13) Total ranking: \{ FtBIN \NONI \} » PARSE » ALIGN

(14) hakirújíkágàjà 'after he pulls taut'

<table>
<thead>
<tr>
<th>Input: ( \mu \mu \mu \mu \mu \mu )</th>
<th>FtBIN</th>
<th>NONI</th>
<th>PARSE</th>
<th>ALIGN</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ( (\mu \dot{\mu}) (\mu \dot{\mu}) (\mu \dot{\mu}) )</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. ( \mu (\mu \dot{\mu}) (\mu \dot{\mu}) (\dot{\mu}) )</td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. ( \mu (\dot{\mu}) (\mu \dot{\mu}) (\mu \dot{\mu}) )</td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. ( \mu (\mu \dot{\mu}) (\mu \dot{\mu}) \mu )</td>
<td>**</td>
<td></td>
<td>****</td>
<td></td>
</tr>
<tr>
<td>e. ( \mu (\mu \dot{\mu}) \mu (\mu \dot{\mu}) )</td>
<td>**</td>
<td></td>
<td></td>
<td>****</td>
</tr>
<tr>
<td>f. ( \mu (\mu \dot{\mu}) \mu \mu \mu )</td>
<td>***</td>
<td></td>
<td></td>
<td>**</td>
</tr>
</tbody>
</table>

(15) NONI » FtBIN for hapé 'he waited for him'

<table>
<thead>
<tr>
<th>Input: ( \mu \mu )</th>
<th>NONI</th>
<th>FtBIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ( (\mu \dot{\mu}) )</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>b. ( \mu (\dot{\mu}) )</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

(16) FtBIN » NONI for hapé 'he waited for him'

<table>
<thead>
<tr>
<th>Input: ( \mu \mu )</th>
<th>FtBIN</th>
<th>NONI</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ( (\mu \dot{\mu}) )</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. ( \mu (\dot{\mu}) )</td>
<td></td>
<td>*!</td>
</tr>
</tbody>
</table>
3. **Stress patterns of words with Dorsey’s Law vowels**

The distribution of epenthetic vowels in Winnebago is governed by what has been called Dorsey’s Law. A tautomorphemic cluster of a voiceless obstruent and a sonorant is split by an epenthetic vowel that is identical to the following vowel (Miner 1990). Dorsey’s Law vowels (henceforth 'DL vowels') are typographically marked here by underlining. A DL vowel is always high-toned (Susman 1943, Miner 1979, 1981):

\[
\begin{align*}
(17) & \quad \text{H} \\
& \quad \text{kere} = \text{kere}
\end{align*}
\]

If a DL vowel falls on what would be the weak position of a foot under the normal footing pattern and there are at least two moras following the DL vowel, the footing pattern is 'shifted' one mora to the right (as in (20c), (21a,b), (22a,d), (23a), and (24)). If a DL vowel falls on what would be the weak position of a foot but there are less than two moras following the DL vowel, the normal footing pattern obtains (see (18), (19b), (21d), (22a), and (23c)). If a DL vowel falls on what would not be the weak position of a foot under the normal footing pattern (i.e. the strong position of a foot or the initial mora), the normal footing pattern obtains (see (19a), (20a,b,d), (21c), (22b,c,e), and (23b)).

\[
(18) \quad (\mu \ ˘) \quad \text{pgrás} \quad \text{‘flat’ Susman 1943, Miner 1990} \\
& \quad \text{pgró} \quad \text{‘spherical’ Susman 1943}
\]

\[
(19) \quad \text{a. } (\mu \ ˘) \quad \text{šgawažók} \quad \text{‘you mash (potatoes)’ HWE, Miner 1990} \\
& \quad \text{šýrugás} \quad \text{‘you tear’ Miner 1990}
\]

\[
\text{b. } (\mu \ ˘) \quad \text{ruxýrůk} \quad \text{‘to earn’ Susman 1943} \\
& \quad \text{wakírí} \quad \text{‘small animals, insects’ Susman 1943}
\]

\[
(20) \quad \text{a. } (\mu \ ˘) \quad \text{šgawažókji} \quad \text{‘you mash hard’ Miner 1990} \\
& \quad \text{kgrejúsep} \quad \text{‘Black Hawk’ Miner 1990}
\]

\[
\text{b. } (\mu \ ˘) \quad \text{pgrópóro} \quad \text{‘spherical’ Miner 1990} \\
& \quad \text{šýrugúrůk} \quad \text{‘you earn’ Miner 1990}
\]

\[
\text{c. } (\mu \ ˘) \quad \text{hikoroñó} \quad \text{‘he prepares’ HWE, Miner 1990, Susman 1943} \\
& \quad \text{ruxýrúké} \quad \text{‘he often earns it’ Susman 1943}
\]

\[
\text{d. } (\mu \ ˘) \quad \text{hirakára} \quad \text{‘he took care of it’ Susman 1943} \\
& \quad \text{rookéwe} \quad \text{‘he dressed him’ Susman 1943}
\]
(21) a. \( \mu \mu \mu \mu \) hakewéhášge 'six times perhaps' Susman 1943
    hok'awanege 'so that he could come in' Lipkind 1945
b. \( \mu \mu \mu \mu \) gikanákánap 'shiny' Miner 1990
    wakiripáras 'flat bug' HWE, Miner 1990
c. \( \mu \mu \mu \mu \) harakéwegi 'if you enter' Susman 1943
    hirakírhohó 'you prepare' HWE
d. \( \mu \mu \mu \mu \) roorákewé 'you dressed him' Susman 1943
    wiirákará 'he took care of them' Susman 1943

(22) a. \( \mu \mu \mu \mu \) hákevéákgášcí 'he is entering (moving)' Miner 1979
b. \( \mu \mu \mu \mu \) hirakírhohóní 'you don’t dress' Halle and Vergnaud 1987
c. \( \mu \mu \mu \mu \) waapóropóró 'snowball' Miner 1979
d. \( \mu \mu \mu \mu \) boonjárášip 'I shot you (my own) down' Susman 1943
e. \( \mu \mu \mu \mu \) hákiríjíkíšá 'he pulls taut' HWE, Miner 1990

(23) a. \( \mu \mu \mu \mu \) gikanákánapášá 'it is shiny' Miner 1990
    wakiripáropóró 'spherical bug' HWE, Miner 1990
b. \( \mu \mu \mu \mu \mu \) hirakírhöhónírát 'the fact that you do not dress' Halle 1990
c. \( \mu \mu \mu \mu \mu \) haakítújíkíšá 'I pull it taut' HWE, Miner 1990

(24) \( \mu \mu \mu \mu \mu \) harakíšurújíkíšá 'you pull it taut' HWE

I propose that the general strategy for footing DL vowels is to avoid putting them in the weak position of a foot. Recall that high tone is a primary correlate of stress in Winnebago: the strong position of a foot is realized as high-toned. Since a DL vowel is always high-toned, it makes a poor weak member of a foot.

Winnebago, then, is a prominence system. A syllable with a DL vowel is the analog of a heavy syllable in a stress system that respects syllable weight. However, it is not length that makes DL vowels phonetically prominent; rather, it is their high tone.

The Weight-to-Stress Principle (WSP) (Hayes 1980, Prince 1990, Prince and Smolensky 1993) has been used to account for the behavior of heavy syllables in prominence systems:

(25) WSP (Weight-to-Stress Principle): Heavy syllables are prominent in foot structure and on the grid.

WSP may be restated as, 'If \( x \) is phonetically prominent, then \( x \) falls on the strong position of a foot. (If \( x \) does not fall on the strong position of a foot, then \( x \) is not phonetically prominent)' (see Prince and Smolensky 1993). The
problem for Winnebago is that WSP assigns the same penalty to a phonetically prominent light syllable that is not footed (as in the attested but non-optimal (26c)) and a phonetically prominent light syllable that falls on the weak position of a foot (as in the unattested but optimal (26b)). Note that there is no possible ranking of WSP within the established hierarchy \{FtBIN, NONI\} » PARSE » ALIGN that would select (26c) over the other candidates.

(26) Marks assigned by WSP to hikoroḥō 'he prepares'

<table>
<thead>
<tr>
<th>Input: μ μ μ μ</th>
<th>FtBIN</th>
<th>NONI</th>
<th>WSP</th>
<th>PARSE</th>
<th>ALIGN</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. (μ ū) (μ ū)</td>
<td></td>
<td>*!</td>
<td></td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>! b. μ (μ ū) μ</td>
<td></td>
<td></td>
<td>**</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>(脩)c. μ μ (μ ū)</td>
<td></td>
<td></td>
<td>**</td>
<td></td>
<td>**!</td>
</tr>
<tr>
<td>d. μ (ū) (μ ū)</td>
<td>*!</td>
<td></td>
<td>*</td>
<td></td>
<td>***</td>
</tr>
<tr>
<td>e. μ (μ ū) (ū)</td>
<td>*!</td>
<td></td>
<td>*</td>
<td>*</td>
<td>****</td>
</tr>
</tbody>
</table>

The proposed solution to this problem is to restate WSP so that it does not penalize an unfooted phonetically prominent syllable. Rather, it penalizes only those phonetically prominent syllables which fall on the weak position of a foot. This restatement of WSP is referred to as the Prominence-to-Headship Principle (PHP):

(27) PHP (Prominence-to-Headship Principle)

If \( x \) is phonetically prominent, then \( x \) does not fall on the weak position of a foot. (If \( x \) falls on the weak position of a foot, then \( x \) is not phonetically prominent.)

PHP crucially does not assign a mark to the attested candidate (28c):

(28) Marks assigned by PHP to hikoroḥō 'he prepares'

<table>
<thead>
<tr>
<th>Input: μ μ μ μ</th>
<th>FtBIN</th>
<th>NONI</th>
<th>PHP</th>
<th>PARSE</th>
<th>ALIGN</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. (μ ū) (μ ū)</td>
<td></td>
<td>*!</td>
<td></td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>b. μ (μ ū) μ</td>
<td></td>
<td></td>
<td>*!</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>(脩)c. μ μ (μ ū)</td>
<td></td>
<td></td>
<td>**</td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>d. μ (ū) (μ ū)</td>
<td>*!</td>
<td></td>
<td>*</td>
<td></td>
<td>***</td>
</tr>
<tr>
<td>e. μ (μ ū) (ū)</td>
<td>*!</td>
<td></td>
<td>*</td>
<td>*</td>
<td>****</td>
</tr>
</tbody>
</table>

NONI is ranked above PHP in Winnebago, since it is not the case that an initial mora may be footed in order to satisfy PHP (compare optimal (30b) to (30c)). FtBIN is also ranked above PHP: a DL vowel may not form a
degenerate foot in order to satisfy PHP (compare (30b) to (30d,e)). Finally, PHP is ranked above PARSE, since moras may go unfooted in order to satisfy PHP (compare (30b) to (30a)).

(29) Total ranking: \{ \text{FTBIn} \} \gg \text{PHP} \gg \text{PARSE} \gg \text{ALIGN}

(30) hakewehäsge 'six times perhaps'

<table>
<thead>
<tr>
<th>Input: μ μ μ μ μ</th>
<th>FTBIn</th>
<th>NONI</th>
<th>PHP</th>
<th>PARSE</th>
<th>ALIGN</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. μ (μ *μ) (μ *μ)</td>
<td></td>
<td></td>
<td>*!</td>
<td>*</td>
<td>****</td>
</tr>
<tr>
<td>b. μ μ (μ *μ) μ</td>
<td></td>
<td></td>
<td></td>
<td>***</td>
<td>**</td>
</tr>
<tr>
<td>c. (μ *μ) (μ *μ) μ</td>
<td></td>
<td></td>
<td>*!</td>
<td>*</td>
<td>**</td>
</tr>
<tr>
<td>d. μ (μ *μ) (μ *μ) (μ *μ)</td>
<td></td>
<td></td>
<td>*!</td>
<td>*</td>
<td>******</td>
</tr>
<tr>
<td>e. μ (μ *μ) (μ *μ) μ</td>
<td></td>
<td></td>
<td>*!</td>
<td>*</td>
<td>**</td>
</tr>
</tbody>
</table>

How does PHP account for systems that treat heavy syllables as prominent? It seems that PHP incorrectly allows heavy syllables to go unfooted. However, a heavy syllable can be footed by itself without violating FTBIn (unlike a syllable with a DL vowel, which violates FTBIn when footed by itself). Imagine a system identical to Winnebago except that in place of DL vowels there are heavy syllables. The (d) candidate in tableaux (31) and (32) is optimal regardless of whether PHP or WSP is the constraint in question, since the closest competitor of (d), candidate (c), incurs more PARSE violations than (d).

(31) Marks assigned by PHP to a system with prominent heavy syllables

<table>
<thead>
<tr>
<th>Input: μ μ μ μ μ</th>
<th>FTBIn</th>
<th>NONI</th>
<th>PHP</th>
<th>PARSE</th>
<th>ALIGN</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. (μ μ*μ) (μ *μ)</td>
<td></td>
<td></td>
<td>*!</td>
<td>***</td>
<td></td>
</tr>
<tr>
<td>b. μ (μμ *μ) μ</td>
<td></td>
<td></td>
<td></td>
<td>**</td>
<td>*</td>
</tr>
<tr>
<td>c. μ μμ (μ *μ)</td>
<td></td>
<td></td>
<td></td>
<td>*<em>!</em></td>
<td>**</td>
</tr>
<tr>
<td>d. μ (μ*μ) (μ *μ)</td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
<td>*****</td>
</tr>
<tr>
<td>e. μ (μμ *μ) (μ *μ)</td>
<td></td>
<td></td>
<td>*!</td>
<td>*</td>
<td>*****</td>
</tr>
</tbody>
</table>
(32) Marks assigned by WSP to a system with prominent heavy syllables

<table>
<thead>
<tr>
<th>Input: μ μμ μ μ</th>
<th>FtBIN</th>
<th>NONI</th>
<th>WSP</th>
<th>PARSE</th>
<th>ALIGN</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. (μ μμ) (μ μ)</td>
<td>*!</td>
<td></td>
<td></td>
<td>***</td>
<td></td>
</tr>
<tr>
<td>b. μ (μμ μ) μ</td>
<td></td>
<td>***</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. μ μμ (μ μ)</td>
<td></td>
<td>***</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. μ (μμ) (μ μ)</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. μ (μμ) (μ μ)</td>
<td>*!</td>
<td>*</td>
<td>*</td>
<td></td>
<td>***</td>
</tr>
</tbody>
</table>

The only situation in which failing to foot a phonetically prominent syllable is optimal is when the syllable is light and would therefore violate FtBIN if footed alone; this is the case for Winnebago.

3.1 Violations of PHP

PHP is violable in Winnebago under certain circumstances: (i) The head foot of the prosodic word may violate PHP in order to avoid violating NONI or FtBIN (e.g. (poró) 'spherical', ru(xyrük) 'to earn'); (ii) A non-head foot may violate PHP when a DL vowel is followed by a single mora (e.g. ro(orá)(kegē) 'you dressed him'). PHP is split into two constraints, PHP(Head-of-PrWd) and PHP(Foot). PHP(Head-of-PrWd) is ranked above PARSE, as argued above. PHP(Foot) is ranked below PARSE, since moras may not go unparsed in order to satisfy PHP for non-head feet (compare optimal (34a) to (34d,e)). PHP(Foot) is ranked above ALIGN, since a word-medial prominent syllable may be 'skipped' in order to satisfy PHP(Foot) (compare (35b) to (35c)). These subrankings give the final total ranking in (33).

(33) Final total ranking:

\[
\{ \text{FtBIN} \} \quad \text{» PHP(Head) » PARSE » PHP(Foot) » ALIGN}
\]

(34) roorákevè 'you dressed him'

<table>
<thead>
<tr>
<th>Input: μ μμ μ μ</th>
<th>FtBIN</th>
<th>NONI</th>
<th>PHP (Head)</th>
<th>PARSE</th>
<th>PHP (Foot)</th>
<th>ALIGN</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. (μ μμ) (μ μ)</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>b. μ (μμ μ) μ</td>
<td>*!</td>
<td></td>
<td></td>
<td>**</td>
<td></td>
<td>***</td>
</tr>
<tr>
<td>c. μ (μμ) (μ μ)</td>
<td></td>
<td></td>
<td></td>
<td>**</td>
<td></td>
<td>*****</td>
</tr>
<tr>
<td>d. μ (μμ) μ μ</td>
<td></td>
<td>***!</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. μ μ (μ μ) μ</td>
<td></td>
<td>***!</td>
<td></td>
<td>**</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
(35) \textit{harakíšurujíkšanq} 'you pull it taut'

<table>
<thead>
<tr>
<th>Input: (\mu \mu \mu \mu \mu \mu \mu \mu)</th>
<th>PHP (Head)</th>
<th>PARSE</th>
<th>PHP (Foot)</th>
<th>ALIGN</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. (\mu (\mu \mu) (\mu \mu) (\mu \mu) \mu)</td>
<td>**</td>
<td>**</td>
<td>*!</td>
<td>***********</td>
</tr>
<tr>
<td>b. (\mu (\mu \mu) (\mu \mu) (\mu \mu))</td>
<td>**</td>
<td>**</td>
<td></td>
<td>***********</td>
</tr>
<tr>
<td>c. (\mu \mu (\mu \mu) (\mu \mu) (\mu \mu))</td>
<td>**</td>
<td>**</td>
<td></td>
<td>***********</td>
</tr>
</tbody>
</table>

4. Conclusion

The Optimality Theoretic analysis presented here offers empirical and theoretical advantages to previous analyses of the patterns of Winnebago stress. The often-overlooked phonetic prominence of DL vowels is crucial to the analysis: characterizing DL vowels as prominent provides an explanation for their seemingly odd interaction with stress. The 'derivational paradox' presented by Winnebago disappears when only surface patterns about stress are considered; Optimality Theory, with its violable constraints, provides the venue for analyzing complex interactions of surface generalizations.

REFERENCES


Many thanks to Mike Hammond, Colleen Fitzgerald, Amy Fountain, Diane Ohala, Chang-Kook Suh, and Keiichiro Suzuki for their comments on previous versions of this analysis.

Department of Linguistics
Douglass 200 East
University of Arizona
Tucson, Arizona 85721
heiberg@ccit.arizona.edu
1. On the Diversity of Polarity Sensitivities

It is well-known that negative polarity items (NPIs) are sensitive to a variety of licensors besides negation, including questions, conditionals, comparatives, certain quantifiers and various adverbs like rarely and hardly. Years of effort and a voluminous literature have not reduced the controversy as to what makes these all a natural class. But in addition to the question of what in general is responsible for licensing NPIs, there is also the problem that not all NPIs share precisely the same sensitivities. NPIs like any and ever are notoriously liberal as to what they will accept as a licensor; but other NPIs can be rather more particular.

Traditionally, the response to this problem has been to posit a hierarchy of ‘negative strength’ such that stronger NPIs will only be licensed by stronger licensors (Edmondson 1981; van der Wouden 1994). But such an approach, if it is even feasible, may have little to say about why different NPIs have the peculiar sensitivities they do.¹

Given the diversity of NPI sensitivities, the goals of this paper are quite modest. I will focus on two classes of NPIs, minimizers and indefinites, arguing that distributional differences between them directly reflect their distinct constructional semantics. Minimizers are a class of moderate NPIs including forms like so much as, the least bit, a hope in hell, and sleep a wink; the indefinite NPIs, including any (anyone, anything etc.) and ever, are slightly more liberal. I will argue that both classes are subject to a general licensing condition, but that they differ crucially in the ways they can meet this condition.

I draw my data from three sorts of environments in which these classes show distinct patterns of sensitivity. The examples in 1 show NPIs in negated because clauses. Here the indefinite any is licensed, but the minimizer drink a drop sounds odd. Note that on the intended reading these sentences presuppose that Zelda did fall asleep and deny only that it was vodka that caused her to do so.

1. a. Zelda didn’t fall asleep because she drank any vodka.
   She was just very tired.

   b. *Zelda didn’t fall asleep because she drank a drop of vodka.
   She was just very tired.

The examples in 2, from Heim (1984), show that while both every and the minimizer so much as can be licensed in the restriction of a universal quantifier, so much as is more particular about which sentences of this sort it will accept.

2. a. Every restaurant that I’ve ever been to happens to have four stars in the handbook.

   b. Every restaurant that charges so much as a dime for iceberg lettuce should be shut down.

   c. *?Every restaurant that charges so much as a dime for iceberg lettuce happens to have four stars in the handbook. (Heim 1984, exs 36-38)
As Heim notes, the difference between 2b and 2c seems to be that while 2b implies that the restaurants should be shut down precisely because of what they charge for iceberg lettuce, in 2c there can be only a fortuitous correlation between what a restaurant charges for lettuce and its rating in the handbook. While this poses a problem for so much as, the more liberal ever sounds fine in 2a despite the apparently fortuitous correlation between the habits of the speaker and the ratings of the restaurants.

Finally, as noted by Lee and Horn (1994), minimizers and indefinites are not equally acceptable in the focus constituent of only. In 3a the indefinite ever is fine; in 3b the minimizer the least bit is atrocious.

3. a. Only people who’ve ever read Heidegger will understand this movie.
   b. *Only people who’ve read the least bit of Heidegger will understand this movie.

Having thus established that minimizers and indefinites do indeed behave differently, we can now turn to the question of why this should be.

2. Representing the Constructions

Both the minimizers and the indefinite NPIs, I claim, are scalar operators and, as such, depend on the availability of a scalar model in order to be happily interpreted (cf. Fauconnier 1975, 1979; Israel 1994). Following Kay (1990), a scalar model is understood as a structured set of propositions ordered so as to support inferences between them. The basic difference between the minimizers and the indefinites resides in the precise structures they require of a scalar model.

Figure 1 gives an example of a scalar model. The model consists of a propositional schema, for example the schema P, ‘Norm can solve y,’ paired with a set of values for y ordered (at least partially) along a dimension of difficulty.

\[
\begin{align*}
y_5 & \quad \text{most difficult} \\
y_4 & \quad \text{P: 'Norm can solve y.'} \\
y_3 & \quad \text{R: 'Norm cannot solve y.'} \\
y_2 & \quad \text{least difficult} \\
y_1 & \quad \text{least difficult}
\end{align*}
\]

**Figure 1: A Scalar Model of Puzzles**

In general, ability to solve a puzzle ranked high in the model will entail the ability to solve any lower ranked puzzle; conversely, inability to solve puzzles low in the ranking will entail an inability to solve any puzzles ranked higher.²

Two properties can be distinguished which characterize any proposition in the scale: (quantitative) q-value refers to a proposition’s position, high or low, within the scalar ordering; (informatively) i-value distinguishes emphatic (high i-value) propositions entailing other propositions within the model from weak (low i-value) propositions which are themselves entailed.³

Given this, minimizers can be defined as inherently emphatic forms which encode a low q-value. Normally a scalar model supports inferences only from
high q-values to low q-values, but minimizers will be restricted to just those contexts in which the inferences are reversed, for only in such contexts will a minimizer be appropriately emphatic. As Fauconnier (1975, 1979) has shown, the class of scale reversing contexts, including negation, conditionals, comparatives and much more, corresponds quite neatly to the class of NPI licensors.

It is important to note that a minimizer’s q-value does not define an actual individual but simply marks a phantom minimal element on the scale. I call the element a ‘phantom’ because, lacking any inherent referential force of its own, it refers only by virtue of the entailments it licenses. The phantom is never experienced directly — it cannot, for example, establish discourse referents — but it makes itself known by virtue of its strategic position within the scale.

Now, indefinite NPIs are just like minimizers: they too encode a phantom instance which, in order to refer, must entail other instances within a set. But whereas a minimizer makes crucial reference to some quantitative ordering within which it encodes a minimal q-value, indefinites are pure phantoms encoding only an arbitrary instance randomly selected from an array of possible instances (cf. Langacker 1991: 138). As such, indefinites require only a minimal scalar model with just one element, the phantom, ordered with respect to all other possible values. Thus although indefinite NPIs do trigger entailments over a full set of alternative values, unlike the minimizers, they are not inherently emphatic. 4

The difference between the two constructions is illustrated in figure 2.

![Minimizer Construction](image)

**Minimizer Construction**

![Indefinite Construction](image)

**Indefinite Construction**

Both NPI types yield a proposition which must be construed with respect to a set of entailed, alternate propositions. Minimizers further require that the set of alternate propositions be itself partially ordered; the indefinites are compatible with such an ordering, but do not require it. They are fine so long as the phantom element counts as arbitrarily standing in for and entailing all possible alternatives.

The proposed representations capture the intuitions of Schmerling (1971) and Heim (1984), who suggest that minimizing NPIs, but not the indefinites any and ever, “are semantically equivalent to expressions containing the word even” (Heim, p. 105). Just as the focus constituent of even must represent the least likely, and hence most informative value that would satisfy its sentence’s expressed proposition (cf. Horn 1969; Kay 1990), so too do minimizers encode the scalar element which is least likely to satisfy a given propositional schema and whose predication is thus maximally informative.

3. Phantom Reference and the Implication Constraint
Since both minimizers and indefinites encode phantoms, both will be subject to a general constraint limiting them to environments in which they can achieve what we might call 'phantom reference.' In essence, a phantom is a maximally schematic indefinite which is linked to an array (i.e. a partially ordered set) of elements and which can refer only by triggering entailments over that array.

Normally, predication of any particular instance, P(x), will entail the predication of an indefinite element: if we know that Zelda drank vodka, we know that Zelda drank a (some) thing. For a phantom to be licensed this implication must be reversed. In that case, when entailments run from indefinite values to particular instances, a phantom will be able to refer indirectly by virtue of the inferences it licenses. The requirement is captured by the implication constraint (IC) given in 4.

4. IC: Given a partially ordered set A with elements \{a, b, c \ldots \} and phantom \( \alpha \) linked to A, \( \alpha \) is licensed in a proposition P iff for every element x of A, 

\[ P[\alpha] \rightarrow P[x]. \]

Not surprisingly, negation satisfies the IC: if Zelda did not drink a thing, then clearly Zelda did not drink vodka.

We may note that the IC is rather similar in spirit to Ladusaw's (1979) proposal that NPIs are licensed only in the scope of a downward entailing (DE) operator; however, as will be shown, the IC represents a distinct, and rather less stringent constraint on NPI licensing. Moreover, while I take it for granted that all scale reversing contexts satisfy the IC and thus allow for phantom reference, it turns out that not all environments which satisfy the IC can also reverse entailments on a scale.

4. Differing Entailments: Negated because Clauses

Given the differences between minimizers and indefinites there are at least two sorts of situations in which an indefinite NPI might be licensed where a minimizer is blocked. On the one hand, there may be contexts which satisfy the IC for pure phantoms, but which fail to reverse implications on a scale. On the other hand, we may find contexts which do satisfy the IC both for pure phantoms and for the scalar minimizers, but which nonetheless do not provide a coherent scalar model in which to interpret a minimizer. In the next two sections I will present examples illustrating the latter situation. In this section I argue that negated because clauses provide an example of the former case.

Negated because clauses have played a rather prominent role in discussions of polarity ever since Linebarger (1980). Recently, Kadmon and Landman (1993) have pointed out that this context is not in fact downward entailing. The acceptability of certain NPIs in this context might thus appear problematic; however, while I accept the conclusion that negated because clauses are not DE, I will argue that they do nonetheless satisfy the IC for pure phantoms.

Without negation, entailments in a because clause run from particular individuals to indefinite phantoms. Thus, for any two propositions, P and Q, along with the individual x and the phantom \( \alpha \), the implication in 5 will be valid.

5. \( Q \) because \( P[x] \rightarrow Q \) because \( P[\alpha] \)
Common sense verifies this. If, for example, Elly was punished because she punched Lyle, it follows that she was punished because she punched someone. Similarly, if I know that Zelda fell asleep because she drank vodka, I know that she fell asleep because she drank something.

Not surprisingly, under negation these entailments are reversed and the implications run from phantoms to particular individuals, as in 6.

6. Q not because P[α] —> Q not because P[x]

Thus, if it’s not because she punched someone that Elly was punished, then it cannot be because she punched Lyle. And if Zelda didn’t fall asleep because she drank something, then she didn’t fall asleep because she drank vodka.

Negated because clauses thus satisfy the IC, and as the examples in 7-8 illustrate, the indefinite NPIs any and ever are both licensed in this context.

7. a. Zelda didn’t fall asleep because she drank anything. She was just tired.
   b. Jen isn’t happy because she has any money. She’s just in a good mood.

8. a. Elmo doesn’t get good grades because he ever studies. He’s just smart.
   b. Sally doesn’t like him because he ever buys her flowers. She just suffers from the delusion that he’s a nice guy.

But negated because clauses do not license minimizers, for the simple reason that because clauses in general do not license entailments along a quantitative scale. If, for example, Zelda fell asleep because she drank a pint of vodka, it in no way follows that she fell asleep because she drank a shot of vodka: a pint may suffice where a single shot would not. And in general, whenever some quantity, n, is a sufficient cause for some effect, one cannot assume that any lesser quantity, n – y, will also be a sufficient cause. The implication in 9 is thus not valid.

9. Q because P[n] —/—> Q because P[n-y]

Since there is no implication to be reversed, negation can have no effect and so the implication in 10 fails to hold as well.

10. Q not because P[n-y] —/—> Q not because P[n]

If drinking a shot of vodka was not a sufficient cause for Zelda to fall asleep, it’s still possible that drinking a pint was enough to make her nod off.

The failure of negated because clauses to reverse implications on a scale leads to the prediction that minimizing NPIs will not be licensed in this context. As the examples in 11 show, the prediction is borne out.

11. a. *Zelda didn’t fall asleep because she drank a drop of vodka.
    b. *Jen isn’t happy because she has a red cent.
    c. *Elmo didn’t get an A because he so much as cracked a book.

I conclude that there is good reason to believe that minimizers and indefinites do in fact differ in the ways proposed above, and that the distributions of both depend on satisfying the IC. The moral is that implication reversal is not itself a
homogeneous phenomenon: different contexts may reverse different sorts of entailments, and fine-grained differences in the representations of NPIs may have a significant effect on their distributions.

5. Incoherent Scales: NPIs in restrictor clauses

If an environment does not provide the appropriate inferential structure for an NPI then obviously the NPI will not be licensed. But this is not the only thing that can go wrong. Even with the right sorts of entailments, a variety of factors may prevent the formation of a coherent scalar model in which to interpret an NPI. This is what often happens to NPIs in restrictor clauses.

As noted above (example 2), minimizing NPIs are only sometimes licensed in the restriction on a universal quantifier. As the examples in 12-13 show, similar facts hold for minimizers and indefinites in the restrictions on most and few.

12. a. Most students who’ve ever read Hegel seem to wear hats.
   b. ??Most students who’ve read the least bit of Hegel seem to wear hats.
   c. Most students who’ve read the least bit of poetry will be familiar with Stevens’ “The Emperor of Ice Cream”.

13. a. Few of the guests who ate any trout dressed well.
   b. ??Few of the guests who ate a bite of trout dressed well.
   c. Few of the guests who ate a bite of trout enjoyed the meal.

In 12a the assertion of an unexpected and fortuitous correlation between the reading of Hegel and the wearing of hats poses no problem for the indefinite ever; but in 12b, substitution of the least bit into the same context sounds jarring. In 12c, with a natural connection between the two clauses, the least bit is fine.

Much the same thing happens with few in 13. Where there is apparently only a fortuitous correlation between the two clauses, as in 13a and 13b, the indefinite any is fine but the minimizer eat a bite is blocked. But when a causal relationship can be established the minimizer suddenly becomes acceptable.

Given the difference between minimizers and indefinites, the explanation for these facts is fairly straightforward: one needs more than a fortuitous correlation in order to construct a coherent scalar model. As suggested above, minimizers presuppose, by virtue of their scalar semantics, both a set of higher values which could fill in a propositional schema, and an ordering relation such that higher values will increase the likelihood of the proposition’s validity.

A sentence like 2c, repeated below, thus presupposes a correlation between the price of lettuce and the quality of restaurants such that the more a restaurant charges for iceberg lettuce, the more likely it is to have four stars.

2. c. ??Every restaurant that charges so much as a dime for iceberg lettuce happens to have four stars in the handbook.

But such a correlation clashes with our background knowledge about restaurants and lettuce prices: no matter how much one charges for it, iceberg lettuce is still not very fancy and charging more for it does not make a restaurant any better.

Similarly, in 12b the minimizer the least bit presupposes a scalar model according to which the more one reads Hegel, the more likely one is to wear a hat. And in 13b the minimizer eat a bite suggests a correlation between gastronomic indulgence and sartorial splendor such that the more trout you eat, the less likely
you are to dress well. In as much as one can accommodate such presuppositions, these sentences may be acceptable, but the correlations they require do not fit neatly with our normal assumptions about the way the world works. And for this reason, the sentences tend to sound, at best, rather peculiar.

With the indefinites, on the other hand, no such unlikely correlations need be assumed. A sentence like 12a with ever requires only that of all the students who have read Hegel, most seem to wear hats. The notion of random selection encoded by ever simply emphasizes the claim that students seem to wear hats no matter when or how much Hegel they might have read: it does not suggest that the more one read, the more likely one would be to wear a hat.7

While the differences between minimizers and indefinites in these contexts may be straightforward, we might still ask why either of them are licensed here in the first place. After all, the restrictions of few and most are non-monotonic, and it is not obvious how or why they can satisfy the IC. While I do not claim to have definitively solved this problem, I can at least make a few suggestive remarks.

I assume that the crucial generalization uniting these quantifiers is that they are all in some important sense proportional: in order to be verified they require a global evaluation over the set of possible instances. Thus, to determine the truth of a sentence like 14a one must know what happened to every guest who ate trout. Only then can one be sure that a majority got sick.

14. a. Most of the guests who ate trout got sick.
   b. Several guests who ate trout got sick.

With determiners like some and several this is not the case: to determine the truth of 14b one need only build up enough positive instances of sick trout-eaters to count as several, and it doesn’t matter what happened to the rest of them.8

Some evidence for the importance of proportionality comes from the fact, previously unnoticed in the literature, that quantifiers like many and some, on their proportional readings, also seem to allow NPIs in their restrictions. Thus a number of people I’ve surveyed find the sentences in 15 to be at least marginally acceptable, and certainly much better than the non-proportional equivalents in 16.

15. a. ?Sóme of the guests who ate so much as a bite of trout got sick.
   b. ?Mány of the guests who ate so much as a bite of trout got sick.

16. a. *Sóme (sm) guests who ate so much as a bite of trout got sick.
   b. *Mány (mny) guests who ate so much as a bite of trout got sick.

If this is accurate, it suggests that the proportional readings of these quantifiers may involve the same kind of global evaluation as is necessary for most, few and all, and that this global evaluation may indeed be what licenses NPIs.

A natural way to capture the requirement of global evaluation would be to represent sentences containing few, most or all in terms of an implication from the restriction to the nuclear scope. Since conditionals are scale reversing and, in a limited sense, downward entailing (Fauconnier 1979; Heim 1984), we might appeal to the effects of this hidden licenser to explain the licensing properties of the proportional quantifiers. But while the use of a conditional is straightforward in the representation of universal quantifiers, it might seem problematic for most and few. A sentence like All the guests got sick can be verified if one takes all the guests and checks that each one got sick, but if we substitute most for all, even
when the sentence is true, one could easily select a set containing a majority of the
guests for which it would not be the case that each one got sick.
In order to get around this problem, the quantifier must be understood as
applying only after the implication has been checked for all possible instances in
the domain of quantification. Thus for a sentence like 14a, one must evaluate for
every guest, G, the validity of the proposition If G ate trout then G got sick. If a
majority of the possible values for G satisfy this proposition then 14a will count
as true. And substituting few for most, if a minority of possible values for G
satisfy the proposition, then Few of the guests who ate trout got sick will be true.
If this is correct, it gives us a way of viewing these contexts as at least
secondarily scale reversing by virtue of the conditional. The proposal, while
admittedly rather sketchy, is hopefully at least not implausible. In any event, it
seems clear that there is something about the proportional determiners as a class
which allows them to license NPIs. Unfortunately, any more detailed
examination of this licensing will have to await another paper.

6. Contradictory Scales: NPIs in the focus of only
In this last section I will briefly consider why indefinite NPIs are licensed in
the focus of only and why minimizing NPIs are not. Again, the first question
seems to be rather the more problematic of the two. As the examples in 17
suggest, the focus of only does not appear to reverse implications.

17. Only people who drank something enjoyed the party.
18. Only people who drank guava nectar enjoyed the party.

17 clearly does not entail 18 and so it would appear that implications in this
context do not run from indefinites to particular individuals as the IC would
require; however, there is a sense in which an important part of the meaning of 17
does in fact entail an important part of the meaning of 18.
Roughly following Horn (1969), I assume that a sentence like 17, with term-
focus only, encodes both of the two propositions expressed in 19, and that the
first of these is presupposed, while the latter is asserted.

19. a. All the people who drank something enjoyed the party.
   b. None of the people who did not drink something enjoyed the party.

In the asserted proposition, 19b, the indefinite does not occur in an environment
which reverses implications. 19b does not entail 20, since someone who didn’t
drink guava nectar could still have drunk something else and enjoyed the party.

20. None of the people who didn’t drink guava nectar enjoyed the party.

But while the assertion does not satisfy the IC, the presupposition does. In 19a,
the focus of only has the form of the restriction on a universal quantifier, and as
noted above, this context is scale reversing. The hypothesis then is that NPIs are
licensed in the focus of only by virtue of this presupposition: that is, for
essentially the same reasons that they are licensed in the restriction on a universal
quantifier.
It may seem odd that a presupposition should license an NPI, but then there is
a sense in which 19a is not like other presuppositions: 19a need not be assumed as
part of the common ground in order for something like 17 to be felicitous. The sentence is perfectly acceptable as a response to the question *Did anyone enjoy the party?* in which clearly there is no assumption that anyone did. Thus while the proposition in 19a might not be asserted by 17, it can at least be conveyed as new information. I assume that only when this is the case will the presupposition be able to license an NPI.

But now the question arises: if the presupposition in 19a is available to license NPIs, why does it only license the indefinites? After all, the restriction on a universal quantifier generally does license minimizers, at least so long as an appropriate scalar model is available. But in 3b, the minimizer is clearly not licensed, even though a scalar correlation can presumably be made between the amount of Heidegger someone knows and their understanding of the movie.

3. a. Only people who've ever read Heidegger will understand this movie.
   b. *Only people who've read the least bit of Heidegger will understand this movie.

The problem, I suggest, lies not in the lack of a scale, but rather in a clash between the meaning of *only* and the scalar semantics required by the minimizers. As argued above, a minimizer incorporates the semantics of *even*: it represents the scalar element which is least likely to satisfy a given propositional schema and whose predication is thus maximally informative. But, as König (1991) points out, the scalar semantics of *only* has just the opposite effect, requiring that the element in its focus be the one most likely to make its sentence true. This is illustrated by the fact that in 21, *only* is fine with *the tallest* but odd with *the shortest*, since generally taller monkeys are the ones most likely to reach higher.

21. Only {the tallest/# the shortest} monkey could reach the banana.⁹

The jarring effect of 3b thus arises from the fact that the sentence makes two contradictory claims at once. On the one hand, because of the minimizer, it says that the people who are least likely to understand the movie are those who have read only "the least bit" of Heidegger (or none at all). On the other hand, because of *only*, it also says that the people who have read "the least bit" of Heidegger are the ones most likely to understand the movie.

This scalar contradiction explains the fact, noted by Lee and Horn (1994: 134), that *even* cannot occur in the focus constituent of *only*. This poses a problem for their claim that *any* incorporates the semantics of *even*, since, as 22b (Lee and Horn's 152b) shows, *any* can occur in this environment.

22. a. #Only those students who have even a single sibling need to complete the survey.
   b. Only students who have any siblings need to complete the survey.

On the present analysis, however, this fact is a virtue: *any* and *ever* are acceptable in the focus of *only* because, unlike the minimizers, they do not encode the least likely element within a scalar ordering. Rather, by virtue of their phantom semantics, they serve to trigger implications over an array of possible alternatives. Thus in 22b *any* can emphasize that everyone with a sibling must complete the
survey without suggesting that those who have more than one sibling are in any way more obligated to do so.

7. Conclusions

My goal in this paper has been to motivate what should be an intuitive claim: that polarity items occur in certain contexts and not in others because only certain contexts are compatible with their meanings. Along the way, I have run rather quickly through some quite complicated phenomena, but while important problems may remain, some points, hopefully, will be clear.

The first of these is that the semantics of NPIs is more subtle than what might be suggested by a simple constraint that they appear only in downward entailing contexts: it is not enough for a context to reverse entailments if the entailments reversed are not the ones an NPI demands. More importantly, it should be clear that the semantics of NPIs is not particularly complicated. The fact that NPIs are phantoms, and consequently subject to the implication constraint, is not in any way mysterious. By virtue of the IC, phantoms provide an economical way of making very general statements ranging over an array of instances. Thus while the status of any given form as a phantom may be a matter of lexical stipulation, the existence of phantoms in general is functionally motivated: they allow us to say a lot without a lot of effort. The moral is that simple functional considerations may have complicated grammatical consequences.

Endnotes

* This paper has benefitted from the comments of Raul Aranovich, Chuck Fillmore, Ron Langacker and John Moore. Special thanks are due to Gilles Fauconnier for his tireless and invaluable advice. Any remaining defects are entirely my own fault.

1. Differences between NPIs do not always lend themselves to a simple ordering of NPI strength. For example, semi-modal need and punctual until are both strong NPIs, but while need is fine in questions and bad with the negative adverb rarely (i-ii), until is fine with rarely but bad in questions (iii-iv):

   i) Need you be so rude to your grandmother?
   ii) *You rarely need study for these sorts of exams.
   iii) *Did Zelda leave the party until midnight?
   iv) Zelda rarely leaves a party until midnight.

Given such facts, it’s hard to know which forms are strong and which are weak. Unfortunately, that’s about all I can say about need and until.

2. Such an entailment is not strictly logical: there exist possible worlds, including this one, in which someone might be able to solve hard problems without being able to solve the easy ones. But the entailment does represent a default inference: it holds in all possible worlds which, all things being equal, one is likely to consider in making deductions. It is thus a ‘pragmatic entailment.’

3. Israel (1994) presents a taxonomy of four types of polarity sensitive items based on the interaction of these features. The taxonomy includes emphatic and understating versions of both negative and positive polarity items.
4. Note that this contradicts the recent proposals of Kadmon and Landman (1993) and Lee and Horn (1994), both of which represent *any* as encoding an essentially emphatic sense.

5. I am indebted to Fauconnier (p.c.) both for help in formulating the IC and for the general notion of phantom reference.

6. Roughly, a downward entailing operator is one which allows subset for superset substitutions *salva veritate*. For detailed discussion see Ladusaw 1979, Barwise and Cooper 1981, or van der Wouden 1994, *inter alia*.

7. Note that focus stress on *ever* forces the scalar reading and leads to ungrammaticality.

8. I am indebted to Ron Langacker for discussion of these points.

9. Note that, as one would expect, with *even* the judgements are reversed:

   21'. Even {# the tallest/ the shortest} monkey could reach the banana.
References


Word Order, Mutation, and Topic in Welsh
Nicholas Kibre
University of California, Santa Barbara

Overview:

Although Welsh has generally been described as a language with VSO or VSX constituent order, a number of clause constructions in the language deviate from this pattern. In this paper I argue that these reflect an organizing principle of Welsh constituent order, that the position immediately following a clause-initial verb or complementizer serves as a grammatical Topic. Under this hypothesis, the language’s tendency towards VSX clauses follows from the typically high topicality of subjects, and I will demonstrate that other orders occur precisely where subjects are un-topical, or where non-subjects are topical.

Several important lines of evidence corroborate the role of this Topic category in the organization of the Welsh clause. First, it can function as a pivot in conjoined clauses. Second, a comparative argument can be made based on a somewhat different, but plausibly related, aspect of clause structure in Irish. Finally, and perhaps most interestingly, it can serve as the basis for an elegant account of an interesting rule of consonant mutation.

1. Constituent Order in the Welsh Clause:

Welsh is generally described as a VSO language. In fact the situation is somewhat more complex; an increasing preference for periphrastic conjugations has meant that VSO clauses like 1.1, typical of literary Welsh, have come to be replaced by ones better described as Aux-S-(Particle)-V-O in spoken and more current written Welsh, as in 1.2. (For a overview of relevant Welsh sociolinguistics see Ball, Griffiths et al. 1988, Davies 1988, B. P. Jones 1988, D. G. Jones 1988, and other works collected in Ball 1988).

1.1 Febwthodd Huw Lgar.
PRT sell.3SPST NAME car
“Huw sold a car.” (Tallerman 1990)

1.2 Gwnaeth Huw Lwerthu car.
do.3SPST NAME sell.INF car
“Huw sold a car.” (ibid)

Note that words prefixed by $^L$ are ones which have undergone a consonant mutation termed Lenition—a point I will return to later.

Another non-VSO order occurs in all varieties of Welsh in clauses governed by certain complementizers, which must be SVO.

1.3 Disgwyliaidd Huw i Olwen Lwerthu car.
expect.3SPST NAME COMP NAME sell.INF car
“Huw expected Olwen to sell a car.” (Tallerman 1990)

These clause types differ in their relative orderings of subjects and main verbs. A number of authors have argued that these alternations can be explained transformationally, including Awbery 1976, Sproat 1985 and Sadler 1988. Less attention has been paid to the relative order of subjects and other clause arguments. Objects and obliques generally follow the subject in verb-initial clauses, and follow
the verb in verb-medial clauses, but in some cases, elements besides subject can immediately follow the verb. This is only a property of specific constructions, making it tempting to write them off as idiomatic or irregular; nevertheless, there is a functional cohesion to them. Three of these constructions are particularly common, and will be reviewed in the sections below.

A Presentative Construction:

Although Welsh has a special presentative copula, this has been supplemented and partly replaced by a construction borrowed from the English “there + be”. However, this construction has been nativized to Welsh’s verb-initial syntax by reversing the order to “be + there”, as in 1.4 and 1.5.

1.4 roedd 'na laddwy yn cystadlu COP3SPST there two PRT compete.INF
“There were two competing…” (Radio Cymru)¹

1.5 bod 'na stori tu allan i 'r taled L-ryddiaith COPINF there story side back to the medal prose
“(we know that) there’s a story behind the prose medal…” (Radio Cymru)

The Possessive Construction:

Welsh does not have a verb “to have”. Like Finnish and Russian, it uses the copula to describe possession (when the possessum is indefinite, this is the same presentative copula which elsewhere is being replaced by bod + 'na), making the possessed item the subject and the possessor an oblique, marked by one of the prepositions gan or (gy)dà, both of which mean “with”.

The order of a possessive clause can be the usual VSX, that is “be+thing+with+owner”, as in 1.6a; but more often the possessor comes first, as in 1.6b.

1.6a Mae llawer o arian gyda nhw. COP3S a.lot of money with 3P

1.6b Mae gyda nhw llawer o arian. COP3S with 3P a.lot of money
Both: “They have a lot of money” (Jones 1992).

An Obligative Construction:

In Welsh, obligation (“to have to”), is traditionally expressed with an expression we can loosely sketch as “be+need+for”+NP+VP. This can be seen in example 1.7.

1.7 efallai fydd rhaid ini L-gynnal ysgolion perhaps COP3SFUT need for.1P support.INF schools
“Maybe we will need to support schools.” (Radio Cymru)

In speech and less formal writing, however, the copula is usually dropped in the present tense, as in 1.8.

1.8 rhaid i ni L-geisio rhedeg ei eglwys need for 1P try.INF run.INF 3MGEN church
“we need to try to run his church” (Hedyn)

In a sense, rhaid has been grammaticized into an auxiliary. But what seems to function as its subject, (“us” in 1.9) is still an oblique, governed by i.
2. Evidence for a Topic Category in Welsh

In each of these three cases, some sort of oblique or adverbal element is in the position immediately following the clause’s main verb (in terms of dependency), or verbs in the case of 1.7. The grammatical roles of these three obliques are different, but this should not keep us from asking whether some functional similarity unifies this similarity in their surface order.

The hypothesis I want to propose is the following:

— The position immediately following the (hierarchically) main verb of a clause, or sequence of verbs, or the complementizer i, codes a grammatical Topic category.

— This category expresses the discourse function of clause Topic (Givón 1983), or starting point (Chafr 1976, 1994).

In a transformational analysis, we might unify reference to “main verb” and “complementizer” with some kind of V-to-COMP-raising rule. For now I will just assume that both verbs and complimentizers serve as some kind of “cognitive reference point” which arguments can be positioned relative to.

As in English, the subject in Welsh is largely the home of topical participants. Thus it is unsurprising that subjects are generally realized in the Topic position. Notably, though, the non-VSO-type constructions we have examined are all ones in which the relationship between subject and topicality breaks down.

In the rhaid obligatory construction, there is no true subject to compete in topicality with the NP marked with i, and this is also the “logical” subject of the semantically main verb. In the presentative, we can surmise that the subject is bumped from the Topic position by the pleonastic yna because it is new and discontinuous, or quintessentially un-topical, expressing this fact is the point of a presentative construction. In the possessive, the topicality of the possessor seems intuitively correct because it is coded as a subject in languages which do have a verb “to have”. And whereas a possessor is usually inanimate, possessors are human.

Intuitions, of course, are only the beginning of an explanation. In the next section I will make a more empirical argument for this hypothesis in terms of discourse patterns associated with one of these constructions, the possessive.

Discourse Evidence for the Topic Category:

Topicality can be measured. In Givón 1983’s methodology, topicality is quantified in terms of the quantity of co-referential NP’s in preceding and following clauses. Chafr 1994, on the other hand, has suggested less empirical but more subtle measures in which the analyst, partly intuitively, must classify referents as given, accessible or new.

Another way is to observe how speakers code NP’s. Givón 1983:18, 1990:905 has proposed that the “weight” of an NP is proportional to its discontinuity, where weight ranges on the following scale:

Zero Anaphora < Bound Pronominals < Free Pronouns < Lexical NP’s

If we accept this, then we can measure the tendency of sentence positions to be have topical referents by quantifying the relative frequencies of lexical NP’s, pronouns, and nulls.

Zero anaphora occur only in restricted environments in Welsh (such as in the “gapped” arguments of relative clauses), and the distinction between free and bound pronouns is not always clear. But the proportion of lexical NP’s vs.
“lighter” alternatives still gives a good measure of topicality. To support the Topic-category hypothesis, I have measured the proportion of lexical NP’s in subject and oblique positions in “normal” clauses, and compared this with the level of anaphora in the two arguments of possessive clauses.

For the first measure, I selected 128 clauses from an informal discussion session broadcast on Welsh radio which were finite, active, and non-relative. Within this set, obliques were lexical NP’s significantly more often than subjects; obliques were full NP’s in 45 out of 61 tokens (74%), while subjects were full NP’s in 36 out of 127 (28%). (Only object arguments from transitive verbs were measured, and one weather-type verb was considered to be inherently subjectless).

In possessives, this trend was reversed. Here the subject was usually a full NP, and often a indefinite/new one at that, as in 2.1.

2.1 Felly does gyda chi lddim straeon newydd ini
    thus COPNEGEXIST with you no stories new for.1P
    heno te?
    tonight PRT

    “so you don’t have any new stories for us tonight then?” (Radio Cymru)

This text had only four examples of possessive clauses, so another written text, a (semi?)-autobiographical novel (Rhwng Dwy; see note 1), was searched for examples of the possessive construction in particular. Here, too, possessives proved to have topicality patterns opposite to those measured for non-possessive clauses: Possessor obliques were full NP’s in 3 out of 24 tokens (12.5%), and possesa subjects were full NP’s in 23 out of 24 tokens (96%).

The tendency towards the V-Possessor-Possessem pattern was fairly strong in this data. In 21 of the 24 examples (87.5%), the V-Possessor-Possessum constituent order was found. Interestingly, the lone anaphoric possessum in this data set, hon in 2.2, was one of the remaining three.

2.2 y mae hon gen i ers saith mlynedd
PRT is this with 1S since seven years

    “I’ve had this for seven years.” (Rhwng Dwy).

    This suggests that speakers can actively manipulate these word order
alternatives for information flow purposes, but for now this has to be left as a
hunch which larger corpora will be needed to test.

In this section I have examined the relevance of the proposed Topic category
to information flow. In the next I will examine its importance to other aspects of

**Topic as the Pivot Category In Welsh:**

Welsh generally seems to be a subject-pivot language in the sense of pivot
developed by Dixon 1972, 1979, and Comrie 1978. If two conjoined clauses have
a co-referential subject, it may be omitted in the second (even in informal varieties
where subjects are not usually dropped). That is, we find the pattern: (Verb Subj
Obj) conj (Verb Subj; Obj), but not: (Verb Subj; Obj) conj (Verb Subj Obj), nor,
(Verb Subj Obj) conj (Verb Subj; Obj).

But this generalization does not hold when impersonal verbs are involved.
In addition to first through third person singular and plural forms, each tense and
mood paradigm of the Welsh verb has a special impersonal form. The impersonal
resembles a passive (it can have a demoted agent as a chômeur) but is best
described as a subjectless active; that is, a voice in which the subject is demoted or dropped but the object is not promoted. There are several arguments for this analysis.

First, intransitive verbs can be impersonal, as in 2.3.

2.3 *Rhedwyd* yno. run.IMPRSPST there "One ran there."

Second, in impersonal periphrastic clauses, the patient comes in the usual position for objects, following the infinitive main verb.

2.4 *Yr ydys yn gweld ci.* PRT COPIMPRS PRT see.INF dog "One sees a dog."

On the other hand, in non-periphrastic clauses, pronominal patients can be coded by a special set of "infixed" object pronouns cliticized between the main verb and a sentence-initial particle.

2.5 *Fe ‘i darledir L-bum noson yr wythnos.* PRT 3F broadcast.IMPRS five night the week "It is broadcast five nights a week." (Chi biau)

However, Sadler 1988 has observed that the objects of impersonals seem to count as pivots, just like subjects. They can be deleted if conjoined to a clause with a co-referential subject, as in 2.6, and subjects can be deleted in conjoined clauses if co-referential to them as in 2.7.

2.6 *Canodd Siôn-i a cymeradwywyd Ø gan y gyntuleidfa. sing.3SPST John and applaud.IMPRSPST with the audience. "John sang and was applauded by the audience." (Sadler 1988:228)

2.7 *drawyd y cwch-i gan L-don anferth a dymchwelodd Ø-i hit.IMPRSPST the boat with wave huge and capsized.3SPST "...the boat was hit by a huge wave and capsized" (Hedyn)

Shibatani 1985’s analysis of passives and impersonals claims that both involve “agent defocusing”. Assuming that defocused agents are non-topical, these examples can be explained if we assume the following:

— Objects of Welsh impersonals are not promoted to subject, but as the most topical NP’s available they do advance to Topic.
— The pivot category of Welsh is not actually Subject but Topic. The subject just appears to be pivot most of the time because subjects are usually in the Topic position.

The relevance of the Topic for other aspects of syntax is strong evidence that this category does exist, and plays an important role in the organization of the Welsh clause.

**Comparative Evidence:**

Evidence for the proposed Topic category can be found not only within Welsh but in comparison with at least one of its relatives. Noonan 1994 has proposed an analysis of several voice constructions in Irish in terms of an
"Information Structure Map" underlying certain types of clauses. Particularly, he makes the following proposal:

"In clauses with two arguments, immediate postpredicate position is reserved for arguments of high topicality; the position following this and prepredicate position are used for arguments of low (or, at least, lower) topicality."

Noonan 1994:295

My proposal for Welsh differs from this in several ways, but the "Information Structure Map" I have proposed seems like a plausible cousin to this one. Both claim high topicality for a position immediately following an element which may be thought of as central to the structure of the clause; in Noonan’s Irish proposal, this centrality is defined semantically, whereas in the present hypothesis for Welsh centrality is defined in terms of structural dependency. In non-periphrastic main clauses these will be conflated, and while it would go beyond the scope of this paper and the available evidence to try to trace the topicality-marking strategies of Irish and Welsh to a common ancestor, it seems reasonable to conclude that they have one.

3. Consonant Mutation:

So far, I have defined the Topic category in positional terms, and looked for evidence linking the function of Topic to this position. In this section I will argue that the Welsh Topic is also coded morphologically through a rule of initial consonant mutation.

Grammaticized Sandhi:

Initial consonant mutations are a set of gradations in the manner of articulation of word-initial consonants. The phonological category of the mutation of interest to us here is traditionally termed Lenition. Lenition consists of the following changes:

- Voicing of stops: p, t, c → b, d, g.
- Voicing of ll (=l/l) and rh (=r/h) into l and r.
- Frication of voiced stops: b, d, m → f (=f/vl), dd(=d/f), f (=f/vl).
- Elision: g → Ø.

Mutations are phonologically opaque: they are usually triggered following specific particles, but the shape of a particle does not predict which if any mutation it will cause. For example, the form ei means "his" if followed by Lenition, or "her" if the word following undergoes another mutation called Aspiration; furthermore, the homophonous (at least in Southern Welsh) pronoun eu "their" can only be distinguished from these in speech by the fact that it causes no mutation at all.

Mutations can also grammaticize into grammatical markers in their own right. For example, attributive adjectives are lenited when the noun they are modifying is feminine singular.

3.1a cimasc bach
    dog  small
    "A small dog."

3.1b cathFem l-fach
    cat   small
    "A small cat."
This mutation has sometimes been described as triggered by feminine adjectives, in the same sense that a particle such as ei. "his" is a Lenition trigger (see, for example, Ball & Müller 1992:161-164). However, the cases in 3.2 show that this mutation is based on the structural relation of dependency rather than the simple presence of a preceding feminine noun in front of the adjective.

3.2a \[ \text{gwr}_{\text{Masc}} \circ \text{ldeddylfryd}_{\text{Fem}} \circ \text{cyffelyb} \]
man of action similar
"A similar man of action."

3.2b \[ \text{gorsaf}_{\text{Fem}} \circ \text{radio}_{\text{Masc}} \circ \text{l\text{oblogaidd}} \]
station radio popular
"A popular radio station." (Chi biau)

For a thorough overview of mutation and approaches to the problem see Ball & Müller 1992 generally. For a more theoretically-driven look at the subject, as well as further arguments for the agreement analysis of this mutation, see Kibre 1995.

**Syntactic Mutation and the XP-Trigger hypothesis:**

The mutation of interest here is another one triggered in syntactically defined environments. In the past it has been analyzed a kind of morphological marker akin to the gender agreement marker described above (Comrie 1975, 1976, Lieber 1983, 1987, Perlmutter & Postal 1983, Zwicky 1984) but never entirely successfully. A number of other accounts have been proposed, but since their solutions (although interesting) fail the test of descriptive adequacy, I will not go into them here (Fife 1992, Hannahs 1993; for a more detailed review of these proposals see Kibre 1995).

The most successful approach to date has been the NP- or XP-trigger hypothesis proposed in various forms by Harlow 1981, 1989, Sproat 1985, Borsley 1986, Tallerman 1990 and Borsley & Tallerman 1993. Since this model can at least account for most of the data, it seems like a good place to start. As its name suggests, this hypothesis states that words are subject to Lenition when immediately preceded by a maximal projection: NP, AP, PP, or VP.

This model makes a number of predictions which will be seen to be correct, if I may direct the reader to glance back at some of the earlier examples in this paper. Recall that words which have undergone Lenition have been marked by a preceding superscript \( L \) (for completeness I have marked all lenited words, including those whose mutations have been triggered by particles, so it should be understood that the statements below are not meant to account for all of the lenitions in the examples above). The environments for which it predicts Lenition include the following:

- Initial words in objects are mutated in VSX clauses (\( \text{gar}<\text{car} \) in 1.1).
- Non-finite verbal complements are mutated in the object position of VSX clauses as well (\( \text{werthu}<\text{gwerthu} \) in 1.2).
- Objects are not mutated in periphrastic constructions: (\( \text{car} \) in 1.2); and neither are infinitives in this position.
- In presentatives, the subject is mutated after \( \text{yn}a \), which could be treated as a PP. (\( \text{ddwy}<\text{dwy} \) in 1.4)
In possessives, the possessum subject is mutated when the possessor oblique, a PP, comes first (lawer<llawer in 1.6b) but not when the VSX order is maintained (1.6a).

In the rhaid-obligative construction, the main verb is mutated after the PP with the agent: (gynnal<cynnal in 1.7; geisio<ceisio in 1.8)

There seems to only be one case where an XP is not followed by mutation, which I will turn to shortly. There are, however, several cases where this mutation applies even though there is not a preceding XP. Literary, although not colloquial Welsh, allows null subjects in main clauses. As seen in 3.3, Objects are nevertheless mutated even though the subject NP is missing, as are verbal complements.

3.3 Gwerthodd Ø Lgar.
sell.3SPST car
“He/she sold a car.” (Tallerman 1990)

Objects and verbal complements are mutated after the empty subject position of clauses relativized off of subject, and in a number of subject-fronting constructions.

3.4 dyn (a) Lwerthodd Lgar
man REL sold.3SPST car
“a man who sold a car” (ibid)

This includes cases where the subject is a fronted wh-word, or is focus-fronted.

3.5 Pwy (a) Lwerthodd Lgar.
who REL sold.3SPST car
“Who sold a car?” (ibid)

Imperative verbs also lack overt subjects, but their objects and infinitive complements are mutated as well.

3.6 Rho Lwyddiant i mi heddiw.
give success to me today
“Give me good fortune today.” (Gen 24:12)

3.7 Cofiwch Llaw yn Siôp y Penhe.
remember call in Shop the things
“Remember to stop by a Siôp y Penhe.” (Advertisement)

Proponents of the XP-trigger hypothesis propose to account for this data by allowing null NP’s to trigger mutation as well as overtly realized ones. Although this solution works, it may be problematic from a phonological standpoint; Nespor & Scoretti 1985 and Nespor & Vogel 1986 have argued that empty categories are phonologically invisible. Although not fatal to the XP-trigger hypothesis, this is an objection which should be addressed.

Another problem for the proposal is that the empty subjects of impersonal verbs do not cause mutation to following objects and complement verbs. For example, note that pob has not become Lbob in 3.8, and darlledu has not lenited to Lddarlledu in 3.9.

3.8 Cedwir pob hawl.
keep.IMPRS every right
“All rights reserved.” (Publisher’s notice)
3.9 cychwynnwyd darledu 'n Lfyw y Cwestiynau Seneddol
start.IMPRS broadcast PRT live the questions senatorial
Cymreig Welsh
“live broadcasting of Welsh Parliamentary Questions began” (Chi biau)
A solution to this problem, proposed by Harlow, is to say that cased NP’s
(which include “real” NP’s, the “little pro” of pro-drop clauses, and wh-traces)
cause mutation, but that this empty subject is uncased “big PRO”, so it does not.
The one remaining descriptive hurdle for this approach I am aware of arises
in conjoined NP’s like those in 3.10, where the first NP fara fails to trigger
Lenition on the following menyn.

3.10 Bwytais i Lfara. menyn/*L-fenyn a chaws
ate.1S PST 1S bread butter and cheese
“I ate bread, butter, and cheese.” (Borsley & Tallerman 1993)

Borsley & Tallerman 1993’s solution to this problem is to propose that bara
and menyn are actually separated by a phonetically null conjunction, and that this
conjunction absorbs the mutation triggered by the first NP. There are several
problems with this proposal. First, as with the proposal that empty NP’s can
trigger mutation, there is the issue of the phonological relevance of phonologically
null constituents which needs to be addressed.
The second appears when we consider this proposal in relation to the
feminine-agreement mutation of adjectives discussed earlier. In 3.11, we must ask
why no invisible conjunction is present between the adjectives “tall” and “strong” to
prevent the mutation of the latter.

3.11 merch *dal L-gref a doeth
girl tall strong and wise
“a tall, strong, and wise girl” (Ball and Müller 1992)

A Topic-Oriented Mutation Analysis:
Roughly speaking, the Harlow, Tallerman, Borsley proposal makes the
following descriptive statement:

— In a series of XP’s, the first word of all but the first XP is subject to
Lenition.

I have argued that among the constituents following a verb or
complementizer, the first has the special status of a grammatical Topic. Putting
these two observations together, a new analysis of the mutation suggests itself: All
the rightward dependents of a verb or a complementizer are subject to Lenition on
their first word, except for the Topic. In essence, this claims that the Topic
category is coded both by constituent order, and the use of a “mutation-morpheme”
similar to that of gender mutation to mark non-topics.
Notably, this solution correctly differentiates between the empty subjects of
relative clauses, pro-drop clauses, sentences with fronted subjects, and those in
impersonals. In the first group, the subject is known, and still highly topical—the
fact that its referent can be recovered from zero demonstrates its topicality. On the
other hand, I have already argued that the objects of impersonals are topical, and we
have seen structural evidence that they can function as Topics—thus their non-
mutation is explained.
Furthermore, this proposal escapes the conundrum which conjoined NP’s present to the XP-trigger hypothesis. In 3.10, for example, it will correctly assign Lenition to the first word of the object noun-phrase *y fara. menyn a chaws*. but leave *menyn* unaffected since it only deals with constituents of clauses, not of compound NP’s.

**Form beyond Motivation:**

The Topic-centered model of mutation I have proposed is as descriptively adequate as the XP-trigger hypothesis, but the question remains open how well it explains the facts we have observed. In particular, we need to ask precisely how the Topic/Non-Topic distinction claimed to be encoded by position and mutation is related to actual discourse topicality.

First, note that although I have argued for the topicality of possessors in the possessive construction and the obliques marked by *i* in the *rhai’d*-obligative, and for the un-topicality of subject in the be+*yma* presentative, alternatives to the standard VSX clause order sometimes arise through stylistic variation. It is possible, for example, for adverbial obliques to appear between verbs and subjects, mutating the latter.

3.12 *Roedd yr bryn yferch.*
**COP3SPST** on the hill girl
“A girl was on the hill.” (Tallerman 1990)

I have also found one example, unusual but too colorful to omit, in which the subject and verb are separated by a non-finite VP, and here the subject mutates too:

3.13 *Yma mae yn gorwedd ygorph Richard Roberts.*
**here COP3S** PRT lie.INF body NAME
“Here lies the body of Richard Roberts.” (Tombstone near Llanfairfechan, North Wales)

It has been suggested to me (Tom Shannon, P.C.) that such sentences might be fit into the current model of the Welsh clause with a slightly more broadly defined notion of Topic, expanded include something like the Prague school notion of Theme (see, for example, Firbas & Golková 1975, Firbas 1986, and Danes 1984). That is, although *ar y bryn* and *yn gorwedd* do not code topical entities in the Chafe/Givón sense (that is, high continuity referents) they do present background information into which the new referents “a girl”, and “Richard Roberts” can be added.

Another apparent problem for the topicality model of consonant mutation are presented by sentences such as 3.9. In order to explain how the infinitive verbal complements of impersonals escape mutation, such as the verb *darlledu* here, we have to assume that infinitive VP’s are selected as Topic; infinitives have many nominal characteristics in Welsh, and are traditionally called verb-nouns, but the discourse category of topicality is only really relevant to true NP’s.

These cases are informative, but do not really disprove the functional explanation I have proposed. This hypothesis attempts to explain the motivation for certain patterns in Welsh grammar, but it does not claim that this mutation functions as an “untopic marker” in each and every sentence. The mutation may have been developed to express aspects of information flow, and still does so in most cases—but once absorbed into the grammar of a language, a category may take on a life and momentum of its own.
In addition to previous linguistics publications, which have been cited, examples are taken from several written and spoken texts:

-Chi Biau: Chi biau BBC Cymru/Wales. BBC Cymru/Wales. Canolfan y BBC, Llandaf, Caerdydd CF5 2YQ. (Pamphlet guide to radio and television programs)
-Hedyn: Hedyn: O Gymru, Edigarhech. Mudiad Efengylaid Cymru, Bryntrion, Pen-y-bont ar Ogwr, Canol Morgannwg CF31 4DX. (Evangelical pamphlet)

Formalist accounts of Welsh syntax have been proposed in this vein by Jones & Thomas 1977, Harlow 1981, and most recently Sproat 1985—although it should be noted that Sproat classifies i and auxiliaries as instances of INFL, and considers lexical inflected verbs to be raised to INFL rather than COMP.

3For example, /genin/ "with me" can be analyzed as a single inflected form of gan, "with" or as an inflected preposition gen plus a free pronoun i; it is freely written as one word or two.

4Eve Sweetser has suggested an alternative explanation for the subject-like characteristics of impersonals' objects: the impersonal is being reanalyzed as a true passive. This is not unreasonable, since the impersonal is a largely literary form in modern Welsh, and is being replaced by a true passive in the spoken language. Space does not permit me to address this point here at length, but I have taken up the issue elsewhere (Kibre to appear).

In any case, a crucial question here is whether impersonals' objects were pivotal before the construction's decline or have only acquired this quality recently. Unfortunately I do not have an answer to this at present, and will have to leave the matter open to future research.

References:


English Negation from a Non-Derivational Perspective*

Jong-Bok Kim
Stanford University

Introduction

A different perspective on grammatical research, even on well-known phenomena, can sometimes provide us with arguments for a quite unexpected and new analysis. This paper claims that if we accept the view that the English negator *not* can be either a modifier or a complement, we can offer a more straightforward and explicit explanation for English negation than those couched in terms of head-movement and functional projections including NegP (e.g., Pollock 1989).

1 Basic Properties of English *Not*

The English negator *not* behaves in very much the same way as negative adverbs like *never*. Their similar distribution is particularly clear in nonfinite verbal constructions, such as gerundive, infinitival and bare verb phrases, as can be seen, by comparing (1) and (2).

(1) a. Kim regrets [never [having seen the movie]].
   b. We asked him [never [to try to call us again]].
   c. Duty made them [never [miss the weekly meeting]].

(2) a. Kim regrets [not [having seen the movie]].
   b. We asked him [not [to try to call us again]].
   c. Duty made them [not [miss the weekly meeting]].

But there also exist several properties which distinguish *not* from a negative adverb like *never*. One obvious fact that distinguishes the negator from the negative adverb is the so-called *do*-support phenomenon in English. The examples in (3) show that in non-auxiliary finite verb phrases, the particle *not* requires the dummy verb *do*. But a true negative adverb like *never* has no such a requirement.

   b. Tom did not borrow Mary's book.
   c. Tom never borrowed Mary's book.

A more striking property of the negator *not* lies in VP deletion. As noted by Baker (1972), Sag (1976) and others, VP deletion immediately after adverbs like *never* is not permitted, as illustrated in (4).

(4) a. Tom has written a novel, but Peter never has ___.
   b. *Tom has written a novel, but Peter has never ___.
However, a peculiar property of not is found in finite clauses. Consider the examples in (5).

(5) a. Tom has written a novel, but Peter has not __.
    b. Mary has finished her homework, but Peter has not __.

It is possible to elide VPs after a not following a finite auxiliary, unlike VPs that follow adverbs.

Any analysis of English negation thus needs to account for the essential properties of not: On the one hand it behaves like the negative adverb, but on the other hand it exhibits certain differences. My purpose here is to provide a non-derivational analysis of both these properties.

2 Not as a Modifier

In capturing the adverbial and non-adverbial properties of the negative particle, the proposed analysis starts from the three basic assumptions summarized below.

- The negative particle not and the negative adverb never are both preverbal adverbs and hence both modify a VP (cf. Ernst 1992, Baker 1991).

Given these basic assumptions, let us consider how we can deal with English negation. The first claim I made is that not is an adverb that modifies a VP. This claim then implies that the negation can randomly occur as a VP modifier unless otherwise constrained. However, one peculiar property of not is its restricted occurrence: first of all it cannot precede a finite verb.

(6) a. *John [not [left]].
    b. *John [not [has gone]].
    c. *John [not [is leaving]].

In capturing this empirical generalization, I adopt the framework of HPSG (Head-driven Phrase Structure Grammar) which allows lexical representations to be formulated with considerable precision. (7) is a simplified lexical entry for the negator not.¹
(7) not: \[
\begin{array}{c}
\text{HEAD} \quad \text{adv}\text{[MOD VP[nonfin]:2]} \\
\text{CONTENT} \quad \text{RELATION} \quad \text{not} \\
\text{ARG} \quad \text{2}
\end{array}
\]

The lexical entry (7), following Pollard and Sag’s (1994) analysis, specifies that *not* selects for the nonfinite VP that it modifies. This selectional relation between the modifier (adverb) and the modified element (nonfinite VP) is manifested by the value of the head feature MOD(IFIER). Also the CONTENT value represents that the negation semantically takes the meaning of the modified VP (2) as its argument.

The lexical entry (7) readily captures various distributional possibilities of *not*, first of all.²

(8) a. *John [not VP[fin]has gone]].
    b. *John certainly [not VP[fin]talked to me]].
    c. *John [not VP[fin]always agreed with me]].

As shown in (8), *not* cannot modify a finite VP. But, as is clear from the examples in (9), it can modify any nonfinite VP.

(9) a. I saw John acting rude and [not VP[ger]saying hello]].
    b. I asked him to [not VP[base]leave the bar]].
    c. Kim has [not VP[part]been drinking the wine]].

Further we need no additional statement to account for its position in coordination sentences like (10).

(10) a. John will [[not [walk]] and [talk]]]
    b. John will [[not [walk]] and [talk]]
    c. John will [[walk] and [not [talk]]
    d. You can [[walk for miles] and [not [see anyone]]

In each case the negation modifies a base form VP, satisfying our lexical specification. The analysis also correctly predicts that the postfinite *not* can either scope over only the first conjunct or over both conjuncts, as shown in (10)a and (10)b. But if we accept the general assumption that only categorically identical constituents can be coordinated, an analysis adopting the NegP hypothesis seems to run into a problem: examples in (10) would force the NegP hypothesis to change this well-accepted assumption. Under the NegP hypothesis, where the overt negation occupies the head of NegP, the sentences in (10) would be NegP and VP or VP and NegP coordinations. And even one allows these non-identical constituents to be conjoined, an explanation must still be provided for the impossibility of coordinations like CP and IP.³

3  *Not* as a Complement

My analysis has provided a clean and simple way of accounting for much of the distributional possibilities of *not*. But there still remain facts to be accounted for, especially with respect to VP deletion.
3.1 VP Deletion

As shown in section 1, one peculiar property of not comes from VP deletion. Unlike other adverbs, the negation immediately after a finite auxiliary can be stranded after VP deletion. Before laying out our analysis, let us first consider general cases where VP deletion has been applied. Consider the contrast between (11) and (12).

(11) a. Kim can dance, and Sandy can __, too.
   b. Kim has danced, and Sandy has __, too.

(12) a. *Mary considered joining the navy, but I never considered __.
   b. *Mary wanted to go and Sandy wanted __, too.

These illustrate the standard generalization that VP deletion is possible only after an auxiliary verb. In capturing this, I first assume the VP deletion lexical rule given in (13), incorporating the main idea from the VP deletion metarule of Gazdar et al. (1982).

(13) VP Deletion Lexical Rule:

$$\begin{bmatrix}
\text{HEAD} & \text{verb} [+\text{AUX}] \\
\text{COMPS} & \langle \text{VP} \rangle
\end{bmatrix} \Rightarrow \begin{bmatrix}
\text{COMPS} & \langle \quad \rangle
\end{bmatrix}$$

The lexical rule in (13) guarantees that VP deletion applies only to an auxiliary verb selecting a VP complement. Taking an auxiliary verb as its input, this lexical rule thus gives us as output another lexical entry whose VP complement is not realized. The verbs such as can and has in (11) are all auxiliary verbs ([+AUX]) and subcategorize for a VP complement. Thus, the VP complement of all these verbs can undergo the lexical rule. But the lexical rule cannot be applied to verbs such as consider and want in (12), simply because they are non-auxiliary verbs ([-AUX]).

Now, let us return to the issue of VP deletion after an adverb. One important constraint on VP deletion is that it cannot apply immediately after an adverb, repeated here in (14).

(14) a. *Kim has never studied French, but Lee has always __.
   b. *Tom has written a novel, but Peter has never __.

One simple fact we can observe from (14) is that adverbs cannot modify an empty VP. In the framework of HPSG, VP modifying adverbs carry at least the lexical information given in (15).

(15) $$\begin{bmatrix}
\text{HEAD} & \text{adv}[\text{MOD VP}: \text{adv}] \\
\text{CONTENT} & \begin{bmatrix}
\text{RELATION} & \text{adv-rel} \\
\text{ARG} & \text{2}
\end{bmatrix}
\end{bmatrix}$$
The lexical entry in (15) simply states that the adverb with this lexical information modifies a VP. The head feature MOD guarantees that the adverb selects the head VP it modifies. Given Sag and Fodor's (1994) traceless theory, an ungrammatical example like (14)a would then have the structure given in (16).\(^5\)

(16) \[
\begin{array}{c}
\text{have} \\
\text{Adv[MOD VP]} \\
\text{always}
\end{array}
\]

HPSG has a small set of schemata, analogous to X' schemata, which specify partial information about universally available types of phrases. The adjunct schema is one of the universally available options for well-formed phrases. This adjunct schema roughly says that an adjunct and the head it selects for through its modifier feature (MOD) forms a well-formed phrase. Now look at the structure in (16). In our lexical theory where a VP modifier (e.g., always and never in (14)a,b) selects for its head VP through the head feature MOD(IFIER), the absence of this VP then means that there is no VP the adverb can modify. And this results in an ill-formed structure: no universal schema in HPSG renders such a structure acceptable, thus explaining the ungrammatical of (14)a,b.

One more striking property of not with respect to VP deletion that we have not discussed yet is that the negator not after a finite auxiliary can be stranded, as illustrated in (17).

(17) a. Kim said he could have heard the news, but Lee said that he could not

   b. *Kim said he could have heard the news, but Lee said that he could have not ___ .

If the negation not in (17)a and (17)b were taken to be a modifier, we would predict both of these examples to be unacceptables since in both cases there is no VP for the negation to modify.

To account for the puzzling contrast between (17)a and (17)b, I adopt Warner's (1994) suggestion that there is another way to introduce the adverb not into syntax. Borrowing the general technique of categorial grammar which allows a functor to be type-shifted to an argument or vice versa,\(^6\) I assume that a verb can be 'converted' into another verb selecting for the negation not as an additional complement (or equivalently, converting the modifier not to a syntactic complement) via a lexical rule as given in (18).\(^7\)
(18) English Adverb Conversion Lexical Rule:

\[
\begin{align*}
\text{HEAD} & \quad \text{verb[+AUX, fin]} \\
\text{COMPS} & \quad \langle \Pi \text{VP} \rangle \\
\text{CONTENT} & \quad \overset{2}{\text{2}}
\end{align*}
\implies
\begin{align*}
\text{COMPS} & \quad \langle \text{Adv}_{I, \Pi} \text{, VP} \rangle \\
\text{CONTENT} & \quad \overset{3}{\text{3[ARG} \overset{2}{\text{2}}} \text{]}
\end{align*}
\]

The lexical rule in (18) takes as input any finite auxiliary which selects for a base VP complement and yields as output another verbal entry which adds the negation as another complement, i.e., adds it onto the finite verb’s COMPS list. The lexical rule also has a semantic effect: the converted complement adverb’s content (3) becomes the main content of the lexical rule output, with the meaning of the input (2) being its argument.

Now notice that the output of this lexical rule in (18) can be the input of the VP-deletion lexical rule. We have seen that any auxiliary selecting for a VP complement can undergo this lexical rule. Thus, nothing prevents the output of the Adverb Conversion Lexical Rule from undergoing the VP-Deletion Lexical Rule, as illustrated in (19).

(19) Applying the VP Deletion Lexical Rule:

\[
\begin{align*}
\text{HEAD} & \quad \text{verb[+AUX, fin]} \\
\text{COMPS} & \quad \langle \text{Adv}_{I, \Pi} \text{, VP[bse]} \rangle
\end{align*}
\implies
\begin{align*}
\text{HEAD} & \quad \text{verb[+AUX, fin]} \\
\text{COMPS} & \quad \langle \text{Adv}_{I} \rangle
\end{align*}
\]

Now let us turn our attention to VP deletion after the negator not (relevant data repeated in (20)).

(20) a. Susan may have been studying, but Mary may not __.
    b. *Susan may have been studying, but Mary may have not __.
    c. *Susan may have been studying but Mary may have been not __.

Given the input domain of the Adverb Conversion Lexical Rule, the negation not in (20)a can be converted to the complement of the finite auxiliary, allowing the structure in (21).

(21) \[
\begin{array}{c}
V[+\text{AUX}] \\
\text{could}
\end{array} \quad \overset{\text{VP}}{\overbrace{\text{Adv}[\text{MOD VP}]}} \\
\text{not}
\]

Notice that the phrase [could not] in (21) forms a well-formed head-complement structure where not is the complement of the head could. Nothing blocks this structure. Under this analysis, the ungrammaticality of (20)b and (20)c also falls out naturally. The negations in (20)b and (20)c are just modifiers. They cannot be complements. We have seen that an adverb requires the VP it modifies to be present in order to form a well-formed structure. But the VPs that the adverbs modify are absent here. Our non-derivational analysis thus gives us a simple and explicit explanation for these VP deletion facts.
3.2 Scope

In section 1, we have observed that the negation not displays adverbial properties concerning scope relations. But there are exceptions to the generalization that the linear order of two adverbs determine their relative scope: a negation that immediately follows the finite auxiliary may take scope over the latter (cf. Gazdar et al. 1982).

Consider the example in (22) first.

(22) Kim may not drink the wine on the table.

Example (22) can have two different relative scope readings between the modal and the negation, as illustrated in (23).

(23) a. Kim is not permitted to drink the wine on the table.
    b. Kim is permitted to not drink the wine on the table.

Recall that not can be either a VP modifier or a complement of the finite auxiliary via the proposed lexical rule. This then allows us to generate two possible structures for the sentence (22), as represented in (24).

(24) a. Kim $VP[may\ VP[not\ VP[base][drink\ the\ wine\ on\ the\ table]]]$.
    b. Kim $VP[v[may]\ Adv[not]\ VP[base][drink\ the\ wine\ on\ the\ table]]$.

In the structure (24)a, not is a base VP modifier governed by the modal. Under the structural determination of scope, then, the reading where may scopes over the negation is naturally expected. In (24)b, the negation is the complement of the finite auxiliary via our proposed lexical rule. Since the incorporated negation is syntactically a complement but is semantically still the ‘functor’ ([not'(Modal'(VP'))]) according to the proposed lexical rule, not takes scope over the auxiliary may in this case.

But there are well-known lexical exceptions concerning the scope of negation. As noted by Horn (1972), not all finite auxiliaries exhibit scope ambiguity with a following negation. Especially the epistemic verbs such as may and must, and semi-auxiliary verbs like need do not induce this kind of scope ambiguity with the negation immediately following.

(25) a. Kim must not drink the wine on the table.
    b. Kim need not drink the wine on the table.

The not in (25) can only have narrow scope with respect to the modal verb must. But the not in (25)b has only wide scope.

We might try to simply bar the epistemic may and must from undergoing the Adverb Conversion Lexical Rule. But this treatment would create a problem in accounting for VP deletion facts:

(26) a. Lee must go back to his country, but Kim must not __.
    b. Lee needs to go back to his country, but Kim need not __.
Since VP deletion after the negation not is possible if and only if it is the complement of the finite auxiliary in our system, we cannot claim that the not in (25) is a modifier. The solution to this problem lies in rather positing a lexical restriction: we can lexically specify that when the epistemic verbs such as may and must take the negation as a complement, they take wider scope over the negation. But an auxiliary like need is lexically specified so that it takes narrower scope than the following negation. Given the general assumption that lexical rules can have exceptions, these scope facts provide further support for our analysis.

3.3 More on the Justification of Not as a Complement

Though there seems to be no direct and obvious evidence for the complement status of not, several pieces of indirect evidence clearly support this proposal. There are cross-linguistic phenomena where adverbs behave like complements in certain contexts such as adverb incorporation in Chukchee, Modern Greek, and Nahuatl, syntactic case marking on adverbs in Finnish and Korean, and adverbial agreement in Italian. Though space limitations prevent us from elaborating these facts here (see Kim and Sag (1995) and the references therein), we could argue that phenomena like these, exhibiting certain parallels between complements and adverbs, motivate analyses in terms of a conversion rule similar to the one I adopt here.

Even in English, we can find some cases where adverbs act like complements. One obvious similarity can be found in subcategorization facts. Though adverbs are not usually selected by the verb, there are certain verbs which subcategorize for an adverb, as noted by Jackendoff (1972), McConnell-Ginet (1982), and others.

(27) a. Tom behaved *(rudely) to Marcia.
b. The job paid us *(handsomely).
c. John worded the letter *(carefully).
d. The management has treated John *(contemptuously).

The presence of the adverbs in (27) is obligatory. The omission of the adverbs here renders the sentences in (27) unacceptable.

Such parallels between complements and adverbials in terms of subcategorization again make it reasonable to allow certain adverbs to function as complements. We have in particular seen that the English particle not has dual properties: adverbial properties and non-adverbial properties. In capturing the non-adverbial properties of the negation, we have allowed the negation immediately after a finite auxiliary to become a complement of the auxiliary via the lexical rule. This lexical rule then predicts certain differences between the postfinite auxiliary negation not and the negation in other positions. There are more cases showing differences between the two.

VP fronting, a phenomenon similar but not identical to VP deletion, shows another peculiar property of the negator not (cf. Ernst 1992).
(28) Mary said she would not be eating broccoli, and
   a. [be eating broccoli] she will not __.
   b. [not be eating broccoli] she will __.

(29) Mary said she would be not eating broccoli, and
   a. *[eating broccoli] she will be not __.
   b. [not eating broccoli] she will be __.

As can be seen from the contrast (28) and (29), only the negation immediately following the finite auxiliary can be stranded as in VP deletion.

Examples of tag questions given in (30) also exhibit another difference between the negation immediately following the finite auxiliary and the one not following it.

(30) a. He has not spoken to her for days, \{\begin{align*}
   &\text{has he?} \\
   &*\text{hasn't he?}
\end{align*}\}
   b. He has often not spoken to her for days, \{\begin{align*}
   &*\text{has he?} \\
   &\text{hasn't he?}
\end{align*}\}

The choice of tag types is dependent upon the structure of the main clause. The contrast given in (30) shows that only the negative immediately after the finite auxiliary can affect the choice of the tag types.

One may argue that these differences just reflect the distinction between sentential negation and constituent negation, as has been traditionally assumed. Why can't we just follow this dichotomy? Why can't we accept the view that negation immediately after the finite auxiliary is sentential negation and the one in other positions is constituent negation? Leaving aside the question of how this dichotomy can deal with the nonadverbial as well as adverbial properties of not, the question arises as to what determines that the not following the finite auxiliary is a sentential adverb.

There are convincing cases showing that the not in the post-auxiliary position is different from sentential adverbs in several respects. For example, the emphatic usage of sentential adverbs provides us one clear difference. Sentential adverbs like always and never can be repeated for emphasis, as illustrated in (31).

(31) a. Mary always always always goes home.
   b. Mary never never never goes home.

It seems to be possible to repeat the negator not also, as shown in (32).

(32) a. Mary could not not not go home.
   b. Kim may not not not go home.

But the repetition of not is different from that of sentential adverbs. Each negator crucially contributes to composing the meaning of the whole sentence. Thus, (32)a and b mean, respectively, that it is not possible for Mary not to go home and Kim is not permitted not to go home. This indicates that unlike sentential adverbs, the negator not can not used as emphatic.¹¹
Various properties of the negation *not* immediately after the finite auxiliary distinguish it either from adverbial modifiers or from sentential adverbs. And direct and indirect empirical facts we have seen provide convincing arguments for allowing the negation to serve as a complement in a restricted context, i.e., when following a finite auxiliary verb.

4 Conclusion

We cannot deny that grammar is to some extent an indeterminate system. Categories and structures, for example, often do not have neat boundaries. We are tempted to overlook such uncertainties, or to pretend that they do not exist. In particular, the English particle *not* displays dual properties: adverbial properties and non-adverbial properties. I have proposed that one plausible way to capture these dual properties is to allow the negator *not* to be converted to a complement in a particular lexically restricted environment, namely following finite auxiliary verbs. This ‘conversion’ lexical rule mechanism has been well justified by phenomena such as VP deletion, VP fronting, and Scope.

It is true that a derivational grammar whose chief explanatory resources are functional projections including NegP and syntactic movement might be able to account for the phenomena I have dealt with here. But, in this paper, I have exploited a non-derivational and surface-oriented framework of Head-driven Phrase Structure Grammar, whose foundations lie in a concrete conception of constituent structures, a limited set of universal principles, and enriched lexical representations. The proposed analysis has shown how the interaction of a concrete $X'$-theory and the strict lexicalism that HPSG employs can draw effects similar to those of head movement and functional categories, and further permit a simple and explicit explanation for negation (and several related phenomena) that provides many descriptive and explanatory advantages.

Notes:

*I am grateful to Ivan Sag, Peter Sells, and Tom Wasow for their valuable comments on an earlier version of this paper. I especially thank Ivan Sag for his enormous contributions to the paper during various stages of its growth. I have also benefited through suggestions and help from Elizabeth Bratt, Tony Davis, and Rob Malouf. All errors and misinterpretations are, of course, mine.  

1 The negator *not* can modify other categories such as AdjP, PP, and AdvP, as in (i).

1. a. This is a not unattractive doll in some ways.
   b. I like beer, but not in the morning.
   c. I visit them not very often.

For such cases, we can loosen up the restriction of what *not* can modify.

2 I assume that English VPs are partitioned according to two VFORM values, *finite* and *nonfin(ite)*. The sort *nonfin* has *base*, *infinite*, *ger(undive)*, and *part(ictle)* as its subsorts.

3 One possible solution seems to posit an additional functional projection such as PolP (Polarity Phrase) or $\Sigma$ Phrase (cf. Laka 1990) and generate a null element as its head for each declarative sentence.

4 Examples like (i) are not VP deletion but null anaphora complement cases.
i. a. The children began singing songs, and the adults began __ too.
   b. Tom continued being noisy, although Terry stopped __.

Hankamer and Sag (1976) classify anaphoric processes into two groups: deep and surface anaphora. The difference is that the former permits nonlinguistic antecedents, while the latter allows only linguistically expressed antecedents. Given this distinction, VP deletion is an instance of surface anaphora, whereas null complement such as in (i) is an instance of deep anaphora.

Sag and Fodor (1994) reexamine empirical motivations for phonetically empty categories which have been important theoretical foundations in modern GB analyses. They show that all independent arguments for the existence of traces such as auxiliary contraction, wanna contraction, and position of floated quantifiers are neither satisfactory nor well-grounded. They also present positive arguments for terminating filler-gap dependencies by lexical heads, not by traces. See Sag and Fodor (1994) for details.

For concrete examples of type-shifting (or raising) rules in categorial grammar, see Partee and Rooth (1983), Dowty (1988), among others. This idea of converting adverbs into complements has been independently proposed for various phenomena by Miller (1991), Rida et al. (1994). Especially see Warner (1993), Abeillé and Godard (1994), and Kim and Sag (1995) for using the same technique for English and French negation.

Notice here that this lexical rule mechanism, though adopting its basic idea from the type-shifting method in categorial grammar, is basically different from it. While a categorial grammar allows type-shifting as a general principle, our system permits it only in limited cases: the system lexically controls its application. Thus in a strict sense, the lexical rule does not type-shift a modifier to a complement, but allows the ‘conversion’ of a modifier into a complement in the given environment.

I assume that Advf restricts adverbial complements to only a small subset of adverbs like not and so in English.

One may ask whether it is acceptable not to satisfy the MOD feature of the adverb not in such a case. But note here that the structure (32) is not an adjunct structure, but a head-complement structure because the negation is now converted to a complement. The HPSG theory says nothing about what happens when a complement has a MOD value. Thus its presence in a complement does not affect the well-formedness of the given phrase.

But notice that the repetition of not is not always possible:

i. a. *John wants to not not go.
   b. *John could not not go.

Whether one attributes no immediate recursive of the modifier not as in (ia) to a syntactic, semantic or pragmatic source, the question remains of why then the negator can be repeated in cases like (32). The analysis proposed here can provide a simple answer to this. The two not's in (32) are different: the first one is the complement not and the second one is the modifier.

References


0. **Introduction**

The theory of phonological representations has standardly been guided by the Jakobsonian view that predictable properties are excluded from the phonological representation (see Anderson 1985, ch. 5), and in particular, that phonetic properties which are not contrastive in any language are excluded from the inventory of phonological features, so as not to predict unattested contrasts. McCarthy (1994), for example, states, "An adequate theory of phonological distinctive features must ... be able to describe all and only the distinctions made by the sound systems of any of the world's languages." I will argue against the Jakobsonian treatment of contrastiveness, showing (a) that the contrastive or predictable status of features in a sound system falls out from the interaction of certain classes of constraints, rendering the representational restrictiveness of the Jakobsonian approach superfluous; and (b) following Ohala (1990, 1983) and Steriade (1994b), that the phonology must refer even to universally predictable phonetic properties, taking as an illustrative case the duration of voiced and voiceless stops and its role in spirantization.1

1. **Contrastiveness from Constraint Ranking**

1.1. **Contrastiveness.** Assume the definition of contrastiveness in (1):

(1) **Dfn. contrastive:** (a) A feature F is contrastive in context C iff for all $\alpha \in \{+,-\}$ an underlying $\alpha F$ specification is always realized in the output as $\alpha F$ in C.
(b) F is contrastive (tout court) iff there is some C such that F is contrastive in C.

Any underlying featural distinction which does not meet this definition is, intuitively speaking, unlearnable, and therefore cannot be used to signal distinctions in meaning (the traditional test of contrastiveness). "Feature" is used here in the broadest possible sense: any property of the phonological representation, including prosodic properties. And though the following discussion is couched in terms of binary features, it extends trivially to privative features, by substituting $<F, \emptyset>$ for $<+F, -F>$. To simplify the exposition, I will assume full underlying specification.

1.2. **PARSE$_F$.** I assume an Optimality Theoretic approach (Prince and Smolensky 1993) in which faithfulness is formalized (at least in part) in terms of a set of featural PARSE$_F$ constraints (Kirchner 1993, Cole and Kisseberth 1994, and Jun 1995).

(2) **PARSE$_F$:** Preserve the underlying value of F in the output.
Thus, a PARSEF violation is incurred if an underlyingly +F specification is changed to -F on the surface, or vice-versa. As we will see, the existence or non-existence of a corresponding PARSEF constraint has interesting consequences for the status of F in sound systems.

1.3. The Proposition. The contrastive or predictable status of features within a sound system is determined by the ranking of PARSEF constraints with respect to other constraints which restrict the distribution of the features. More rigorously,

\[(3) \quad \textbf{The Contrastiveness Theorem} \]

\[\text{For all features } F, F \text{ is contrastive iff} \]

\[(1) \quad \text{there is a constraint PARSEF and} \]

\[(2) \quad \text{for all constraints } K \text{ which restrict the values of } F \text{ in some context} \]

\[(a) \quad \text{PARSEF } \gg K \text{ or} \]

\[(b) \quad \text{there is some feature } F' \text{ s.t. } K \text{ refers to } F' \text{ and} \]

\[(i) \quad \text{PARSEF } \gg \text{ PARSEF'} \text{ or} \]

\[(ii) \quad \text{there is no constraint PARSEF'}. \]

To prove (3), it is sufficient to show that (A) if the conditions in (3) hold, F is contrastive, and (B), if the conditions in (3) do not hold, F is not contrastive.

A. Case 1: Assume that conditions 1 and 2a are true w.r.t. F. To indicate a distributional constraint which prohibits the occurrence of some value of F in combination with certain values of some number of other features, I use *[αF,βF',...]*. (This notation is standardly used for segment-internal ("feature cooccurrence") constraints, but clearly it does not matter for our purposes whether the relation among the features which the constraint prohibits is segment-internal or not.)

<table>
<thead>
<tr>
<th>Input: [αF,βF',...]</th>
<th>PARSEF</th>
<th><em>[αF,βF',...]</em></th>
<th>PARSEF'</th>
</tr>
</thead>
<tbody>
<tr>
<td>[αF,βF',...]</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>[αF,βF',...]</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>[-αF,βF',...]</td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

(Thick vertical lines indicate crucial ranking.) As shown in the tableau above, PARSEF is satisfied, either at the expense of *[αF,βF',...]* or PARSEF', depending on their relative ranking. Since, under this ranking, underlying αF is always realized as αF on the surface in this context, by definition (1) F is contrastive.

Case 2: Assume that conditions 1 and 2b are true w.r.t. F. PARSEF is then satisfied, regardless of the ranking of *[αF,βF',...]* w.r.t. PARSEF or PARSEF'.

<table>
<thead>
<tr>
<th>Input: [αF,βF',...]</th>
<th><em>[αF,βF',...]</em></th>
<th>PARSEF</th>
<th>PARSEF'</th>
</tr>
</thead>
<tbody>
<tr>
<td>[αF,βF',...]</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[αF,βF',...]</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[-αF,βF',...]</td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
(We have already seen in Case 1 that if PARSEF is ranked above *[αF, βF', ...], F is contrastive.) A fortiori, PARSEF is satisfied if there is no constraint PARSEF'. Since underlying αF is always realized as αF on the surface, by definition (1) F is contrastive.

Consequently, the conditions in (3) are sufficient to show that a feature is contrastive.

B. **Case 1:** Assume condition 1 is false w.r.t. F. If there is no constraint on the distribution of F, F occurs in free variation.

(6)

<table>
<thead>
<tr>
<th>Input: [αF,...]</th>
<th>(no relevant constraints)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[αF,...]</td>
<td></td>
</tr>
<tr>
<td>[-αF,...]</td>
<td></td>
</tr>
</tbody>
</table>

If there is a constraint on the distribution of F, *[αF, βF', ...], but one or more of the features referred to in the constraint lack a corresponding PARSE constraint, F again occurs in free variation.

(7)

<table>
<thead>
<tr>
<th>Input: [αF, βF', ...]</th>
<th>*[αF, βF', ...]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[αF, βF', ...]</td>
<td>*</td>
</tr>
<tr>
<td>[αF, -βF', ...]</td>
<td></td>
</tr>
<tr>
<td>[-αF, βF' , ...]</td>
<td></td>
</tr>
</tbody>
</table>

Since free variation means that underlying αF is not always realized as αF on the surface, by definition (1), F is not contrastive.

If, however, there is a constraint *[αF, β1F1,...βnFn], and all features F1 through Fn do have corresponding PARSE constraints, then F is predictable in the context β1F1,...βnFn.

(8)

<table>
<thead>
<tr>
<th>Input: [αF, β1F1,...βnFn]</th>
<th>PARSEF1</th>
<th>...</th>
<th>PARSEFn</th>
<th>*[αF, β1F1,...βnFn]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[-αF, β1F1,...βnFn]</td>
<td></td>
<td></td>
<td></td>
<td>!</td>
</tr>
<tr>
<td>[αF, β1F1,...βnFn]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[αF, β1F1,...βnFn]</td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[αF, -β1F1,...βnFn]</td>
<td></td>
<td></td>
<td>!</td>
<td></td>
</tr>
</tbody>
</table>

That is, both αF and -αF are realized as -αF on the surface, failing to meet definition (1). As shown above, in other contexts, where there is no such relevant distributional constraint, F occurs in free variation. Therefore, if there is no PARSEF constraint, F is either freely varying or predictable, but never contrastive.

**Case 2:** Assume condition 2 is false w.r.t. F. If PARSEF' and *[αF, βF', ...] are both ranked above PARSEF, F is predictable in this context.
If either PARSEF or *[αF,βF,…] is unranked w.r.t. PARSEF, then F occurs in free variation.

Consequently, for any context in which there is a relevant constraint *[αF,βF,…] which dominates or is unranked w.r.t. PARSEF, and PARSEF dominates or is unranked w.r.t. PARSEF, F is either freely varying or predictable, but never contrastive.

Therefore, the conditions in (3) are both necessary and sufficient to show that a feature is contrastive, Q.E.D.²

1.4. **Universally noncontrastive features.** Recall that in the Jakobsonian treatment of contrastiveness, phonetic properties which are never contrastive in any language are excluded from the phonological representation. However, I have shown in Part B (Case 1) of the previous section that it suffices to assume that such properties lack a corresponding PARSEF constraint; regardless of ranking, their surface realization will be either predictable or freely varying. Consequently, we may include any and all phonetic properties in the phonological representation, without thereby expanding the range of contrasts available to UG.

1.5. **Gradiency of representations.** Further consider the familiar phonological strategy of decomposing a continuous phonetic dimension (e.g. vowel height) into a set of binary features.

(11) Vowel height:

```
<---------------------------------------------------- high l +high ------------------>
<----------------- + low l -low --------------------->
[--------------------------------------------------]
    lower                                      higher
```

Note that if each "step" along the scale need not be contrastive per se, it is possible to subdivide a phonetic continuum into any number of features, each of
which corresponds to some range (in principle, even approaching infinitesimaly) within that continuum.

(12) Phonetic dimension X:

\[
\begin{array}{c}
\text{<-----------------+B | -B -------------------------------->}
\text{<---- +A | -A -------------------------------->}
\text{[-----------------+n | -n -------------------->}
\text{not at all X}
\end{array}
\]

completely X

Thus, in (12), the X dimension is carved up into n binary features. The implicational relations among the features (e.g. if -B then -A) follow from their definition as ranges within a particular dimension. “Categorical” effects can be obtained, notwithstanding the gradient representation of the dimension, by means of feature cooccurrence constraints. For example, in (12), if there is an undominated constraint *[-A,+n], this would rule out segments with some degree of X-ness which lies between the +A and -n cutoffs.

(13)

<table>
<thead>
<tr>
<th>Input: [-A,...,+n]</th>
<th>*[-A,+n]</th>
<th>PARSE_A</th>
<th>PARSE_n</th>
<th>(Underlyingly *[-A,+n] segment realized either as +A or -n depending on relative ranking of PARSE_A and PARSE_n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[-A,...,+n]</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[-A,...,-n]</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>[+A,...,+n]</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

Therefore, gradient phonetic distinctions may be represented in the phonology without expanding the range of contrasts available to UG. In sum, one can envision a phonological representation which, in its detail and gradience, could be equivalent to what has standardly been called a phonetic representation, generally presumed to be the output of a distinct phonetic component of the grammar.3 I will not argue here that there is no phonetic component distinct from the phonological component. Nevertheless, I have shown that one of the central arguments for positing such a component -- that phonological representations cannot include gradient distinctions and other non-contrastive phonetic detail -- is without force.

2. Motivation for Predictable Phonetic Properties in the Phonological Representation

Although I have demonstrated that Jakobsonian representational restrictiveness is superfluous to an adequate account of phonological contrastiveness, it could still be the case that, as an empirical matter, universally predictable phonetic properties play no role in conditioning phonological phenomena, therefore the phonological representation need not refer to such properties. However, Ohala (1983) and Steriade (1994b) have presented evidence against this claim, namely that the predictable aerodynamic properties of voicing play a large role in explaining the distribution of voiced segments. Similarly, Browman and Goldstein (1992) have argued that subphonemic distinctions in degree of overlap among articulatory gestures can explain a variety of assimilation phenomena. The remainder of this paper concerns itself with a
further case of a universally predictable phonetic property which plays a role in conditioning phonological phenomena.

2.1. The problem. The relation between voicing and lenition is a long-standing problem of phonological theory (Foley 1977, Lass and Anderson 1975, Harris 1990, Bauer 1988). For example, in most dialects of Spanish (Harris 1969, Lozano 1979, Castillo and Bond 1987), voiced stops spirantize in certain environments (14a, 15a), while voiceless ones never do (14b, 15b).

(14) a. paða 'turkey hen' laðo 'side' toya 'toga'
    b. papa 'potato' lato 'I throb' toka 's/he plays'

(15) a. barkos 'ships'
    dewwas 'debts'
    ganaðo 'cattle'
    aj barkos 'there are ships'
    aj dewwas 'there are debts'
    aj yanaðo 'there is cattle'
    b. palmas 'palm trees'
    toros 'bulls'
    kokos 'coconuts'

Similarly, in Tümpisa Shoshone (Dayley 1989), spirantization (in certain environments) is obligatory in (singleton) voiced stops (16a), but optional in voiceless ones (16b).

(16) a. taʃefjį 'sun'
    tsiðoːhį 'push'
    tuˈyʷanːi 'night'
    b. taha(ʃ/pj) 'snow'
    huʃiari(ʃ/kj) 'sing'

In fact, I am aware of no languages in which spirantization which applies to voiceless stops to the exclusion of voiced stops. Nevertheless, despite the well-known and widely attested relation between voicing and lenition, no previous phonological framework has done more than stipulate, by some means or other, that voiced stops are "weaker" than voiceless ones, therefore in some sense closer to continuants.

2.2 Stop duration. The problem can be solved once we take into consideration certain predictable phonetic properties: voiced stops are phonetically shorter than voiceless ones (Lehiste 1975). In Breton, for example, average closure durations for intervocalic voiced stops (averaging across place of articulation) is 49.9 msec, whereas for voiceless stops it is 102.3 msec. Similarly, Homma (1981) reports that in Japanese, voiced stops have an average closure duration of 44 msec, whereas for voiceless stops it is 67 msec. As discussed in the previous section, we can carve up the duration continuum into any number of binary features; but for our purposes a single cutoff point suffices, which we can refer to as [+longer duration].

(17) [longer duration] ([ld]): a segment is [+ld] if its duration is greater than k msec. A segment is [-ld] if its duration is less than or equal to k msec. (For the sake of concreteness let k = 60).
It is not crucial to this analysis why voiced stops are [-ld], though Ohala (1976, 1983) has suggested some plausible aerodynamic bases for this pattern. I will simply stipulate a feature cooccurrence constraint, *[αvoi, ald, -cont]. Since [ld] is universally non-contrastive, there is no PARSE[ld] constraint, therefore this feature is universally predictable in stops from the specification of [voi], regardless of ranking.

(18) | Input: [+voi,+ld] | PARSEvoi | *[αvoi,ald,-cont] |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[+voi,+ld]</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>[+voi,-ld]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[-voi,+ld]</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>[-voi,-ld]</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

(19) | Input: [-voi,-ld] | PARSEvoi | *[αvoi,ald,-cont] |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[-voi,-ld]</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>[-voi,+ld]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[+voi,+ld]</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

2.3. Spirantization as undershoot. Following Zipf (1949), Lindblom (1984), and others, I assume that articulation (and all other human physical activity) is governed by a basic imperative of effort minimization, which I formalize as the following Optimality Theoretic scalar constraint.

(19) LAZY: Minimize articulatory effort

Regardless of precisely how articulatory effort is determined (cf. Westbury and Keating 1986, Lindblom 1990), it seems uncontroversial that for a given closure gesture, the more the duration of the gesture is reduced, the more effort is required to achieve it (at least, provided that the closure is not so long that special effort is required to maintain it, as is perhaps the case in geminates). By the same token, if effort is held constant, the more reduced the duration, the less the magnitude (constriction degree) of the gesture. The tendency of voiced stops to spirantize can now be explained: in a [-ld] segment, in the interest of conserving effort, complete closure may be sacrificed, yielding a continuant. In other words, spirantization can be naturally viewed as a case of articulatory "undershoot" (Lindblom 1963).

2.4. Spanish. To formalize this in OT terms, let X equal the amount of effort required to achieve complete closure in a [-ld] segment. Like all scalar constraints in OT (see Prince and Smolensky, ch. 5), LAZY may decomposed into a set of binary constraints, whose ranking w.r.t. each other is determined by Pāṇini's Theorem (i.e. the Elsewhere Condition).

(20) LAZYX: Do not exert effort ≥ X.

(21) ... » LAZYX+1 » LAZYX » LAZYX-1 » ...
The Spanish spirantization facts can now be accounted for in terms of the following constraint ranking:

(22) \( \{ \text{PARSE}_{\text{voi}}, *[\alpha_{\text{voi,alld,-cont}}], \text{LAZY}_X \} \gg \text{PARSE}_{\text{cont}} \)

Tableaux (23) and (24) demonstrate that under this ranking, voiced stops spirantize, whereas voiceless stops do not.

<table>
<thead>
<tr>
<th>Input: [+voi,-ld,-cont]</th>
<th>PARSE_{voi}</th>
<th>*[αvoi,alld,-cont]</th>
<th>LAZY_X</th>
<th>PARSE_{cont}</th>
</tr>
</thead>
<tbody>
<tr>
<td>[+voi,-ld,-cont]</td>
<td></td>
<td>![image]</td>
<td></td>
<td>*[image]</td>
</tr>
<tr>
<td>[+voi,+ld,-cont]</td>
<td></td>
<td>![image]</td>
<td></td>
<td>![image]</td>
</tr>
<tr>
<td>[-voi,-ld,-cont]</td>
<td></td>
<td>![image]</td>
<td></td>
<td>![image]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Input: [-voi,+ld,-cont]</th>
<th>PARSE_{voi}</th>
<th>*[αvoi,alld,-cont]</th>
<th>LAZY_X</th>
<th>PARSE_{cont}</th>
</tr>
</thead>
<tbody>
<tr>
<td>[-voi,+ld,-cont]</td>
<td></td>
<td>![image]</td>
<td></td>
<td>*[image]</td>
</tr>
<tr>
<td>[-voi,+ld,+cont]</td>
<td></td>
<td>![image]</td>
<td></td>
<td>![image]</td>
</tr>
<tr>
<td>[-voi,-ld,+cont]</td>
<td></td>
<td>![image]</td>
<td></td>
<td>![image]</td>
</tr>
</tbody>
</table>

More generally, in this sort of framework, lenition is analyzed as some degree of LAZY dominating some PARSE_manner feature constraints (see Kirchner 1994).6

The environments for spirantization can be obtained by blocking spirantization (or even requiring fortition) in particular environments, by means of higher-ranked constraints, which are not directly relevant here.

2.5. Tümpisa Shoshone. This analysis can readily be extended to account for the optional spirantization of voiceless stops in Tümpisa Shoshone, while still capturing the relation between voicing, closure duration, and spirantization. We simply need to identify the amount of effort required to achieve complete closure in a voiceless stop: call this \( Y \). The optionality of spirantization can be captured by leaving LAZY\( Y \) and PARSE_{cont} unranked w.r.t. each other.7

<table>
<thead>
<tr>
<th>Input: [-voi,+ld,-cont]</th>
<th>*[αvoi,alld,-cont]</th>
<th>LAZY_X</th>
<th>PARSE_{cont}</th>
<th>LAZY( Y )</th>
</tr>
</thead>
<tbody>
<tr>
<td>[-voi,+ld,-cont]</td>
<td>![image]</td>
<td></td>
<td>*[image]</td>
<td></td>
</tr>
<tr>
<td>[-voi,+ld,+cont]</td>
<td>![image]</td>
<td></td>
<td>![image]</td>
<td>![image]</td>
</tr>
<tr>
<td>[-voi,-ld,+cont]</td>
<td>![image]</td>
<td>![image]</td>
<td>![image]</td>
<td>![image]</td>
</tr>
</tbody>
</table>

Crucially, since the duration of the voiceless stops is longer than the voiced stops, \( X \) is less than \( Y \); so by Pāṇinī’s Theorem LAZY\( X \) is universally ranked above LAZY\( Y \). Consequently, it is impossible to have a system in which the longer (voiceless) stops spirantize, while the shorter (voiced) ones do not.
3. **Conclusion**

I have shown that, contrary to the Jakobsonian view, an adequate treatment of contrastiveness does not require the exclusion of universally predictable features from the phonological representation. Rather, the predictable or contrastive status of features falls out from the ranking of PARSEF constraints w.r.t. constraints which restrict the distribution of these features; and universally predictable features simply lack a corresponding PARSEF constraint. Consequently, phonological representations may contain an unlimited amount of phonetic detail, including gradient distinctions, without thereby increasing the range of contrasts available to UG. Furthermore, I have presented an example of a phonological phenomenon, the relation between stop voicing and spirantization, which is conditioned by a universally predictable phonetic feature, namely the durational distinction between voiced and voiceless stops. Therefore, enrichment of phonological representation to include some predictable phonetic features is not only feasible: it is empirically necessary. The question of which predictable phonetic features, beyond stop duration, are relevant to phonological phenomena becomes a wide-open field of empirical inquiry, now that the blinders imposed by the Jakobsonian treatment of contrastiveness have been removed.

---

1. My approach has adverse implications for underspecification theory as well, since underspecification theory is essentially a language-particular implementation of the Jakobsonian treatment of contrastiveness with respect to early stages of phonological derivation. Nevertheless, because the inadequacies of underspecification theory, in light of OT (Smolensky 1993, Inkelas 1994, Steriade 1994b) or otherwise (Mohanji 1991, Steriade 1994a), have already received attention, I will not explicitly address them here.

2. In a multi-stratal grammar (if such exist), F will be contrastive just in case (3) characterizes the constraint ranking w.r.t. F at each stratum. We have shown that, on the first round of evaluation, underlying $\alpha F$ maps to output $\alpha F$ just in case (3) holds w.r.t. F. The output, $\alpha F$, is then taken as the input for the next round of evaluation. But if (3) characterizes the next stratum as well, the same result obtains, and so on, regardless of the number of levels of computation.

3. Interestingly, this view is consistent with recent research on speech perception, e.g. Pisoni 1992, which suggests that speakers retain in long-term memory all sorts of non-contrastive perceptual information associated with particular tokens of lexical items, including voice characteristics and speaking rate.

4. The feature [ld] is obviously reminiscent of the notion that voiceless stops are specified [+tense] or "fortis"; although [ld] refers to pure duration, whereas [tense] ostensibly refers to the tension of the vocal tract, and the terms fortis and lenis have never had consistent phonetic definitions. However, it matters little whether the [ld] feature is viewed as an original proposal or a revival of an old idea. By Jakobsonian reasoning, since neither [ld] nor [tense] is contrastive in consonants, neither feature should be included in the representation of consonants. Consequently, phonological motivation for either feature constitutes a refutation of the Jakobsonian position.

5. Briefly, voiced stops must be short so as to avoid passive devoicing (cessation of Bernoulli vibration of vocal cords due to build-up of oral air pressure during a stop). Voiceless (unspirated) stops, on the other hand, must be long so as to be simultaneous with the glottal abduction (devoicing) gesture, which has a relatively fixed duration, varying somewhat from speaker to speaker, but rarely less than 60 msec (Weismer 1980) if the timing is not simultaneous, the devoicing will "spill" onto neighboring sonorants, violating the constraint *(+son,-voi]). If this is the correct explanation, we would expect the value of $k$ in the definition of [ld] to vary somewhat among speakers, due to variation in size of the oral cavity, as well as varying depending on place of articulation (the more anterior the closure, the longer the voicing can last).

The problem of optional rules is a non-trivial one in constraint-based formalisms. The device of indeterminate ranking seems too powerful, in that it fails to characterize just the sorts of variation typically encountered within a given idiolect. Lindblom (1990) has observed that intra-speaker variation typically involves variation along a hypoarticulation - hyperarticulation continuum, where hypoarticulation maximizes ease of articulation, and hyperarticulation maximizes preservation of acoustic cues. In Kirchner (1994), this notion is modeled within OT by assuming that the input to phonological computation contains not only the underlying representation, but also some information about the current extralinguistic state of the system, including tiredness, preoccupation, etc. This information might take the form of a numerical index, which augments or diminishes by some constant function the "effort" cost associated with each articularatory gesture. Variation in the value of this index would, in effect, correspond to adjustment of feedback gain in Lindblom's H&H model. In the present case, it suffices to assume that under hypoarticulation conditions, the "effort" index boosts the cost of a voiceless stop gesture to X (and the cost of a voiced stop to something greater than X).

REFERENCES

Kirchner, R. (1994) Lenition in Phonetically Based Optimality Theory, Dissertation Prospectus, UCLA.


(1994b) *Voicing, Underspecification, and the Small Feature Set*, ms., UCLA.


Germany vs. "The South" or:
Should (and can?) second language learners be taught
how to be "rude" or "sweet"?

Elisabeth D. Kuhn

Virginia Commonwealth University

When I moved from Germany to Berkeley about a dozen years ago, my command
of English was quite adequate, I thought, but nobody had prepared me for the
pragmatic differences between the ways Germans and Californians used language. I
was puzzled by a number of things -- for example, by people constantly asking me
how I was and then not stopping to hear the answer -- all of it. If they did stop, I
would notice their eyes glaze over after a minute or two. I quickly learned that the
expected answer was 'fine'. They did not realy want to know the truth.

For a few weeks, I was also amazed at people's friendliness. 'My, am I popular!' I
thought. However, I quickly learned that smiles and "friendliness" in California
did not mean what I had grown used to expect them to mean. In Germany, they
signal genuine affection. In California, they often merely signal 'I don't bite!'  

Knowing what I know now, I wish somebody had told me these things and a
variety of others, even practiced them with me. A class on applied cross-cultural
communication would have been very helpful. I am grateful that I had at least a
chance to practice -- in California -- before I moved to Virginia, because it turned
out that Virginia style politeness contains just about every element of California
style politeness I found difficult to learn, only in a much more elaborate version. I
would have been an unmitigated social disaster had I tried to survive in Virginia
straight off the plane from Frankfurt.

In this paper, I want to discuss some major differences in communicative
conventions between Germans and Virginians -- from what I have been told, they
hold for much of the South -- and make a case for teaching those conventions
explicitly, for practicing them, and for giving people who move from one to the
other a chance to become familiar and more comfortable with those conventions
before, or soon after, their arrival. This could prevent a lot of hurt feelings and
insecurity.

A number of scholars (Byrnes (1986), Lakoff (1975), Kuhn (1981, 1992),
Kotthoff (1988), House and Kasper (1981), House (1989), and Hall (1984), have
discussed the fact that in the U.S., Germans have a reputation for being rude or
abrupt. Part of this reputation can no doubt be traced to German military personnel
barking Achtung 'attention' in old war movies. Some of it stems from German
exchange students bristling at questions from friendly Americans trying to make
small talk, responding with 'why do you want to know' or some version of 'none
of your business'. Some of it stems from Germans who, when asked how they
liked the U.S., happily expounded on the pros and cons, especially the cons, since
the controversial aspects made for much more interesting discussion.

Another important source of that reputation lies in differing conventions for the use
of speech acts and their modifications. In German, directives and other utterances
are frequently modified by particles such as **mal**, **doch**, **ja**, and **halt**, a technique that allows the speaker to use a relatively direct speech act without projecting too rude an illocutionary force.

_Gib mir das Salz_ 'give me the salt' is a command.

_Gib mir mal das Salz_ 'give me _mal_ the salt' is a request.

**Mal**, which does not translate well into English, plays down the cost of any request, making it less imposing. Thus, a more direct speech act can be used. After all, you ask someone more directly when you need a quarter than when you need $50. Unfamiliarity with the importance and proper use of those particles and how their function does and does not translate into English causes problems for Germans who speak English as well as for Americans who speak German. Germans who speak English cannot find proper translations for _mal_ and thus may simply leave it out. **Voilà!** A rude command. Americans, on the other hand, who go to Germany, will hear commands, but the _mal's_ and _halt's_ won't really register because they are easily overlooked when you have not been alerted to their presence, importance, and proper use. The result? Americans will probably perceive the Germans as rude. And they might even try to use the commands the way they hear them -- without particles -- and then sound too rude in return. Since that can be uncomfortable, though, they might simply use more indirect versions of the respective speech acts instead, which, in turn, causes them to be perceived as too unassertive or indirect.

Having said all that, however, it is important to note that, particles or not, Germans do tend to use more direct and blunt speech acts, whether in requests, apologies, or statements. Kasper and House in particular have pointed this out (1981), and so has Byrnes (1986), and also Kotthoff (1988), who gave a good contrastive example of student paper critiques at German and American universities. Below is an approximate translation of her example. The German professor would write something like the following in response to a very badly done book report:

>This paper is totally unacceptable as is. You have not grasped the key concepts of the book. It is not clear which ideas are yours, and which ones are from the book. Read the book again and see me if you have questions.

(1988:14)

The American professor would respond to the same review much more kindly:

>You have taken on a challenging task with this book review. One can see that you have put a lot of effort into dealing with the content and into doing justice to the author. There are, however, a few points that are not quite clear yet. You might also want to make a clearer distinction between...

(1988:15)

Even German students, writes Kotthoff, find the U.S. critique more palatable. However, there is the danger --and I can tell that from personal experience and from that of a number of friends -- that the German students may not get the message that the paper really is not acceptable and will be stunned to find out later from the grade
just how unacceptable it was. On the other hand, encouragement provided by positive feedback works for Germans as well.

As a rule, however, Germans tend to stay with the facts, as unpleasant as they may be, and strive to express them as clearly and explicitly as possible. When I first learned about Grice's (1975) four maxims, quality, quantity, relevance, and manner, I felt they reflected how Germans communicated quite adequately. Virginians, on the other hand, make extensive use of Grice's conversational implicatures -- you have to work at understanding what is really meant.

I will now present a brief analysis of some features of polite behavior in Virginia in terms of Lakoff's politeness rules (1975, 1990), comparing and contrasting them with their counterparts in California and Germany. It is my hope that this paper will demonstrate how polite behavior in Virginia "works" and how some coaching in its main elements will help newcomers to its culture whose styles may be quite different, to communicate more effectively by helping them to avoid alienating the people they speak with or being alienated by them. Some coaching can also help Virginians who plan to go to Germany or who work with German colleagues and bosses.

As Gumperz (1981, 1982), Lakoff (1975, 1990), and Tannen (1986), have argued, speakers tend to interpret what they hear within their own framework of reference. They assume that everyone who speaks the same language functions just like them and thus interpret whatever the other may say or do to mean what they would have meant, had they said or done it themselves. Since speakers from different speech communities (and not just from other countries -- see for example Tannen (1981 and 1986)) communicate according to different rules, this leads not only to clashes, misunderstandings, and hurt feelings, but also makes repair of such problems difficult because speakers rely on their own respective repair mechanisms to resolve the problem (Gumperz 1982; Tannen 1986). This usually makes things even worse and can lead to dislike and prejudice.

While there are many variations on communicative conventions, Lakoff (1975, 1990) argues that there are three major strategies around which they are built: distance (don't impose), deference (give options), and camaraderie (be friendly). She cites German and British culture, and European middle and upper classes in general as examples for distance politeness, in which conflict is assumed as inevitable and to be avoided by preventing "participants coming into direct contact with one another." (1990, 35). A variety of rules on what is considered acceptable behavior and talk help preserve people's space, and protect their privacy and their feelings. The second strategy, deference, she argues, is the one preferred by the Japanese, and by women in general in many other societies as well. It is characterized by the use of euphemisms and circumlocution, and by the liberal use of hedges and questions. The third strategy, camaraderie, values friendliness and personality higher than properness and non-offensiveness. Lakoff cites the Californian style with its appearance of instant intimacy as a prime example of conventional camaraderie (1990, 18). Lakoff further argues that those strategies are especially distinct in casual encounters and casual acquaintance relationships. As people get to know each other better and become friends, they incorporate type three elements while dropping the ones from the others if they are incompatible (especially in case of strategy one) or incorporate them alongside their own
(strategy two).

Virginia politeness conventions (and those of the South in general), however, seem to incorporate an interesting combination of strategies one, two, and three, including conventional camaraderie and distance, which, Lakoff argues, are really incompatible. This makes it very difficult for outsiders to always respond appropriately to the cues they get in conversations. Considering that many cultures view strategy three as the one appropriate to closer relationships, the potential for mixed signals in a society that intertwines strategy three (conventional camaraderie) with one (distance) is enormous.

In order to illustrate what happens, let me compare and contrast a few encounters in Germany, California, and Virginia. Overall, casual encounters, while generally involving some version of "how are you" and "have a nice day" in both California and Virginia, generate a more involved and detailed version of that in Virginia than they would in California. Below you will find composite examples of typical service encounters in a store in Germany, in California, and in Virginia (the more elaborate versions are more likely if there is no line, even though a line is no guarantee for brevity)(C = customer; S = salesperson):

**Germany:**

C: Guten Tag. 'good day' (maybe)
S: Guten Tag. 'good day'
       (task)
C: Auf Wiedersehen. 'good bye'
S: Auf Wiedersehen. 'good bye'

**California:**

S: Hi. How are you?
C: Fine. And you?
S: Fine.
       (task)
S: Have a nice day. (maybe -- It seems that the frequency of these has greatly decreased in the last few years. It used to be standard about 10 years ago and now they are used only some of the time (then again, maybe my sensitivities have changed after 4 years in Virginia.))
C: Thank you. You too. (maybe)

**Virginia:**

S: Hi. How are you today?
C: Pretty good. How are you?
S: Just fine. What can I do for you today?
       (task)
S: Thank you for ......
C: Thank you. (or: You're welcome)
S: Have a nice weekend/evening/....
C: Thank you. You too.
S: And come back and see us again.
C: Okay. Bye.
S: Bye.

The above, however, is the short version. Not infrequently, such service
encounters are expanded into virtual pseudo-relationships, as S follows up with a series of questions such as the following:

Where are you from?
Is your family still there?
Do you miss them?
Do you go back often?
What brought you here?
Where do you work?
What do you teach?
How many languages do you speak?
Are you married?
Have you met anyone yet?

More detailed information is elicited and given. All "free" info is turned into more conversation. At the end of the exchange there is a more elaborate 'Nice talking to you. Please come back and see us again. Good luck with (something specific) or have a good (something specific that had been mentioned in the previous conversation).'

While such questions are not uncommon in small talk situations at parties and such in California either, I had never encountered them in service encounters before I came to Virginia. Here, however, even service encounters are frequently turned into pseudo relationships. I have even taken to avoiding certain cosmetics counters -- or banks -- because their salespersons established such elaborate relationships that I feel guilty when I buy something at another counter. And sometimes I avoid them because I simply don't have the time or inclination for a lengthy visit. This has never happened in California even though sales clerks there were friendly as well.

As this example shows, the trickiest part for newcomers is small talk in Virginia -- in service encounters and with just about everybody who is not a close friend. Here, much more detailed and personal questions are asked than in California -- with follow-up question probing ever deeper. The problem? It is considered rude and inappropriate to divulge negative or problematic information. On the other hand, one also should not be so positive as to appear to be bragging.

To explain this in terms of Lakoff's politeness strategies: Virginians ask questions which would be off-limits in all but the most intimate conversations in Germany. The intention behind this is camaraderie -- the person should be made to feel like a friend. And you are supposed to answer them and sound as if you were telling the truth. However, you can't tell the truth if the answer would be negative or otherwise problematic. This is part of distance politeness -- don't impose your problems on others.

Why is this difficult? The personal questions signals to newcomers that this may be a budding friendship. If that is welcome, they will happily respond in kind, and quickly realize that what they have said is not appreciated. The other person changes the topic, plays down the significance of what has been said ('Well, I'm sure it'll be okay') which our German speakers in turn would find very offensive. If such closeness is not welcome to German speakers, they will feel imposed on and respond with 'None of your business,' which makes them look rude.
A similar problem involves the expression of opinions. In Virginia, you have to chat amiably and agreeably, without letting too explicit a contrary opinion "spoil the atmosphere." Germans do not have that rule. Honesty is valued above all. In fact, Germans consider it a sign of integrity to stand up for their opinion and to say what they think, whereas they really have only bad words for people who do not do that. Those people are considered deceiving brownnosers.

Now what about teaching newcomers these strategies? Byrnes (1986) bristles at the thought of actually teaching someone to be what she considers rude. I believe the key is an understanding of what this really means within the culture where it is practiced. Standing up for one's opinion is valued in Germany. In German eyes, a conversation would be hardly worth having if there were not some differences of opinion. If everybody agrees, then why talk? Obviously, Germans use talk much less for its relationship maintaining function than Americans do -- which may be, at least in part, because relationships are assumed to be more stable and are taken for granted to a higher degree than is the case in the U.S. Moreover, their arguments and acerbic remarks often function as solidarity signals, similar to those Schiffrin (1984) describes in her work on Jewish speakers. As Schiffrin writes: 'They seem designed to show that the interactants' relationship is close enough to withstand what would be considered by outsiders to be verbal assaults' (1984:331). As Schiffrin also points out, Tannen (1981b, 1986) has argued that Jewish speakers create rapport by mutual complaining -- another tool that Germans use in a similar way, and another one that does not work as well in Virginia, at least not in casual encounters.

The German style has also similarities to what Kochman describes in Black and White Style in Conflict as the discussion style of African Americans. Standing up for one's opinion is valued. Anything else is considered devious and problematic. This shows that the problem is not at all one-sided. Germans coming to the South will have to learn to express their opinions more cautiously, certainly in conversations with speakers of Anglo-Saxon heritage. On the other hand, the typical "polite" Southerner will likewise have problems in Germany. Being agreeable and being able to conduct lengthy conversations without voicing any differences of opinion is not a valued trait in Germany. Unless speakers voice an opinion now and then, and respond appropriately to those uttered by others, they will not function at their full potential in a German setting.

The question I posed in the title is 'Should -- and can this be taught?' I hope I have answered whether this should be taught. It will make cross-cultural encounters much more rewarding for all parties involved. But can it be taught? Byrnes (1986) argues that those ways of using language are akin to deeply engrained traits, learned at childhood, and that therefore they really can't be taught.

I agree that they are deeply engrained, and this is precisely why they cause emotional reactions such as dislike in encounters with someone who does not share the style. But I don't agree that they cannot be taught. Virginia, California, and presumably the rest of the U.S. as well, is full of people who have learned to make adjustments to their homegrown style and now function quite adequately pragmatically. That most of them have had to learn how to the hard way does not diminish the argument, but rather strengthens it.
The first step the student of either background has to take is to fully understand how the system works, and develop an understanding and appreciation for the value system on which it is built. This will not be an easy task, but understanding that arguments can be meant as a friendly solidarity reinforcement will make it much easier to learn to become comfortable with a more argumentative style than simply treating the task as 'having to learn to operate in a way one finds detestable.' And the success that comes with operating appropriately in a new culture reinforces the learning and makes it easier as one goes along.

Second, overcoming lifelong conditioning takes work. It is therefore necessary to practice the behaviors that are to be learned for the target speech community. This means that Germans need to learn to have lengthy agreeable conversations about a wide range of topics. They have to learn to respond to questions in a way that sounds sincere without divulging problematic information. This can be practiced. To prove the point: Countless Germans have learned how to do this -- or else they would not be able to operate as successfully as they do in Virginia or California. Some even embrace that style, and find it much to their liking. When I found in California that I had trouble making small talk, I took a course that taught me how to do it, and the quality of my social life immediately skyrocketed. In Virginia, I had to expand my repertory substantially, but I managed -- after a while. And so have a number of others I know. One can learn to be nice and sweet enough to get by, without losing one's self-respect.

The same is true for Southerners who learn German. Modeled after those courses on how to do small talk, one can also practice how to argue and give clever and sometimes acerbic repartees -- and how to cope with getting them. This is something that will help Southerners cope with living in Germany, or cope with their German bosses and colleagues at German companies located in this country. I even found examples for exercises that could be used: a little booklet called *Rede und Antwort* by Hueber. It presents a wide range of possible speech situations, and provides just as wide a range of answers, all of which are likely to be encountered in German settings. Unfortunately, this booklet lacks any explanatory notes -- while one of its recommended uses is self-study. I can imagine that it often meets with consternation. However, with proper explanations about the social function and significance of what is found in it, such a book can be very helpful.

Moreover, this is not a matter of replacing one's own style completely, but a matter of simply expanding ones repertory and learning to use its full range judiciously. Besides, Southerners (and Californians) have an advantage: their ability to start conversations and to keep them going can help in Germany as well, as long as the speaker comes across as sincere and not as phony. The traditional German style evolved in a very non-mobile society that had to defend itself against intruders from East and West, and thus has a strong in-group vs. out-group polarization. Now, Germany is not nearly that stable anymore, and Germans are beginning to learn how to do the things that Americans know how to do as part of their normal style. With an expansion and practice of a more direct repertory, Southern speakers can be very successful communicators in Germany. And they don't even have to be rude to do it.
Bibliography.


A dialect in the face of the standard: a Japanese case study

Yoshimi Miyake
University of Michigan

0. Introduction. The relationship between a local dialect and a national standard has been the topic of many sociolinguistic studies. Aside from the prestige which national standards invariably have, it has also often been found that non-standard dialects too can have local prestige (Labov 1963, Trudgill 1972). The present paper, a study of the Japanese city of Yonezawa, shows the conflicting feelings of its residents: (1) the prestige which they have had for hundreds of years as a regional center, and (2) the feeling of being marginalized as speaking a nationally stigmatized dialect. However, the situation is somewhat different here than what has been described in studies of local prestige in Western society such as those of Labov and Trudgill, because of the feudal origins of the local prestige group, so that even today the prestige of the local dialect extends to all social classes. In order to understand this, it will be necessary to analyze historical, sociopsychological, political, and ethnological aspects of the situation.

Yonezawa is a city of 91,000 people, located in Yamagata Prefecture, in the Tohoku District, which is on the north tip of the main island of Japan. Upon hearing someone speaking the Yonezawa dialect, Japanese from other regions would recognize it as some variety of the Tohoku dialect, which is generally stigmatized at the national level. After briefly describing the characteristics of the Tohoku dialect, I will discuss its social aspects, that is, how the society defines its own dialect and what kind of emotional value the people feel for it as opposed to Standard Japanese. The second part of the paper will discuss the specific situation of the Yonezawa Dialect and how its speakers react to opposing pressures to maintain and eliminate this dialect. Its national stigmatization comes from its being an isolated and underdeveloped area. Its local prestige comes from its association with the Uesugi Clan, which had been nationally powerful prior to the establishment of the Tokugawa Shogunate at the beginning of the 17th century. At that time the Uesugi Clan's territory was drastically decreased and they built their new castle in Yonezawa, transforming it from a village of 800 households to a city of 130,000 people in the next 100 years (Kamimura 1977). Though the importance of the clan has decreased over the years, the city still feels the prestige of being associated with one of the major clans of Japan's feudal times, and this prestige extends to their feelings about the local dialect.

The data used in this paper were collected in the summer of 1994. They consist of tape recordings made at various places, e.g. school conversations, private tutoring, conversations in the waiting room of a hospital, and interviews with people of different ages; I will also discuss the results of a questionnaire survey.

1. Tohoku Dialect. At the time of the Meiji Restoration in 1867, the dialect which was spoken by 'educated people' (Joo'o 1992) in Edo (present-day Tokyo) was designated as Standard Japanese. However, the majority of people in the western part of the country, where the historical capital Kyoto is located, opposed or at least were not happy with this decision. In the study of language attitudes, Japanese have focused on the Kansai Dialect, spoken in the Kyoto area,
which is considered to be the main rival of Standard Japanese. For this reason, the study of Tohoku Dialect, especially the aspect of its perception, has been neglected to a large extent.

1.1. General linguistic features of the Tohoku Dialect. Some dialectologists state that the Tohoku Dialect is the one which is farthest from Standard Japanese (Shibata 1978). This statement represents most people's attitude toward this dialect; it is not popularly considered to be an ordinary Japanese dialect. In this section I will discuss some of the features of the Tohoku Dialect.

1.1.1. Voicing of postvocalic voiceless stops and affricates. Stops and affricates are voiced following vowels, as in (1) (in this and following examples, I will give the dialect (D) in the first line and Standard Japanese (S) in the second):

(1) D: Údzu-do kaedde-guru.
    S: Útsu-to kâette-kuru.
    hit-if return-come
    'If (you) hit (it), (it) comes back.'

Depending upon the rate of speech, this voicing can take place across a morpheme boundary or even a word boundary.

1.1.2. Absence of pitch accent. Pitch accent, a general feature of Japanese, does not exist in the Southeastern Tohoku Dialect. The standard language constrains between hâshi 'chopsticks' and hashi 'bridge', kakî 'persimmon' and kâki 'oyster', etc. are neutralized in this area.¹

1.1.3. Phrase-final high-low intonation. In the eastern coastal area of the Tohoku region, the final mora of a phrase is often lengthened, beginning with a higher pitch than the preceding words and then dropping to and holding at midlevel. For example:

(2) Kinôo uji-sa kaeddarak
    Kinoo uchi-e kâettara
    yesterday home-to return-when
    'when (I) got home yesterday...'

1.1.4. Merger of i and u. I /i/ and u /u/ merge as u /ki/. Furthermore, shi /ʃi/, su /su/, and shu /ʃu/, zi /dzi/, zu /dzu/, and dju /dzu/ merge as su /si/ and zu /dzi/ respectively, yielding many homonyms. For example, shûjûsu 'surgical operation', shichiji 'seven o'clock', and shijsutu 'historical fact' are all pronounced as suzuzu. According to Shibata (1978), it is this last merger which is responsible for the popular name of the Tohoku Dialect Zuuzuuben, 'Zuuzuu dialect'.

1.1.5. Usage of tag-particle be. According to Inoue 1985, the particle be functions in the Tohoku Dialect to express intention and guessing, although some subdialects have only one of these two functions. This feature is often exaggerated by mass media in its TV dramas and shows. Be with this function is actually not only found in the Tohoku District, but also in Kantoo District as well as Chuubu (Central) District. Recently, the form be has been subjected to a variety of morphophonemic processes, e.g.: (Inoue 1985b)

(3) okiru + be--> okirube, ogirube, ogimbe, ogippe
okiru + da + be---> okiru-dabbe, okiru-dambe
wake up intention//guess
'Let's wake up.' OR '(Someone) will wake up.'

1.1.6. Relatively little difference between men's and women's speech. Standard Japanese is characterized by relatively large differences between men's and women's speech. However, Kindaichi (1994) claims that this differentiation occurred very recently after the Meiji Restoration, and the difference can be observed only in the metropole, that is, the former capital, Kyoto, and the present capital, Tokyo. Like other non-­metropolitan dialects, the Tohoku Dialect in general lacks a sharp distinction between men's speech and women's speech. Both men's speech and women's speech in this dialect are regarded as similar to men's speech in Standard Japanese, because many features of men's rough speech in Standard Japanese are used by both men and women in the Tohoku Dialect. As a result, Tohoku women's speech is regarded to be rough and unrefined, as opposed to that of Standard Japanese women's speech. For example, the very informal first person pronoun ore, the second person pronoun omae, and the affective particle yo, which are used exclusively by men in Standard Japanese, are common among both men and women in Tohoku Dialect, as in (4)

(4) Ore yo, kinna kaze hiide yo, (dialect, male and female)
Ore yo, kinoo kaze hiite yo, (Standard, male)
I yesterday cold catch
'Yesterday, I caught cold.'

2. Image and perception. The features listed above are connected with the image of backwardness associated with Tohoku Dialect. As noted by Shibata (1978), the name zuuzuuben itself reflects the characteristics of the dialect which are salient in popular perception: voiced sounds, confusion among shi, su and shu and between zi, zu and dju, and heavy and slow talk which makes it sound as though syllables are being lengthened. Kindaichi (1994) states that negative feelings in Japanese tend to be expressed with voiced sounds and positive feelings with voiceless sounds, so that in a number of Japanese onomatopoetic words, negative feelings are expressed with voiced sounds; this would cause speakers of Standard Japanese to have a generally negative feeling towards the voicing processes of the Tohoku Dialect. The voicing, slow talk, lower pitch, and distinctive intonation of the Tohoku Dialect gives the appearance of backwardness while the minimal difference between men's and women's talk creates the impression that Tohoku women are unrefined tomboys. The general feeling of non-Tohoku people is that the Tohoku Dialect sounds kurai 'dark' or 'dismal', and omoi 'heavy'. As a result, the Tohoku Dialect is stigmatized. Various writers from Tohoku have described this stigmatization from their own experience (Dazai n.d., Inoue n.d. et al.)

As in other societies, the low status of the Tohoku Dialect reflects the low prestige of its speakers. The Tohoku area has suffered from poverty, its average income is the lowest in the country (outside of Okinawa), it has poor education, it has relatively little modern industry, and it relies heavily on rice cultivation.

The Tohoku Dialect sometimes appears in the mass media as a 'marked' dialect. The TV soap opera 'Oshin' --a biography of a woman born to a very poor peasant family in a small village in Tohoku, sold as a housemaid at the age
of six, and having an extremely hard life—was the most popular show in Japan in the 1980's. The hunger, poverty, and misery of her life were expressed vividly in this dialect, and this reinforced its low status.

Another TV character is an American actor, who speaks the Tohoku Dialect of Yamagata Prefecture. This young, blond Californian, who was originally sent to Tohoku for an English education program, has become extremely popular and now appears on several national TV shows. He is amusing because he speaks this unvarnished country dialect fluently and unabashedly in front of fashionable Japanese celebrities. Comedians who speak this dialect are laughed at because of their dialect and hillbilly-like appearance rather than the content of their talk.

In spite of this stigmatization at the national level, it cannot be said that Tohoku Dialect lacks local prestige (Labov 1963). This was reflected in an incident in 1993, when a middle school student who had just moved to Yamagata and so could not speak the dialect properly died as a result of *ijime*, the ritual teasing given to social outsiders, which involves verbal and physical abuse.

3. Definition of the Yonezawa Dialect. Since the end of the 19th century, dialectologists as well as people in Yonezawa have defined the Yonezawa Dialect as the language spoken in the castle of the Uesugi Clan and the immediately surrounding town. The dialects spoken in the countryside around Yonezawa have been excluded from this definition, and these neighbor dialects were labeled as *inaka-go* or *zaigou-go* 'country languages' (Kamimura 1977). This attitude still remains among people of Yonezawa, even though the new administrative area of Yonezawa covers a much larger area. Contrary to the dialectologists' claim that the Tohoku area lacks honorifics (Kindachi 1977), the Yonezawa Dialect is characterized by the frequent usage of honorific and politeness forms which were developed by the caste system. Example (5), from a conversation between two old women, illustrates these honorifics:

(5)  
\[
\begin{array}{cccc}
\text{Iya} & \text{iya} & \text{sappari} & \text{kawayananne} & \text{besshi?} \\
\text{Āra} & \text{ara} & \text{sappari} & \text{kwararetēnai} & \text{deshō?} \\
\text{oh} & \text{oh} & \text{at-all} & \text{not-change (honorific)} & \text{aren't you} \\
\end{array}
\]

'Oh, you haven't changed at all, have you?'

(It will be noted that postvocalic voicing is variable, e.g. *sappari* in (5) rather than *subbari*; I will return to this variability below.) The speaker uses a dialectal honorific form, *kawayananne* (which means 'not change') instead of the plain form, *kawanne*. Also, the speaker uses combination of a tag particle be explained in 1.1.5. and affective particle *shi*. This polite particleshi marks Yonezawa dialect. Examples (6) and (7) are taken from a music lesson:

(6)  
\[
\begin{array}{cccc}
\text{Chotto,} & \text{mīdede} & \text{kudai} & \text{na?} \\
\text{Chōtto,} & \text{mītete} & \text{choodai} & \text{ne?} \\
\text{a-while} & \text{watch} & \text{please} & \text{TAG} \\
\end{array}
\]

'Please, watch for a while.'

\[
\begin{array}{cccc}
\text{Rogu} & \text{o} & \text{motto} & \text{osanai} & \text{to dame da.} \\
\text{Rokū} & \text{o} & \text{mōto} & \text{osănāi} & \text{to damē yo.} \\
\text{six} & \text{PAR more} & \text{not-push} & \text{if bad} & \text{PAR} \\
\end{array}
\]
'You have to push the #6 string harder.'

Sogo nana da ba?
Soko nana de sho?
'there seven PAR TAG
'It's the #7 string, isn't it?'

(7) Koto ya san ni renraku tsukkara ya.
Koto ya san ni renraku tsūku-kara ne.
koto shop HON to contact because TAG
'We can contact the koto tuner.'

hai, son dogi wa oidatte kudai.
hai, sono toki wa ōdenināte kudasai
yes that time PAR come (honorific) please
'Yes, please come at that time.'

Aside from the honorific usage in (7), (6) and (7) are also characterized by some specific dialect features, i.e. the usage of kudai 'please' (kudasai in Standard Japanese) at the end of (7), and the interrogative particle ba at the end of (6) which is used only by local women of upper class households.

In these examples we can observe class and gender stratification of the dialect. Besshi, as in (5), is used by non-upper-class women and men of any social class, while ba, as in (6), is only used by upper-class women. Local honorific terms of this type are associated with solidarity and are particularly used in conversations between people who are familiar with each other. These usages are considered to be exclusive to Yonezawa, and serve to mark the speakers as different from those of surrounding areas. As a result, people living outside of the city often describe the people of Yonezawa and their language as otakai 'full of themselves'.

Even though the direct descendants of the Uesugi Clan no longer live in Yonezawa, those people who accumulated their wealth under the Shogunate have flourished. In reality, it is the merchants and manufacturers, especially of the safflower-dyed silk textile and sake, and the traders who have kept their wealth and remained at the top of the prestige hierarchy since the Meiji Restoration. During the period of rapid economic growth as well as the oil-shock of the 1970's, these traditional elite suffered a crisis. Some of them closed their factories and started different businesses, while others tried to stay in business by modernizing their products. Some are no longer at the top economically, but their family names still carry high prestige.

Usages which have local honorific and politeness forms are often labeled as hataya kotoba 'language of weavers'. This is the speech of the traditionally wealthy and prestigious families, and it is idealized locally as the most sophisticated version of the Yonezawa Dialect.

4. Age differences. There is a clear difference between the speech of younger and older speakers as can be found in many societies. For example, Table 1 gives data on postvocalic voicing (e.g. midai for Standard Japanese mitai) for 11 individuals of different ages.
As can be seen, postvocalic voicing is much more common for older speakers than younger speakers. In the speech of the younger children I recorded, I found considerable variation in use of dialect features. Toddlers spoke in a range of styles, sometimes even using colloquial Standard Japanese, and the 5th graders I recorded used dialect words but not phonological features (except for lack of pitch accent). Compared with the speech of their forties and fifties, that of high school students lacks local honorifics and polite forms; they do use dialectal words and phonological features (although, as can be seen from the data in Table 1, they do not have postvocalic voicing as much as older people). So we see that, in my data, the older the speaker, the stronger the features of the dialect are.

5. Apparent age grading. Since Labov's study of Martha's Vineyard in 1963, sociolinguists have paid particular attention to the distinction between real time and apparent time; an observed difference in synchronic usage between young and old people may be due to a change in language (a real time difference) or consistent changes over the lifetimes of individual speakers (an apparent time difference). It might appear that the phonological features as well as lexicon of Yonezawa Dialect is disappearing, and this is indicated by the children's speech. However, there is a sharp distinction in dialect usage in my data even between children just a few years apart in age, and it seems unlikely that a real time change could be instantiated so quickly and clearly.

Consider the data in Table 2, taken from a questionnaire survey of high school students:
Table 2--Use of the dialect among high school students

<table>
<thead>
<tr>
<th>Do you use the dialect:</th>
<th>% of yes</th>
<th>10th</th>
<th>12th</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaking to your teacher in class?</td>
<td>46%</td>
<td>78%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaking to your teacher outside of class?</td>
<td>51%</td>
<td>81%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>35</td>
<td>35</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In class--chi. sq.=8.27, p<.01, out of class--chi. sq.=6.17, p<.025

We could not expect a real time change to affect 30% of the population in only two years. A more plausible interpretation is that the social context of interaction between students and teachers changes significantly between 10th grade and 12th grade; by the time they become seniors, students are on a more equal basis with their teachers, joke with them more, and use more dialect. Additionally, I observed that former classmates of mine still living in Yonezawa now have more dialectal features than they did 15 years ago, and other local people have also told me that they have observed this pattern.

In this situation, the most likely conclusion is that we have apparent age grading. It appears that young children are exposed to both the dialect and the standard and mix dialectal words into their usage, but they do not have dialectal phonological patterns such as merging of i and u, postvocalic voicing or local honorifics. As they get older, they apply dialectal phonological rules more frequently and learn local honorific forms, reserving the standard language for interaction with outsiders, written language, and certain types of social situations.

While increased dialect usage over the course of an individual's life has not often been reported, it is not unknown. In Labov's 1963 study of Martha's Vineyard, he found that local residents who went away to college and then returned with a commitment to spend their lives on the island showed greater dialect usage over the course of time (Labov 1972: 31):

(E.) is a college graduate who tried city life, didn't care for it, came back to the island and built up several successful commercial enterprises on the Chilmark docks. He shows a high (ay) at 21, considerably more centralized (=dialectal Y.M.) than anyone else I have heard at Chilmark. One evening, as I was having dinner at his parents' house, the conversation turned to speech in general, without any specific reference to (ay) or (aw). His mother remarked, "you know, E. didn't always speak that way...it's only since he came back from college. I guess he wanted to be more like the men on the docks..."

There is a clear reason why maintaining the dialect is possible in modern Japan. Written Japanese is always Standard Japanese; reading, writing, and examination is conducted in Standard Japanese. If Yonezawans have to talk with outsiders or become national figures and speak on Radio or TV, they base their speech on the standard language.

6. Prestige vs. Stigmatization. It is generally accepted that the Tohoku Dialect has been a symbol of backwardness and an object of ridicule. However, analysis of my questionnaire survey shows that Yonezawans of all age groups feel a strong pride in their dialect.
Unlike other sociolinguistic studies, my studies found that pride in the Yonezawa Dialect is most strongly expressed by people in their forties, who currently hold economic and political power in the city. At meetings about local business and politics, they use the dialect. This pride comes from their historical background, as the descendants of the great Uesugi Clan, who fought against the Tokugawa Shogunate. The Uesugi Clan was humiliated by the central government after their defeat, but rather than causing a feeling of shame, this served to reinforce their hostility toward the central government, and this is reflected in their language attitudes.

Maintenance of the dialect is also strengthened by the traditional kinship system. The patrilocal system, whereby the oldest son to live in or near his parents' household in order to continue the family name and the family business, has a strong stabilizing influence on the social structure of the city.

Studies of Western cities such as Trudgill's (1972) analysis of Norwich have found that women have a tendency towards more standard usage. I did not observe this tendency in my study of the Yonezawa Dialect; men and women equally speak the dialect, although some lexical items are used exclusively by women. I believe that this difference is due to the fact that, in Western societies, members of the local higher classes would speak a more standard variety, so that women hoping to enter these higher classes would also speak in a more standard fashion. In Yonezawa, on the other hand, the local higher class people are as strongly committed to the local dialect as anyone else, and speaking the standard (other than in specific contexts) is not seen as elegant or educated but simply alien.

Milroy and Milroy (1992) note that in societies where informal ties are strong, the dialect tends to be maintained. My case study of Yonezawa shows that the local residents' ties of kin or friends are not simply informal but extend to the public domain, such as the economic and political sphere; therefore, the mechanism for maintaining the dialect also extends to the whole public domain. The prospects for the continued survival of the Yonezawa Dialect are therefore better than they would be in a case where the highest class local residents look outside their community for sources of prestige.

NOTE

1This characteristics is also observed in the neighboring Kanto District except Tokyo and Kanagawa.

REFERENCES

SPLIT TENSE AND MOOD INFLECTION IN WAMBAYA*

Rachel Nordlinger
Stanford University

1. Introduction

In this paper, I discuss a phenomenon in the Northern Australian language Wambaya in which functional information, in this case tense and imperative mood, is simultaneously marked on two different heads in the one clause. I will argue that this split inflection challenges the assumption of movement-based frameworks that each piece of functional information present in the clause corresponds to a single phrasal head in phrase structure. Rather, I claim that it strongly supports a unification-based model of grammar in which words enter the syntax fully inflected and each word corresponds to only one node in constituent structure. To substantiate this claim, I will show how Wambaya tense and mood split inflection can be accounted for easily and straightforwardly within the framework of Lexical-Functional Grammar.

2. The Data

Wambaya is syntactically similar to the better known Australian language Warlpiri and has many of the characteristics that are usually associated in the literature with non-configurationality (e.g. Hale 1983, Speas 1990) such as free word order, extensive null anaphora, discontinuous constituents, and lack of evidence for a VP constituent. Like many languages with some or all of these properties, Wambaya has an extensive case marking system and subject and object pronominal agreement.

All finite clauses in Wambaya contain a second position auxiliary which itself contains bound pronouns registering the subject and object arguments of the clause, as well as morphemes marking such things as tense, mood, aspect and direction. The auxiliary is obligatory in every finite verbal clause; only in non-finite clauses and in clauses with nominal predicates does the auxiliary not appear. The auxiliary has no morphological head, but is created by a slot and filler template, the basic structure of which is shown in (1):

(1) Subj (Obj) Tense/Mood (/Aspect/Directional)¹

Both the subject marker and a tense/mood affix are obligatorily present in the auxiliary at all times. The tense/mood suffix may minimally be either a suffix encoding only tense, or one encoding only mood. However, it is also possible to have suffixes which combine tense information with either aspect, directional or other mood information. Object markers are only present when the clause is transitive, and then only when the object is first or second person (third person object is not registered in the auxiliary, see Nordlinger 1993b).

Tense marking on the auxiliary differs according to its morphological structure, as exemplified in (2). Thus, when there is an object marker in the auxiliary, the tense distinction is future/non-future (a), and when there is no object in the auxiliary the
tense distinction is past/nonpast with non-singular subjects (b) or past/present/future with singular subjects (c).

(2)  

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>

This is represented in table form in (3). Since some of the tense morphemes are homophonous, I have distinguished them with subscripts.

(3) Tense distinctions in the auxiliary

<table>
<thead>
<tr>
<th></th>
<th>Past</th>
<th>Present</th>
<th>Future</th>
</tr>
</thead>
<tbody>
<tr>
<td>With Obj</td>
<td>-a₂</td>
<td>-u</td>
<td></td>
</tr>
<tr>
<td>Without Obj</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sg. subj</td>
<td>-a₁</td>
<td>-Ø₁</td>
<td>-u</td>
</tr>
<tr>
<td>non-sg. subj</td>
<td>-a₁</td>
<td>-Ø₂</td>
<td></td>
</tr>
</tbody>
</table>

Due to the fact that the auxiliary contributes no predicate information, it must always be accompanied by a verb, although the two are not required to be contiguous in the clause. (4) and (5) provide examples of some more complex Wambaya sentences in which the verb and the auxiliary are separated even by the subject (4).²

(4) Durraji-ni be.frightened-SIM:SS 3:sg:S-Pst nagarna that:sg:NOM bard-bi run-nFut

The cockatiel ran away frightened.


You will give me that cooked (meat) of yours!

However, since there is quite extensive null anaphora in Wambaya, the clause need not contain any noun phrases at all. Thus the minimal (and quite typical) sentence consists of simply a verb followed by the auxiliary. For simplicity, I will use only these minimal sentences in this paper.

So, we have seen that the auxiliary in Wambaya carries tense marking which differs according to the morphological structure of the auxiliary, as shown in (3). Now, the verb in Wambaya also carries tense marking, in which case the distinction is simply future tense versus non-future tense. Thus, tense is marked twice in all
Wambaya clauses in which tense appears. Some typical Wambaya examples, which illustrate this split inflection are as follows:

(6)  Ngaj-ba  ngu-ny-u.  
     see-Fut  1:sg:A-2:O-Fut  
     I will see you.

(7)  Ngaj-bi  ngi-ny-a.  
     see-nFut  1:sg:A-2:O-nFut  
     I saw/see you.

(8)  Bard-bi  irr-a.  
     run-nFut  3:pl:S-Pst  
     They ran.

(9)  Bard-bi  irri-Ø.  
     run-nFut  3:pl:S-nPst  
     They are running.

(10) Bard-ba  irri-Ø.  
      run-Fut  3:pl:S-nPst  
      They will run.

(11) Gulug-bi  ng-a.  
      sleep-nFut  1:sg:S-Pst  
      I slept.

(12) Gulug-bi  ngi-Ø.  
      sleep-nFut  1:sg:S-Pres  
      I am sleeping.

(13) Gulug-ba  ng-u.  
      sleep-Fut  1:sg:S-Fut  
      I will sleep.

In (14) we see that the verb can appear in the non-future/unmarked form even when the auxiliary is marked with future tense. Speakers say that (14) is simply a paraphrase of (6), and thus there is no (substantial) difference in meaning according as whether or not the verb is inflected with the future tense suffix.

(14) Ngaj-bi  ngu-ny-u  
      see-nFut  1:sg:A-2:O-Fut  
      I will see you.

Thus, as shown in these examples, not only is tense marked on two places in the Wambaya clause, but the system of tense marking is not necessarily the same: each instantiation of tense may have a different value. The table in (15) reflects the possible combinations we have seen so far - I will add to it shortly.
(15) Interaction between tense marking on verb and auxiliary.

<table>
<thead>
<tr>
<th>AUXILIARY</th>
<th>VERB</th>
<th>CLAUSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>-Ø₁</td>
<td>Present</td>
<td>Non-future</td>
</tr>
<tr>
<td>-u</td>
<td>Future</td>
<td>Future/Non-future</td>
</tr>
<tr>
<td>-a₂</td>
<td>Non-future</td>
<td>Non-future</td>
</tr>
<tr>
<td>-a₁</td>
<td>Past</td>
<td>Non-future</td>
</tr>
<tr>
<td>-Ø₂</td>
<td>Non-past</td>
<td>Non-future</td>
</tr>
<tr>
<td>-Ø₂</td>
<td>Non-past</td>
<td>Future/Non-future</td>
</tr>
</tbody>
</table>

Now, the situation as reflected in (15) becomes even more complicated once we consider imperative sentences. Imperative mood, in Wambaya, is part of the same system as tense marking. Thus, it is expressed using some of the same tense markers that we saw in examples (6) to (13). In this respect it contrasts with other mood categories in Wambaya such as ‘irreals’ and ‘hypothetical’ which are marked with distinct morphemes on the auxiliary, as shown in (16) and (17).

(16) Guyala gunu-ny-udi ngaj-bi.  
*He isn’t looking at you.*

(17) Nawu-Ø ngiyny-agba narunguji-ni.  
step.on-nFut 3:sg:fem:A-2:O-HYP car-ERG  
*A car might run you over. [Lit. A car might step on you.]*

In imperative clauses, the verb is marked with the future tense suffix while the auxiliary carries the non-future suffix with auxiliaries containing objects, and the non-past form with other auxiliaries.³

(18) Jiayj-ba girri-ng-a₂ manganyama!  
give-Fut 2:pl:A-1:O-nFut food(ACC)  
*Give (pl) me some food!*

(19) Jiayj-ba girri-Ø₂ manganyama!  
give-Fut 2:pl:A-nPst food(ACC)  
*Give (pl) him/her some food!*

(18) is particularly interesting since two seemingly incompatible suffixes - future and non-future - are combined to signal imperative mood. This example shows more clearly than any others we have seen, that the two tense markers cannot have the same source; one cannot simply be a copy of the other, as would have to be assumed in a movement framework. Rather, they are generated separately and then interact and constrain each other via unification, in a way that I will outline shortly.

Thus, in addition to there being split inflection of tense in Wambaya, these tense markers can also be used in imperative mood, and the functional information is split between the auxiliary and the verb in these constructions also. By adding this information to (15), we can get a more complete idea of the types of interaction between the marking on the verb and on the auxiliary. The table in (20) contains
the basic facts of Wambaya tense and imperative mood inflection that any theoretical account needs to be able to explain.

(20) Interaction between tense marking on auxiliary and verb (impossible combinations are starred)

<table>
<thead>
<tr>
<th>AUXILIARY</th>
<th>VERB</th>
<th>CLAUSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>-∅₁</td>
<td>Present Non-future</td>
<td>Present</td>
</tr>
<tr>
<td></td>
<td>Future</td>
<td>*</td>
</tr>
<tr>
<td>-u</td>
<td>Future Non-future</td>
<td>Future</td>
</tr>
<tr>
<td></td>
<td>Future</td>
<td>Future</td>
</tr>
<tr>
<td>-a₂</td>
<td>Non-future Non-future</td>
<td>Present/Past</td>
</tr>
<tr>
<td></td>
<td>Future</td>
<td>Imperative</td>
</tr>
<tr>
<td>-a₁</td>
<td>Past</td>
<td>Past</td>
</tr>
<tr>
<td></td>
<td>Future</td>
<td>*</td>
</tr>
<tr>
<td>-∅₂</td>
<td>Non-past Non-future</td>
<td>Present/Future</td>
</tr>
<tr>
<td></td>
<td>Future</td>
<td>Future/Imperative</td>
</tr>
</tbody>
</table>

3. The Account

I believe that these complex and unusual facts can be accounted for intuitively and straightforwardly in a unification based framework such as Lexical-Functional Grammar (LFG). LFG separates c-(onstituent) structure from f-(unctional) structure and therefore allows the same type of information to appear in two different positions in the c-structure, but to constrain each other by being unified at f-structure. Thus, there can be two instantiations of the same type of information in c-structure without the assumption that they are from the same syntactic place of origin.

The unification of related information is ensured by the mapping rules which link elements of c-structure to elements of f-structure. In particular, the unification of the information contained in the auxiliary in Wambaya with that in the verb is effected by the assumption that both the verb and the auxiliary are co-heads of the sentence (each annotated with ↑ = ↓) and that therefore, their f-structures will be identified with each other and unified.₄

Partial sublexical entries for the various suffixes discussed above are given in (21), followed by discussion of these entries. I am using the label TM to refer to the category which includes tense and imperative marking. This is distinct from the MOOD system, to which irrealis mood belongs. Only the features relevant to this discussion are included in these sublexical entries. The ways in which these sublexical entries can account for the tense and mood inflection facts given in (20) are demonstrated with some example structures in (22) to (25).

(21) Partial lexical entries for (some) Wambaya tense and mood suffixes.

**Verbal suffixes:**

a. -∅, -bi:  
   (↑ TM)

b. -ba:  
   (↑ TM) = FUT v IMP
Auxiliary suffixes:
c. -Ø: \((\uparrow \text{TM}) = \text{PRES}\)
d. -a: \((\uparrow \text{TM}) = \text{PAST}\)
e. -u: \((\uparrow \text{TM}) = \text{FUT}\)
f. -a: \((\uparrow \text{TM}) = \sim\text{FUT}\)
g. -Ø: \((\uparrow \text{TM}) = \sim\text{PAST}\)
h. -udi, -ushi: \((\uparrow \text{TM}) = \text{PRES}\)
\((\uparrow \text{MOOD}) = \text{IRR}\)

Thus, the fact that the non-future tense form of the verb appears to function more like a general, unmarked form is captured by the fact that its lexical entry requires the presence of a TM value, yet does not provide that value itself. And, the multiple functions of the future tense form of the verb are captured by the disjunction, FUT or IMP in its lexical entry. The auxiliary suffixes have lexical entries corresponding exactly to the earlier descriptions of their functions (as in (3)).

Given these sublexical entries for the tense and mood suffixes, the facts of Wambaya tense and mood inflection follow, as is demonstrated by the example sentences in (22-25). Only the features relevant to this discussion are included in the lexical entries in these examples.

Since Wambaya is a non-configurational language, the issue of its constituent structure is a non-straightforward one, and I will not go into the details of it in this paper. Instead, for the purposes of this discussion, and in order to focus more clearly on the issue at hand, I will use a simplified constituent structure here. I have provided a brief outline of the constituent structure I actually assume for Wambaya, including an account of the second position of the auxiliary, in an appendix at the end of this paper (for a more detailed discussion of this issue see Nordlinger 1995).

(22) Gulug-bi irri-Ø.
sleep-nFut 3:pl:S-nPst
They’re sleeping/ they will sleep.

In (22), the verbal lexical entry requires the presence of a TM value, yet does not provide one itself. This value is provided by the unification of the verbal lexical entry with the auxiliary’s lexical entry which gives the clause as a whole the TM value \(\sim\text{PAST}\), thereby reflecting the ambiguity of this utterance between present tense and future tense readings.
(23)  Gulug-ba    irri-Ø.
sleep-Fut     3:pl:S-nPst

They will sleep.

In (23) the combination of the TM values of the auxiliary and of the verb yield the TM value of FUT for the clause. Note that the fact that the subject is third person prevents the interpretation of the verbal inflection as imperative here. If the subject was second person the sentence would be potentially ambiguous between a future tense indicative and an imperative meaning.

(24)  Ngaj-ba    nyi-ng-a!
see-Fut      2:sg:A-1:O-nFut

Look at me!

In (24), the verb provides the information that the clause is imperative, while the auxiliary simply constrains the clause to not having future tense. Although the auxiliary inflection does not appear to contribute any information to the clause, its presence is very important as it blocks unification of this auxiliary suffix with the verbal suffix that has future meaning. This ensures that the only time this auxiliary suffix can co-occur with the verbal suffix -ba is when the clause is imperative and therefore serves to signal the fact that, in this clause, the suffix -ba is marking imperative not future tense.

(25)  Dagumaj-ba    gun-u.
hit-Fut       3:sg:masc:A-Fut

He will hit him/her/it.
In (25) the presence of the FUT value on the auxiliary suffix means that the verbal suffix -ba in this example must be marking future tense and not imperative since if it were imperative, the two TM markers would not be able to unify. Note that the present and past tense auxiliary suffixes (given in (21c) and (21d) respectively) are similarly blocked from co-occurring with the imperative verbal suffix. Furthermore, their TM values block them from co-occurring with the future tense form of the verbal suffix also.

Thus, in this way, this account can explain all of the possible combinations and their meanings given in (20). Furthermore, it can do it in a simple and straightforward manner, without needing to postulate any additional process or mechanism of information flow. This is in contrast to a movement based account of split inflection such as Mitchell's (1991) analysis of split inflection in Finnish.6 In order to capture the phenomenon of split inflection in Finnish, Mitchell postulates an extra mechanism for information flow, in addition to movement, in which a functional feature is generated in a unique head position in phrase structure and then is ‘percolated’ or ‘spread’ onto another node, where it is ‘spelled out’. In this type of analysis, the ‘future’ tense information in a sentence such as (25) would be generated in the auxiliary position (i.e. Infl) and would then percolate down onto the verb where it is also realized. So, in this type of analysis, the tense information is generated only once - thereby corresponding to only one phrasal head in phrase structure - and then spreads to the other parts of the clause in which it also appears in surface structure.

One serious problem with this approach is that it is not clear how it could cope with cases in which the two tense inflections are compatible, yet different, such as in example (23). In this example, the value of the tense marker on the verb is ‘future’ while that of the auxiliary tense suffix is ‘non-past’. Although these two values are semantically compatible, as shown by the fact that they can unify in the LFG analysis, their difference means that one could not simply be a copy of the other. There is no way that a ‘non-past’ feature on the auxiliary could percolate and in doing so become a ‘future’ tense feature on the verb.

This movement-based analysis has a further disadvantage with respect to the LFG analysis presented here in that it must appeal to an additional ‘percolation’ mechanism by which information flows throughout the structure; a mechanism which is not required by anything else in the framework. Furthermore, this mechanism, in allowing for functional information to spread and appear multiply marked in the clause, seems to undermine the basic premise which the framework strives to maintain: that each functional feature corresponds to a unique position in phrase structure.
4. Conclusion

In this paper I have argued that the tense and mood inflection facts in Wambaya provide strong support for a model of grammar which adopts the principle of lexical integrity and uses unification as the primary method of information flow. I have presented an analysis of the Wambaya data within the framework of LFG and shown how we can make use of the split between c-structure and f-structure to unify tense information that is represented separately in phrasal structure. Furthermore, I have argued that this analysis is preferable to a movement-based analysis since it can account for all the data without the need to postulate an additional and otherwise unnecessary mechanism to do so, and because it can capture intuitions and generalisations about the tense and mood marking system in Wambaya in a simple and straightforward way.

APPENDIX
Wambaya Constituent Structure

Following ideas in Simpson (1991), Kroeger (1993), Austin and Bresnan (to appear) and Halpern (1992) I assume the following constituent structure for Wambaya simple sentences.\(^7\) The basic Wambaya sentence is made up of an IP with an optional SPEC position. The auxiliary is in INFL and the non-projective, exocentric category S is generated as a sister to INFL. The constituents of S are NP and V and order within S is free. The auxiliary is an enclitic, phonologically dependent on the preceding word. Thus, when SPEC, IP is not filled (in which case the auxiliary is the first constituent in the clause), a phonological rule of prosodic inversion (Halpern 1992) causes the auxiliary to follow the first constituent of S. The structure of a sentence such as (13) - repeated here - is therefore as shown in (13'). Prosodic inversion is indicated with an arrow.

(13) Gulug-ba ng-u.
sleep-Fut 1:sg:S-Fut
I will sleep.

(13')

```
  IP
   |
   I'
   |
   I
   |
   S
   |
   ng-u
   |
   V
   |
   l:sg:A-Fut
   |
gulug-ba
   |
sleep-Fut
```

When SPEC, IP is filled, meaning that the auxiliary is not in initial position in the clause, prosodic inversion does not occur; the auxiliary attaches prosodically to the final constituent of the maximal projection in SPEC, IP.
(26) Alanga g-a yarru-Ø.
girl(NOM) 3:sg:S-Pst go-nFut

The girl went.

(26')

IP

NP I'

N I S

alanga ng-u V
girl(NOM) 1:sg:A-Fut
gulug-ba sleep-Fut

* I am greatly indebted to my Wambaya consultants, Molly Nurlanyma Grueman, Minnie Niymarrama Nimara and Mavis Bangarinyra Hogan, for the countless hours spent teaching me their language. I would also like to thank Joan Bresnan, Peter Sells, Will Leben, María-Eugenia Niño and Ian Green for valuable discussions, suggestions and comments on earlier versions of this paper. Naturally, I am solely responsible for any remaining flaws or inadequacies.

1 The marking of aspect and direction is not relevant to the present discussion and so will not be covered here.

2 The abbreviations that I use here and elsewhere in this paper, are as follows:

3 There is no auxiliary in an imperative clause with a singular subject and no object. An example is:

(ii) Jiyaj-ba manganyama!
give-Fut food(ACC)
Give (sg) him/her some food!

4 This is due to the universal principle of structure-function association such that the complements of functional categories are f-structure co-heads. Thus, in a configurational language such as English, the VP is a co-head with I. Since V is the head of VP, this means that the f-structure of the V will ultimately be identified with that of I. Since there is no VP in Wambaya, I co-heads the sentence with V, rather than VP.

5 Note that here, and in (g) the TM values are given with the negation marker (~) before the value instead of before the attribute-value pair (i.e. (~TM) = ~FUT rather than ~(~TM) = FUT). This is to make the distinction between the latter type of equation, which does not necessarily provide a tense value, but merely restricts the possible tense values that it can unify with, and the equation in these sublexical entries, which provides a tense value (and therefore satisfies the requirement of the non-future verbal inflection), but provides only the information that this value is not FUT.

6 See Niño (1994) for arguments against this analysis of Finnish.

7 See Nordlinger (1995) for a more detailed discussion of Wambaya sentence structure.
REFERENCES


What's in a place?
Extended uses of a physical-world noun in Japanese

Kyoko Hirose Ohara
University of California at Berkeley

1. Introduction

1.1. Target of analysis: A closed-class grammatical noun TOKORO 'place'

This paper investigates a family of related expressions in Japanese all of which use a particular closed-class grammatical noun, TOKORO. TOKORO, usually translated into English as 'place', in most contexts requires a modifier of one of the types shown in (1a):¹

(1a) ... Modifier TOKORO ...
where:
Modifier = Demonstrative;
NP-GEN (an NP followed by the genitive case marker);
Pred\textsubscript{attr} (the attributive form of a predicate);
or
Cl\textsubscript{attr} (a clause containing the attributive form of a predicate)

(1b) baakuree ni wa [nigiyakana] TOKORO ga takusan aru Berkeley LOC TOP lively NOM many exist
'There are many lively places in Berkeley.'

My interest in this class of expressions touches on: (i) semantic typology; (ii) characteristics of the grammar of Japanese; and (iii) lexical semantics.

1.1.1. Semantic distinction between place vs. non-place

First of all, TOKORO figures in a semantic typology between place vs. non-place. It has been observed that some languages, including Japanese, Chinese, and Korean, in contrast to languages like English, require explicit identification of places (Takubo 1984, Teramura 1968). In describing a physical movement in Japanese, for example, only place nouns are allowed as goals and sources. Thus, the literal translation of the English sentence (2a) shown in (2b) is ungrammatical, since \textit{doa 'door'} in Japanese is a non-place noun and cannot stand alone in a goal expression:

(2a) I went to the \textit{door}.
(2b) \textit{*}watai wa doa e itta
    \textit{I TOP door GOAL went}
    \textit{Intended meaning:}
    \textit{'I went to the door.'}

TOKORO must be attached to such non-place nouns, in order to change them into place nouns:

(2c) watai wa [doa no] TOKORO e itta
    \textit{GEN}
Lit.
'I went to the door place.'

For a discussion of this see Takubo (1984).

1.1.2. 'Formal nouns' in Japanese

Secondly, a language-internal consideration: TOKORO is generally considered to belong to a closed class of grammatical nouns called keishiki meishi 'formal nouns'. They are 'grammatical' nouns in that they are accompanied by a modifier unlike 'lexical' nouns in Japanese, which can appear on their own as saturated NPs. In addition, they may appear in expressions that do not function as ordinary nominals. For example, TOKORO plus a copula can function as an auxiliary:

(3a) Auxiliary TOKORO-da
kurisu ga benkyoo siteiru-TOKORO-da
Chris NOM is-studying COP
sizukani sina sai.
be-quiet
'Chris is just in the middle of studying. Be quiet.'

The sequence TOKORO-da occurs sentence-finally, adding the meaning 'just in the middle of' to the sentence. Note that it is structurally different from the predicative nominal use of the nominal TOKORO, as shown in (3b). Here, the sentence-final copula da functions to connect the subject baakuree wa 'Berkeley-TOP' and the predicative nominal headed by TOKORO 'the place Chris is studying'. If we were to omit the first NP baakuree wa, the sentence would be interpreted as an ellipted subject-predicative sentence with a contextually-understood subject as in (3c):

(3b) Nominal TOKORO
[NP baakuree wa] [NP kurisu ga benkyoo siteiru TOKORO] da
Berkeley TOP
'Berkeley is the place Chris is studying.'

(3c) Nominal TOKORO
[NP kurisu ga benkyoo siteiru TOKORO] da
'The place under discussion) is the place Chris is studying.'

In contrast to (3c), the auxiliary TOKORO-da sentence in (3a) is a full-fledged sentence without any ellipsis. 2

TOKORO may also appear as a part of clausal conjunctions or sentential adverbials. In (4a), TOKORO connects two clauses serving as a temporal-clause conjunction, while in (4b), it attaches to an adjectival noun and forms a speech-act modifier:

(4a) Conjunctive TOKORO
wata si ga hako o aketa-TOKORO ningyoo ga haitteita
I NOM box ACC opened doll NOM was-inside
Lit.
'When I opened the box, a doll was inside (the box).'</n(4b) Adverbial TOKORO
sottyoku na-TOKORO, kare no gendoo ni wa dare mo ga komatteiru
frank he GEN conduct DAT-TOP everyone-NOM fed-up
'To be frank, everybody is fed up with his conduct.'

Whereas the nominal TOKORO must be followed by a case particle (e.g. (1b) and (2c)) or by a copula (e.g. (3b) and (3c)), the conjunctive and adverbial TOKORO are not followed by any particle.

By appearing as an element of auxiliaries, conjunctions, and adverbials, TOKORO and other formal nouns form the basis of many constructions in Japanese, and they play an important role in the grammar. The auxiliary, conjunctive, and adverbial TOKOROS, however, have semantic and pragmatic properties distinct from the nominal TOKORO, as we will see below. Consequently, describing the constructions related to TOKORO has been problematic, especially in dictionaries and textbooks intended for learners of Japanese (Alfonso 1966 (1980), Makino and Tsutsui 1989, Martin 1975, Morita 1980, Teramura 1992).

1.1.3. Heterosemy

More importantly, TOKORO is interesting from the perspective of lexical semantics (cf. Langacker 1982, Talmy 1978). One of the findings of lexical semantic analyses of closed-class grammatical items is that grammatical morphemes are often shared by expressions that synchronically belong to different syntactic categories (e.g. Brugman 1984, 1988, Traugott 1986, 1988a). Studies of such phenomena, however, have tended to focus on English, with the exception of a recent work by Lichtenberk on Oceanic languages (1991). He specifically uses the term HETEROSEMY to refer to cases where two or more historically-related expressions share a common source element but synchronically belong to different morphosyntactic categories. He has shown that in heterosemy there need not be any property exclusively shared by all the related expressions and that the meanings of the derived expressions tend to be subjective.

Since Japanese grammarians agree that the expressions containing TOKORO are historically related, in the sense of deriving from the noun TOKORO, it is possible to regard the relations among the nominal, auxiliary, conjunctive, and adverbial TOKORO as those pertaining to heterosemy.3 Previous analyses of TOKORO, however, have not satisfactorily dealt with this issue.

1.2. Overview of the paper

The purpose of this paper is to examine the heterosemy structure of TOKORO, that is, the semantic extensions from the nominal TOKORO, by giving a synchronic account of the semantics of the auxiliary, conjunctive, and adverbial TOKOROs and by contrasting their meanings with those of the nominal TOKORO. I will be focusing on what I call temporal-marking and context-providing TOKORO constructions. I will be claiming that in these constructions, the speaker's point of view, commitment to a claim, or attitude toward a speech act is crucial and that TOKORO has undergone subjectification in meaning in the sense discussed by Traugott (1988b, 1989).4 Following Construction Grammar and Frame Semantics, I will assume that it is impossible to describe words in a context-free way, in other words, independently of the constructions in which they occur (Fillmore 1994, Fillmore et al. 1988, Kay and Fillmore 1994).5

The paper is organized as follows: Section 2 introduces the temporal-marking and context-providing TOKORO constructions and contrasts them with the nominal TOKORO. Section 3 examines the temporal-marking TOKORO-da. Section 4 looks into the context-providing TOKORO constructions. Finally, Section 5 concludes the analysis.
2. Temporal-marking and context-providing TOKORO constructions as contrasted with nominal TOKORO

In this section, I will introduce the temporal-marking and the context-providing TOKORO constructions, which will be the main targets of analysis in the rest of the paper, and briefly contrast them with the nominal TOKORO. I will show that due to their own syntactic and semantic constraints they should be characterized as grammatical constructions, that is, form-meaning correlations, different from the nominal TOKORO.

2.1. Nominal TOKORO

In addition to referring to a 'place', i.e. a 'point in space', the nominal TOKORO may refer to an 'aspect of an entity' - either concrete or abstract:

'Place'
(5a) resutoran ni iku nara, [oisii TOKORO] ga ii
restaurant GOAL go COND delicious NOM like
Lit. 'If (we) are to go to a restaurant, I prefer a delicious place.'

'Aspects of an entity'
(5b) kurisu ni wa [okasina TOKORO] ga aru
Chris LOC TOP funny NOM exist
'There is something funny about Chris.'
(5c) kurisu no iken ni wa [sansee dekinai TOKORO] ga aru
GEN opinion LOC TOP agree-cannot NOM exist
'There are aspects in Chris's opinion that I cannot agree with.'

The nominal TOKORO functions as a head noun of NPs. As I mentioned earlier, an NP headed by TOKORO may serve as an argument or adjunct. In that case, the NP is followed by a case particle:

[NP Modifier TOKORO]-case (Argument/adjunct of a matrix predicate)
(6a) baakuree ni wa [nigiyakana TOKORO] ga takusan aru
=(1b) 'There are many lively places in Berkeley.'
(6b) [dare mo inai TOKORO] de hanasitai
somebody even not-present LOC want-to-talk
'I would like to talk where nobody is present.'

NPs headed by TOKORO may also be used as predicative nominals, in which case the NP is followed by a copula (cf. (3b) and (3c)):

[NP Modifier TOKORO]-COP (Predicative nominal)
(6c) kamakura wa [sizukana TOKORO] da
Kamakura TOP quiet COP
'Kamakura is a quiet place.'

2.2. Temporal-marking TOKORO-da construction

What I call the temporal-marking TOKORO-da construction involves the auxiliary TOKORO-da. The sequence TOKORO-da occurs sentence-finally, and serves to locate the situation in time from the speaker's perspective.6 For example, in (7), attaching TOKORO-da to the end of the sentence adds the meaning that the
speaker construes the event to occur in the imminent future. This meaning is somewhat similar to the temporal sense of the English adverb *just:*

(7a)  
\[
\text{kurisu wa bangohan o taberu} \\
\text{Chris TOP supper ACC eat} \\
'\text{Chris will have supper.'}
\]

(7b)  
\[
\text{kurisu wa bangohan o taberu-TOKORO-da} \\
'\text{Chris is just about to have supper.'}
\]

We have seen in (3a) through (3c) that unlike predicative nominal TOKORO the sentence-final *da* in the auxiliary TOKORO-*da* sentences does not serve to connect two NPs in a subject-predicate relation. Moreover, in the temporal-marking TOKORO-*da* sentences, in contrast to the predicative nominal TOKORO, the sentence-final *da* cannot be negated:

(8a)  
\[
\text{Predicative nominal TOKORO} \\
\text{koko wa [tabako o suu TOKORO] de wa nai} \\
\text{here TOP tobacco ACC smoke COP TOP NEG} \\
'\text{Here is not the place to smoke.'}
\]

(8b)  
\[
\text{Temporal-marking TOKORO-*da* construction} \\
*\text{kurisu wa tabako o suu-TOKORO-de-wa-nai} \\
\text{Chris COP TOP NEG} \\
\text{Intended meaning:} \\
'\text{Chris is not about to smoke.'}
\]

Since the temporal-marking TOKORO-*da* sentences are subject to constraints which are absent in the predicative nominal TOKORO, they should be characterized as licensed by a different grammatical construction.

2.3. Context-providing TOKORO constructions

In the context-providing TOKORO constructions, the constituent to which TOKORO is attached supplies a background for the main clause.7 There are three types of such constructions.

The first type involves temporal clause-linking, which encodes a temporal relation between two specific past situations. The subordinate clause, S1, to which TOKORO attaches establishes a temporal setting for a subsequent situation described in the main clause S2:

**Temporal clause-linking: S1 (subordinate clause)-TOKORO, S2 (main clause)**

(9)  
\[
\text{syagai ni deta -TOKORO, ame ga huttekita} \\
\text{out-of-a-car LOC got-out rain NOM started-to-rain} \\
'\text{When (I) got out of the car, it started to rain.'}
\]

In the second type, S1 provides evidence for the speaker's claim made in S2:

**Evidential clause-linking: S1-TOKORO-*de-wa, S2**

(10a)  
\[
\text{sinbun de watasi ga yonda-TOKORO-de-wa syusyoo wa} \\
\text{newspaper LOC I NOM read LOC TOP Prime-Minister TOP}
\]
kaihuku ni mukatteiru yooda
recover GOAL head-PROG seem
'According to what I read in the paper, the prime minister seems to be
recovering (from his illness).'
(10b) sono sima wa mita-TOKورو-de-wa enosima gurai no ookisa desu
that island TOP saw Enoshima about GEN size COP-POLITE
'Judging from what (I) saw, that island is about the size of Enoshima.'
(Alfonso 1966 (1980))

In the third type, TOKورو attaches to a modifier denoting the speaker's
sincerity. The sequence as a whole functions as a speech-act modifier describing the
speaker's attitude toward the speech act in the main clause. That is, it lays the
ground for a speech act:8

Speech-act modifier: Modifier-TOKورو, S
(11) sottyokuna-TOKورو, kare no gendoo ni wa dare mo ga komatteiru
=(4b) 'To be frank, everybody is fed up with his conduct.'

In these sentences, TOKورو exhibits syntactic and semantic properties
different from those of the nominal TOKورو. In the first and the third types, TOKورو
is not accompanied by a case particle or a copula. Moreover, in all of the
three types, S1 cannot be negated. Thus, they involve grammatical constructions
different from the nominal TOKورو.

Having introduced the main targets of analysis in the paper, I will now
discuss the semantics of each of these constructions.9

3. Temporal-marking TOKورو-da construction

In discussing the temporal-marking TOKورو-da construction I will show
that its meanings are subjective, compared to those of the nominal TOKورو.
The sentence-final TOKورو-da serves to locate the event in time from the
speaker's perspective:

(12a) kurisu wa bangohan o taberu
=(7a) 'Chris will have supper.'
(12b) kurisu wa bangohan o taberu-TOKورو-da
=(7b) 'Chris is just about to have supper.'
(13a) kurisu wa bangohan o tabeteiru

is-eating
'Chris is having supper.'
(13b) kurisu wa bangohan o tabeteiru-TOKورو-da
'Chris is just in the middle of having supper.'
(14a) kurisu wa bangohan o tabeta

ate
'Chris had supper.'
(14b) kurisu wa bangohan o tabeta-TOKورو-da
'Chris just had supper.'

I argue that in the temporal-marking TOKورو-da construction the speaker
emphasizes the event's relevance to the present and the event's temporal closeness
to the present. Because of the speaker's emphasis on the two properties of the event described in the sentence, this construction is incompatible with habituals, perfect, and states, as we will see presently.

3.1. **Relevance to the present**

Firstly, **TOKORO-da** is incompatible with the habitual aspect:

(15a) Prospective/Habitual

kurusu wa puuru de oyogu

swimming-pool LOC swim

'Chris will swim in the pool.'  (a **prospective** reading)

or

'Chris habitually swims in the pool.'  (a **habitual** reading)

(15b) Prospective

kurusu wa puuru de oyogu-TOKORO-da

'Chris is just about to swim in the pool.'

Although (15a) may have either a prospective (i.e. aspect which relates some future event with the present) or habitual reading, adding **TOKORO-da** results in a prospective reading only, as the translation of (15b) shows.

Comrie (1976:27-28) defines habituals as describing 'a situation which is characteristic of an extended period of time, so extended in fact that the situation referred to is viewed not as an incidental property of the moment but, precisely, as a characteristic feature of a whole period'. In other words, since habituals are by definition temporally unbounded, they cannot be relevant to the present moment, and therefore they are incompatible with the meaning of the **TOKORO-da** construction.

Secondly, **TOKORO-da** is incompatible with state verbs:

State verbs

(16a) hanako wa eigo ga dekiru

able

'Hanako is able to speak English.'

(16b) *hanako wa eigo ga dekiru-TOKORO-da

Lyons (1977) notes that states last or endure through time and that they are homogeneous throughout the period of their existence. That is, states are usually temporally unbounded, and thus cannot be construed as relevant to the present moment. It is worth noting that although state verbs are not licensed by the **TOKORO-da** construction, not all state predicates are disallowed. Adjectives that denote a temporary state such as 'psych' adjectives are licensed by the construction. The sentences in (17) convey the meaning of 'be right in the state of' or 'at the point of':

Temporally-bounded state predicates

(17a) nodo ga kawaita. biiru demo hosii-TOKORO-da

throat NOM dried beer something-like want

Lit. 'I'm thirsty. I am just at the point of wanting (a bottle of) beer.'

(17b) syusyoo no sin'i o toitadasi-tai -TOKORO-da

prime-minister GEN real-intention ACC question-want-to
Lit.
'I am just at the point of wanting to question the Prime Minister's real intention.'

These sentences show that what is incompatible with TOKORO-da is the temporal unboundedness. As long as a state predicate describes a temporally bounded state, the predicate is licensed by the construction, since a temporary state can be relevant to the present moment.

3.2. Temporal closeness to the present
The perfect aspect is incompatible with the construction in question:

Perfect
(18a) kurisu wa san huransisuko ni tuiteiru
Chris TOP San Francisco LOC has-arrived
'Chris has (already) arrived in San Francisco'

(18b) *kurisu wa san huransisuko ni tuiteiru-TOKORO-da

The perfect is usually defined as the continuing relevance of a previous situation (Comrie 1976:56). Thus, characterizing the semantics of the construction as emphasizing the event's relevance to the present is not sufficient to account for why the perfect aspect is not licensed by the construction. There is another important meaning of the TOKORO-da construction: the speaker emphasizes the event's temporal closeness to the present. In fact, in contrast to (18b), perfect of recent past, or so-called 'hot news perfect', is compatible with TOKORO-da:

Perfect of recent past ('hot news perfect')
(19a) kurisu wa san huransisuko ni tuita
Chris TOP San Francisco GOAL has-arrived
'Chris has arrived in San Francisco.'

(19b) kurisu wa san huransisuko ni tuita-TOKORO-da
'Chris has just arrived in San Francisco.'

In other words, whether or not the event is construed as temporally close to the present is crucial to the acceptability of a TOKORO-da sentence.

3.3 The speaker's point of view
We just saw that the function of the temporal-marking TOKORO-da construction is to encode the speaker's emphasis on two properties of the event described in the sentence: its relevance to the present, and its temporal closeness to the present. The meaning of the construction is thus subjective in that it pertains to the speaker's point of view toward the event.

As a matter of fact, the speaker's point of view may be playing an even larger part in this construction. It has been observed that there is a contrast in use between plain progressive sentences and those with sentence-final TOKORO-da, as shown below (Makino and Tsutsui 1989):

Progressive
(20a) a kurisu ga hasitteiru
oh Chris NOM is-running
'Oh, Chris is running.'

(20b) *a kurisu ga hasitteiru-TOKORO-da
Intended meaning:
'Oh, Chris is just in the middle of running!'

Firstly, the TOKORO-da sentence in (20b) cannot be uttered as a surprise in a monologue, that is, in a situation where the speaker is by him/herself. TOKORO-da sentences seem to presuppose a speaker-hearer interaction.

Secondly, even when the hearer is with the speaker, the sentence (20b) cannot be used to draw the hearer's attention to the scene. Progressive TOKORO-da sentences are typically used when the scene is somehow inaccessible to the hearer, for example, on the phone, as (20c) shows:

Progressive TOKORO-da sentence
(20c) (On the phone)
kurisu wa syawaa o abiteiru-TOKORO-desu
TOP shower ACC is-taking COP-POLITE
ato de kakenaosu yooni iimasyoo ka
later LOC call-back COMPL say-POLITE-PRESUM Q
'Chris is just in the middle of taking a shower.
Can I have her call you back?'

I would like to speculate that these two properties of progressive TOKORO-da sentences, namely, requiring the existence of the hearer and requiring the scene to be inaccessible to the hearer, seem to suggest that the pragmatic function of the TOKORO-da construction is to inform the hearer of the speaker's point of view, or his/her subjective 'location' in space and time.

4. Context-providing TOKORO constructions

The temporal clause-linking, evidential clause-linking, and speech-act modifier uses of TOKORO can be construed as context-providing TOKORO constructions. In all of them, the TOKORO-marked constituent provides a background for the main clause. I will show that the three constructions correspond to uses in the three linguistic domains discussed by Sweetser, namely, the content, epistemic, and speech-act domains, and that the speaker's perspective, commitment to a claim, or attitude is an important aspect of the meanings of these constructions.

4.1. Temporal clause-linking: Uses in the content domain

The temporal clause-linking TOKORO encodes a temporal relation between two specific past situations and thus involves what Sweetser calls the content domain (1990). In this construction, S1 describes a situation which supplies a background for a subsequent change of state or 'discovery' in the speaker's cognition described in S2:

S1 (subordinate clause)-TOKORO, S2 (main clause)
(21a) ie e kaetta-TOKORO,
home GOAL returned
saihu o gakkoo ni oitekita koto ni ki ga tuita
purse ACC school LOC had-left COMP DAT realized
'When I got home, I realized that I had left my purse at school.'

This function of the temporal clause-linking TOKORO is similar to that of the clause-linking TO construction that Fujii has investigated (1991, 1993a, 1993b).10
That S1 functions to establish a temporal setting for S2 can be seen in the unacceptable sentence in (21b):

(21b) *kooen o sanpo siteita-TOKORO, arne ga hutteita
     park   ACC was-walking  rain   NOM was-raining

*Intended meaning*

'When I was walking in the park, it was raining.'

Based on our knowledge of the world, it is difficult to imagine a case in which the subject is walking in the rain but does not realize it until a particular point in time (Fujii 1993:13). In other words, the situation described in S1 in (21b) cannot serve as a temporal setting needed for a change of state or change in speaker's cognition, and the sentence is thus unacceptable. (21c) demonstrates that what is crucial for the construction is a change in the speaker's cognition brought about by the S1 event:

(21c) tutumi o aketa-TOKORO, ningyoo ga haitteita

=(4a) wrapping ACC opened  doll   NOM was-inside

*Lit.*

'When (I) opened the wrapping, a doll was inside.'

Here, S2 cannot be construed as encoding a change of state, since the doll must have been inside the box even before opening it. However, a change is brought about in the speaker's cognition by the S1 event, that is, by opening the wrapping, and thus the sentence is acceptable.11 The meaning of the construction therefore crucially involves the speaker's cognition or point of view.12

4.2. Evidential clause-linking: Uses in the epistemic domain

In the evidential clause-linking TOKORO construction, S1 gives evidence for the content of S2, or indicates the information source that led to the speaker's assertion in S2, and thus this construction may be characterized as supplying a background for an assertion. Since evidence is concerned with the speaker's knowledge or belief, this use pertains to the epistemic domain (see Palmer 1986 for the relation between epistemic and evidential). By providing evidence, S1 qualifies the speaker's commitment to the content of the assertion made in S2, and in this sense the meaning of the construction is subjective.13

S1-TOKORO-de-wa, S2

(22a) kare ga kiita-TOKORO-de-wa kaigi wa
     he   NOM heard      LOC TOP meeting TOP
     san-zi kara da soo desu
     three-o'clock ABL  COP hearsay-POLITE

'According to what he heard, the meeting seems to start at 3.'

(22b) tutaeareru -TOKORO de wa, kare wa mata rikkocho suru rasii
     has-been-conveyed    LOC TOP he    TOP again run-for-election hearsay

'According to what has been told, he seems to be running for election again.'

4.3. Speech-act modifier: Uses in the speech-act domain

The speech-act modifier use of TOKORO may be characterized as a context-providing construction in the speech-act domain. It provides the speaker's sincere attitude toward the speech act at the time of utterance. Note that the meaning of the
construction is subjective in that it reflects the speaker's own attitude toward the speech act.

**Modifier-TOKORO, S**

(23a) syoozikina-TOKORO, watasi ni wa nani ga mondai na no ka honest I at TOP what NOM problem COP NOMINAL Q wakaranai know-NEG

'To be honest, I do not know what is wrong (with it).'

(23b) *kare; no sotyoku na-TOKORO, kare; wa doosite ii no ka wakaranai. he GEN frank he TOP how-to-do NOMINAL Q does-not-know

'To his; being frank, he; doesn't know what to do.'

It is interesting to note that the same morpheme TOKORO is used for the context-providing function in three different domains and that these domains correspond to those that are relevant for other phenomena in other languages such as polysemy of modality in English as discussed by Sweetser. The context-providing TOKORO constructions seem to support Sweetser's view that because any utterance can be viewed in terms of its content, the speaker's reasoning, and the speech situation, the content, epistemic, and speech-act domains are relevant to the analysis of many linguistic phenomena.

5. **Conclusion**

I have given a synchronic account of the semantics of the constructions which involve the formal noun TOKORO in Japanese. In order to examine the semantic extensions from the nominal TOKORO to other TOKORO constructions, I have focused on the meanings of the temporal-marking and context-providing constructions. To conclude, let us contrast their meanings with those of the nominal TOKORO.

I have argued that the meanings of both temporal-marking and context-providing constructions can be characterized as subjective. The temporal-marking construction involves the speaker's subjective judgment toward the event, since it emphasizes the event's relevance and temporal closeness to the present. Likewise, the context-providing constructions pertain to either the speaker's point of view, commitment to the claim, or attitude toward the speech act. In contrast, the nominal TOKORO refers to objects in the real world, either a 'place' or an 'aspect of an entity'. It is thus possible to characterize the semantic extensions from the nominal TOKORO to the temporal-marking and context-providing constructions as a semantic-pragmatic change from less to more subjective. This pattern of semantic extension is in accord with Traugott's generalization of semantic change.

It is hoped that this study not only contributes to a better understanding of extended uses of other formal nouns in Japanese, but also may shed light on the nature of semantic extensions in general.

**NOTES**

* My deepest thanks go to Julia Elliott, Charles J. Fillmore, Katsuya Kinjo, and Eve Sweetser for reading earlier versions of the paper. I would also like to thank Yoko Hasegawa, George Lakoff, Kojiro Nabeshima, and participants of the Japanese Linguistics seminar at the University of California, Berkeley, for helpful
comments. I have also benefited from discussions with Shoichi Iwasaki, Yoshiko Matsumoto, and Shigeko Okamoto.

1 Except in idiomatic expressions such as:

(i.a) **TOKORO** kawareba sina kawaru
change-COND custom change
'Different places, different customs.'

(i.b) kare wa **TOKORO kamawazu** nekorogaru
he TOP mind-NEG sprawl
'He sprawls no matter where he is.'

When it refers to (the addressee's) address, it is preceded by an honorific prefix o:

(i.c) o-name to o-**TOKORO** o kaite kudasai
HON-name AND HON-ACC write please
'Please write your name and address.'

2 There is a syntactic test to distinguish the auxiliary TOKORO-*da* from the predicative nominal TOKORO: In general, when a subordinate clause is headed by the following noun, the case marking on the subject NP in the lower clause alternates between ga 'nominative' and no 'genitive'. The predicative nominative TOKORO sentences exhibit this alternation, but the auxiliary TOKORO-*da* sentences do not:

(i.a) Predicative nominal TOKORO
baakuree wa [kurisu ga/no] benkyoo siteiru TOKORO] da.
Berkeley TOP Chris NOM/GEN is-studying COP
'Berkeley is the place Chris is studying.'

(i.b) Auxiliary TOKORO-*da*
kurisu ga/*no* benkyoo siteiru-TOKORO-da. sizukani sinasai.
be-quiet
'Chris is just in the middle of studying. Be quiet.'

3 According to *Shogakukan Nihon Kokugo Daijiten* (The Japanese Language Dictionary), the nominal TOKORO already existed in the 8th century. The conjunctive TOKORO is attested in the early 11th century. It appears that the auxiliary TOKORO-*da* and the adverbial TOKORO did not appear until the early 20th century.

4 See also Okamoto 1992 for a synchronic account of semantics and pragmatics of sentence-final particles in Japanese including NO and KOTO, which have traditionally been classified as 'formal nouns' along with TOKORO.

5 This assumption is, by and large, compatible with that of Teramura (1978a:300, 1978b:323).

6 I will not be discussing the counterfactual marker use of TOKORO-*da*:

(i) Counterfactual Marker
kare ga kiteireba, komatta koto ni natteita-TOKORO-da
he NOM come-ASPECT-COND come-to-be-in-trouble-ASP
'If he had come, (we) would have been in trouble.'

(Takubo 1994)

7 I will not be discussing the concessive-clause marker TOKORO-de:

Concessive Clause Marker: S1-TOKORO-de, S2
(i) hasitta TOKORO de, ma ni awanai daroo
   run will-not-be-on-time
   'Even if (I) run, (I) will not be on time.'

8 Note that in contrast to the other context-providing constructions, TOKORO in this construction is not preceded by a clause:

(11') *watasi ga sottyokuna-TOKORO, kare no gendo ni wa dare mo ga
      I NOM frank
      komatteiru

9 It may be that metaphors belonging to the Event Structure Metaphor system, such as Properties are Possessions (cf. (5b)), States are Locations (cf. (7b) and (11)), and Perceiving Events are Moving over Time (cf. (9)), are at work in the semantic extensions from the nominal TOKORO denoting a 'point in space' (cf. Lakoff 1992). Indeed, many have considered the notion of 'point' to be central and common in these expressions. However, in this paper I will not try to justify such a claim nor try to use the theory of metaphor to account for all the TOKORO expressions.

10 The TOKORO construction differs formally from the TO construction in requiring the verb in the S1 to be in the -ta ending whereas in the TO construction the S1 verb is in the -ru ending:

(i.a) TOKORO construction
      =4a tutumi o aketa TOKORO, ningyoo ga haitteita.
         -ASP

(i.b) TO construction
      tutumi o akeru TO, ningyoo ga haitteita.
         -ASP
      'When I opened the wrapping, a doll was inside.'

11 As can be seen from the literal translation of (21c), it is common for S2 of the temporal clause-linking TOKORO sentences not to explicitly encode the experiencer and the act of cognition (i.e. 'a doll was inside' rather than 'I found that a doll was inside'). This may in fact be a general tendency in Japanese. By examining English translations of Japanese narratives, Teramura and Ohori have independently discovered a similar contrast between Japanese and English concerning encodings of acts of cognition. Although English also allows the experiencer and the verb of cognition to be left out, Teramura and Ohori have found many cases in which the Japanese original sentence does not mention the experiencer of cognition but its English translation encodes the act of cognition through the use of a cognition verb with an experiencer subject (Ohori To Appear, Teramura 1984).
12 When the agent of S1 is distinct from the speaker, sentences such as (21c) are felicitous only when S2 describes the speaker's point of view and NOT that of the agent of S1.

13 Eve Sweetser has pointed out that where in English has a use parallel to this use of TOKORO:

(i.a) I read (in the paper) where he is going to resign.
(i.b) I heard where it's supposed to flood again.

In these sentences, where is syntactically a complementizer comparable to that, but semantically it is used as a mental-space builder denoting a mental-state location.

REFERENCES


Flat vs. branching morphological structures: the case of suspended affixation
Cemil Orhan Orgun
U.C. Berkeley

1. INTRODUCTION. This paper presents an analysis of an unusual construction in Turkish called Suspended Affixation by Lewis (1967). The analysis has several theoretical implications. First, a principled account requires a contrast between branching and flat nodes in constituent structure. Second, the constituent structure motivated by Suspended Affixation differs from the one needed for the enforcement of a disyllabic minimal size condition. This suggests that morphological and phonological structures do not always match each other perfectly.

2. SUSPENDED AFFIXATION
2.1 DATA. Suspended Affixation in Turkish has been described by Lewis (1967: 35) as a construction in which “one grammatical ending serves two or more parallel words.”

An example is shown in (1a), where the two nouns sihhat ‘health’ and a:fiyet ‘well being’ are conjoined, and the locative suffix, which has scope over both of them, is found only once at the end of the conjoined phrase instead of on both conjuncts. Example (1b) is similar. Further examples can be found in Lewis (1967), Underhill (1976), and Inkelas and Orgun (1994).

(1) a) sihhat ve a:fiyet-te health and well-being-LOC
     b) halk-in [a:fiyet-ler-i] people-GEN sorrow and joy-PL-POSS
     ‘In health and well-being’ ‘the people’s sorrows and joys’

Example (2) shows the null hypothesis for the structure of this construction. I assume that the constituent structure is as implied by the scope relations, with the locative suffix is attached to the whole conjoined NP. At this point, we make this assumption as a notational convenience to use in presenting examples of Suspended Affixation. We will later see that this assumption is justified. Note that I use constituent structure notation solely because of its visual appeal and that the proposal I make is consistent with realizational views of morphology (e.g. Anderson (1992)). The analogue of flat structures is realizational “constructions” referring to a combination of features rather than just one feature.

(2) [ sihhat ve a:fiyet ] te

There are initially puzzling restrictions on the combinations of affixes that Suspended Affixation can target. As seen in (3a), it is possible to suspend ALL eligible affixes. Here, the plural suffix -ler, the possessive -im, and the accusative suffix -i are all suspended. Example (3b) shows that it is acceptable not to suspend any affixes at all. Here, all suffixes are realized on both conjuncts.
(3) a) All affixes suspended: [tebrik ve teşekkûr]-ler-im-i
    [congratulation & thank]-PL-1SGPOSS-ACC
    'my congratulations and thanks (acc)'

    b) No affixes suspended: tebrik-ler-im-i ve teşekkûr-ler-im-i

Example (4) shows the promised puzzling restrictions on Suspended Affixation. In (4a), we see that it is possible to suspend just the accusative suffix -i while realizing the plural and possessive suffixes on both conjuncts. Example (4b) shows that it is NOT possible to realize the plural suffix -ler on both conjuncts while suspending the possessive and accusative suffixes.

(4) Suspension of some but not all affixes (new data; author's judgments; confirmed by three additional native speakers).
   a) [tebrik-ler-im ve teşekkûr-ler-im]-i
   b) *[tebrik-ler ve teşekkûr-ler]-im-i

Our task is to account for this inseparability of the plural and possessive suffixes in Suspended Affixation. That is, we need to find a formal account of the observation that the plural and possessive suffixes are either both realized on all conjuncts or both suspended.

2.2 ANALYSIS. I offer an analysis of this seemingly strange restriction in terms of constituent structure. I claim that the plural and possessive suffixes form a flat (ternary branching) structure with the base they attach to, as shown in (5b), rather than a binary branching hierarchical structure as in (5a).

(5) a) *

        N
       /\          \       
      N   N       im
     tebrik ler

    b) N
       /\  
      N   N
     tebrik ler im

This ternary branching structure is supported by the three way ambiguity of third person plural possessive forms, shown in (6).

(6) it-ler-i
    dog-PL-POSS

    a) 'her/his dogs'
    b) 'their dog'
    c) 'their dogs'

The plural suffix can be interpreted to indicate that the head noun is plural, that the possessor is plural, or both. Since neither affix is required to have scope over the other, it is reasonable to conclude that the plural and possessive suffixes are not in an asymmetric c-command relationship, following the common assumption that c-command implies scope.

An alternative (but incorrect) approach would be to propose that there are two different hierarchical structures corresponding to the two different scope relations possible. The structure of 'her/his dog' would then be as shown in (7a).
Here, the possessive suffix has scope over the head+plural combination. The head is thus taken to be plural. In (7b), the proposed structure of ‘their dog’, the plural and possessive suffixes form a constituent. Thus, neither suffix c-commands the head noun. Rather, the node dominating the plural and possessive suffixes c-commands the head. Therefore, the possessor is interpreted to be plural. So far, there are no problems. However, this account would lead us to expect that ‘their dogs’ would have the structure in (7c), where both the possessor and the head are pluralized. Of course, this form is ungrammatical. Instead, we get the three-way ambiguity described in (6). This supports the flat structure that I am positing, which accounts for the observed “repeated morph constraint” effect in a straightforward and intuitive way.

(7) Incorrect approach: “Scope follows from structure.”

a) *  
\[ it \quad ler \quad i \quad \] 
‘her/his dogs’

b) *  
\[ it \quad ler \quad i \quad \] 
‘their dog’

c) *  
\[ *it \quad ler \quad ler \quad i \quad \] 
Intended: ‘their dogs’

We have thus seen that the plural and possessive suffixes have to be sisters whenever they are both present (5b). Given that the plural and possessive suffixes form a ternary branching structure with the base they attach to, the pattern of suspension in (8) is ungrammatical because it forces the plural and possessive suffixes to be in a hierarchical structure. This example is similar to the one we have seen before in (4b), except that the accusative suffix is not involved here. This further supports the position that the source of the problem is the configuration of the plural and possessive suffixes. There are two possible structures for this form. The first is shown in (8a). Here, the possessive suffix is attached to the conjoined NP, as it has scope over both conjuncts. This configuration violates the condition that the plural and possessive suffixes must be sisters whenever they both have scope over the same head. Therefore, this structure is ruled out. This leaves us with the possibility in (8b), which is structurally well formed. However, this structure does not give us the desired scope relations. In particular, the possessive suffix has scope over the second conjunct but not the first conjunct. Therefore, we explain the fact that the plural and possessive suffixes have to be suspended together, or not suspended at all.

(8) *[tebrik-ler ve teşekkür-ler]-im  
[thank-PL & congratulation-PL ]-1SGPOSS

a) *tebrik \quad ler \quad ve \quad teşekkür \quad (ler \quad im)  
Problem: -ler and -im not sisters

b) tebrik \quad ler \quad ve \quad teşekkür \quad ler \quad im  
Problem: Incorrect scope (-im)
In general, then, suffixes can be separated in Suspended Affixation only if they form a hierarchical structure. If they form a flat structure, they have to be suspended as a group, or not at all.

3. MINIMAL SIZE CONDITION. We have seen that the proposed contrast between flat and branching structures accounts for Suspended Affixation in a simple manner. Now, I will show that this contrast finds further justification in the differential enforcement of a disyllabic minimal size condition. As noticed by Ito and Hankamer (1989), and later studied by Orgun and Inkelas (1992), certain derived forms in Turkish have to contain at least two syllables (9).

Example (9a) shows monomorphemic forms that consist of a single syllable. These forms are acceptable. Example (9b) shows the same forms with consonantal suffixes. Here, we see that the forms are ungrammatical if they contain only one syllable. The last two forms in (9b) confirm this observation: These forms are grammatical because they contain two syllables.

(9) Disyllabic minimal size condition

<table>
<thead>
<tr>
<th>a) Monomorphemic forms</th>
<th>b) Affixed forms (σσ minimum)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ye ‘eat!’</td>
<td>*ye-n ‘eat-PASS!’</td>
</tr>
<tr>
<td>do: ‘musical note do’</td>
<td>*do:-m ‘do-1SGPOSS’</td>
</tr>
<tr>
<td>yut ‘swallow!’</td>
<td>yut-ul ‘swallow-PASS’</td>
</tr>
<tr>
<td>sol ‘musical note sol’</td>
<td>sol-ıüm ‘sol-1SGPOSS’</td>
</tr>
</tbody>
</table>

Orgun and Inkelas noticed a curious fact concerning the possibility of repairing subminimal forms by adding further suffixes. The idea is that one should be able to add more suffixes to subminimal forms to bring the total size to two syllables. The resulting form should then be grammatical.

However, we see that this can only be done for some, not all, of the subminimal forms in (9). In (10b), we see that it is possible to repair the form *ye-n ‘eat-PASS’ by adding an aspect suffix to bring the total size to two syllables. The resulting form, ye-n-ir ‘eat-PASS-IMPRF’ is grammatical. However, (10a) shows that the subminimal possessed form *do:-m ‘do-1SGPOSS’ remains ungrammatical even when we add the accusative suffix to bring the total size to two syllables.

(10) a) *do:-m ‘do-1SGPOSS’   b) *ye-n ‘eat-PASS’
     *do:-m-ı ‘do-1SGPOSS-ACC’ ye-n-ir ‘eat-PASS-IMPRF’

The forms in (10a) bring to mind the issue of cyclicity, or the application of phonology to the output of each morphological operation. In the framework of Lexical Phonology (Pesetsky 1979, Kiparsky 1982, 1985, Mohanan 1982)), the output of each morphological concatenation is submitted to the phonological module, where it is subject to phonological rules and constraints.

With this framework in mind, the Turkish minimal size condition seems to apply cyclically in (10a), as shown more clearly in (11a). Here, the
root+possessive combination is subject to the minimal size condition. It violates the condition, and is therefore ruled out. This appears to be a typical case of cyclicity. However, (10b) suggests that the condition is enforced noncyclically, as in (11b). Here, the root+passive combination is NOT subject to the minimal size condition. Instead, the whole form ye-n-ir ‘eat-PASS-IMPRF’ is subject to the condition only once, after all the suffixes have been added. Since this form contains two syllables, it satisfies the minimal size condition and is grammatical.

(11) The minimal size condition is sometimes enforced cyclically
   a)   do: + m → *do:m   (minimality enforced and violated)

   ... and sometimes noncyclically:
   b)   ye + n   (minimality NOT enforced)
        ye + n + ir → yenir   (minimality enforced and satisfied)

Of course, the form *ye-n ‘eat-PASS’ would have been subject to the condition if it were to be used as a word by itself, and would have been ruled out because it is subminimal.

This contrast between cyclic and noncyclic application of phonological rules or constraints has been noticed in the Lexical Phonology literature. The traditional approach to this problem uses a diacritic feature [+cyclic] (e.g. Halle and Mohanan (1985), Halle and Vergnaud (1987), Czaykowska-Higgins (1993)). In this approach, the output of a given morphological operation is normally not subject to phonological rules or constraints before the word level. However, when an affix bearing the feature [+cyclic] is added, the resulting form is submitted to the phonological module, where it is subject to phonological rules and constraints. That is, we get cyclic phonology only when [+cyclic] affixes are added to a base.

This approach could certainly be used in analyzing the Turkish minimality facts. The passive suffix would not bear the [+cyclic] feature, while the possessive suffix would. This way, forms containing the possessive suffix would be subject to minimality enforcement before the addition of any further suffix.

However, it is in fact possible to DERIVE the cyclic versus noncyclic application of the minimal size condition rather than stipulating it in this fashion. We do not need any additional tools or ad hoc diacritic features like [+cyclic] to achieve this. Constituent structure, a tool we have already used in the analysis of Suspended Affixation, is all we need.

I claim that every nonterminal node in a phrase structure is subject to the minimal size condition (excluding nonbranching (unary branching) nodes in frameworks that allow them), and that the root+passive+aspect combination forms a ternary branching structure (12a), while the root+possessive+case combination forms a binary branching hierarchical structure (12b).
(12) Every nonterminal node is subject to the disyllabic size condition.
   a) \[ \text{V} \]
   \[ \text{ye} \quad n \quad \text{ir} \]
   b) \[ *\text{N} \]
   \[ \text{do:} \quad m \quad u \]

There is only one nonterminal node in (12a)—the node labeled V. The contents of this node equal two syllables. Therefore, this form satisfies the minimal size condition and is grammatical. In (12b), on the other hand, there are two nonterminal nodes. The lower one, labeled \(*\text{N}\), violates the minimal size condition, because its contents equal only one syllable. This form is therefore ungrammatical, regardless of what may be added to it. This is because any additional suffix is not going to add to the contents of this particular node. Thus, we derive the apparent cyclic versus noncyclic application of the minimal size condition from static phrase structure configurations, without any diacritics (cf. Cole and Coleman (1993)).

Support for this analysis comes from a perhaps unexpected source. Recall that the possessive and case suffixes may be separated from each other in Suspended Affixation. This is shown in (13) (repeated from (4a)).

(13) a) \[ \text{[tebrik-ler-im ve teşekkür-ler-im]-i} \]
     \[ \text{[congratulation-PL-1SGPOSS \& thank-PL-1SGPOSS]-ACC} \]
   b) \[ \text{teşekkür} \quad \text{ler} \quad \text{im} \quad \text{i} \]
     \[ \text{thank} \quad \text{PL} \quad 1\text{SGPOSS ACC} \]

We have seen that it is acceptable to keep the possessive suffix on both conjuncts while suspending the case suffix, as shown again in (13a). This implies that that the possessive and case structures must NOT be sisters to each other. In other words, they do NOT form a ternary branching structure with the base they attach to. Rather, they form a binary branching hierarchical structure as shown in (13b). This is, of course, precisely the conclusion we have already arrived at on the basis of minimality enforcement. This agreement between the structures needed for Suspended Affixation and minimality verifies the analysis I am proposing.

By contrast, an approach using exclusively binary branching (e.g. Williams (1981), Kayne (1983), Larson (1988), Baker (1988), Lieber (1992)) would have to STIPULATE the cyclic-noncyclic distinction (e.g. Halle and Mohanan (1985), Halle and Vergnaud (1987), Czyzewska-Higgins (1993)), would lack an account of Suspended Affixation, and could not relate the phonological and morphological phenomena to each other.

Even if one somehow managed to devise an account of Suspended Affixation in such a framework, this would still not be related to the cyclic versus noncyclic enforcement of the minimal size condition or to the suffixal scope ambiguity in third person plural possessive forms. Only an approach that allows
flat as well as binary branching structures is able to offer an insightful analysis of all these phenomena.

4. **INTERACTION OF SUSPENDED AFFIXATION WITH MINIMALITY.** The interaction of Suspended Affixation with minimality reveals an apparent bracketing paradox. As shown in (14), the last conjunct, together with the suffixes at the end, forms a phonological domain which is subject to minimality. This is true even though the suffixes have scope over both conjuncts. The form in (14a) is ungrammatical because the last conjunct (plus the suffixes at the end) contains only one syllable, and thus violates the minimal size condition. In (14b), on the other hand, the last conjunct and the suspended suffix together add up to two syllables. The minimal size condition is thus satisfied, and the form is grammatical.

(14) a. \([do\ ve\ re]-m\)  
\([do\ and\ re]-1SGPOSS\)  
‘my do and re’  
\(*[re-m]\) violates \(\sigma\) requirement

b. \([fa\ ve\ sol]-\ddag\)  
\([fa\ and\ sol]-1SGPOSS\)  
my fa and sol’  
\([sol-\ddag]\) satisfies \(\sigma\) requirement

Thus, the suspended suffixes form a phonological unit with the last conjunct even though they have scope over both conjuncts, and are attached morphosyntactically to the whole conjoined phrase. Morphosyntactic considerations based on scope and Suspended Affixation lead us to conclude that the structure is as shown in (15a). On the other hand, prosodic minimality enforcement suggests the structure in (15b).

(15) a) Morphosyntactic structure (based on Suspended Affixation)  
b) Phonological structure (based on minimality)

There are two common approaches to bracketing mismatches, which I call the “edge feature approach” and the “dual structure approach”.

The edge feature approach assumes that the structure implied by the phonology is the correct one. Certain features and edge feature percolation mechanisms then allow the morphological structure to match the phonological structure by letting the presence of a morpheme on a lower node at a designated edge indirectly influence the features borne by a higher node. This is the approach taken by Poser (1985), Zwicky (1987), Halpern and Miller (1991), Halpern (1992), Miller (1992), Miller and Halpern (1992), etc. The particular mechanism assumed in these works is usually a realizational approach to morphology. In this model, the presence of features on certain nodes triggers the phonological expression of affixes. In Suspended Affixation, the proper version of the edge feature convention would have to state that the relevant features are present on both conjuncts, but it is an option to realize them phonologically just on the
second conjunct.

(16)  Edge feature approach.

\[
\begin{array}{c}
\text{[possessed]} \\
\text{do}
\end{array}
\begin{array}{c}
\text{[possessed]} \\
\text{sol}
\end{array}
\text{ ve } 
\begin{array}{c}
\text{üm}
\end{array}
\]

It may be possible to formalize this approach to make it possible to account for the simple cases of Suspended Affixation where a single suffix is eligible for suspension. However, when we consider the full range of facts, we realize that the edge feature approach is hard put to answer the crucial questions in (17).

(17)  Challenges for the edge feature approach:

a) why can only nonfinal conjuncts have missing affixes?
   \[*\text{tebrik-ler-im-i ve teşekkür}\]

b) why can only the outermost suffixes can be suspended?
   \[*\text{tebri-i ve teşekkür-ler-im-i}\]

c) why do suffixes of a flat node have to be suspended en masse?
   \[*\text{tebrik-ler ve teşekkür-ler-im-i}\]

At this point, I do not know if it is possible to extend the edge feature to account for these observations. In addition, we would demand that such an account explain the connection between suffix suspension, minimality enforcement, and suffixal scope ambiguity. I leave the resolution of this question to future research.

The second common approach to bracketing paradoxes of this sort is to posit a dual structure. In this approach, both of the structures that we have motivated are assumed to be present. They are parallel descriptions of the same object on independent phonological and morphological levels of description. This is the approach taken by Marantz (1988), Sproat (1988), Inkelas (1989), Cohn (1989), Zsiga (1992) (cf. Sadock (1985, 1991)), etc. These two structures are identical by default. However, specific grammatical requirements may force mismatches between the phonological and morphosyntactic structure as needed.

(18)  Dual structure approach.

\[
\begin{array}{c}
\text{m-structure}
\end{array}
\begin{array}{c}
\text{do}
\end{array}
\begin{array}{c}
\text{ve}
\end{array}
\begin{array}{c}
\text{sol}
\end{array}
\begin{array}{c}
\text{üm}
\end{array}
\]

The success of this kind of approach depends crucially on how strong a theory of allowable types of mismatches one has. This issue is as yet unresolved, since we do not at this point know exactly what kinds of mismatches are motivated by empirical evidence. However, two points deserve mention. First, Orgun (1994) has proposed a view of dual structures according to which mismatches are only allowed if a specific morpheme has inconsistent phonological and
morphosyntactic attachment requirements. Suspended Affixation satisfies this
description: The relevant suffixes attach to a phonological word. On the
morphosyntactic side, however, they may attach to a lexical noun (which gives us
the no suspension option), or to an NP. In the latter case, their phonological host
will, in general, differ from their morphosyntactic host.

The second point of relevance is a study by Zsiga (1992). Zsiga claimed
that when there is a morphosyntactic constituent containing more than one
phonological word, affixes (and clitics) that are added to this phrase will
phonologically attach to the edgemost phonological word. This claim, taken
together with Sproat’s (1986) observation that conjoined phrases tend, cross-
linguistically, to contain multiple phonological words, leads us to expect that
Suspended Affixation should give rise to such mismatches. In this light, the
conclusions we have reached in this paper regarding the differences between the
morphosyntactic and phonological properties of Suspended Affixation should not
be surprising.

5. CONCLUSIONS. Regardless of whether we use the edge feature approach or the
dual structure approach, the conclusions summarized below are valid.

Once flat (as opposed to exclusively binary branching) structures are
allowed, cyclic versus cyclic phonological effects follow directly from
independently needed constituent structure configurations. It is not necessary to
introduce ad-hoc diacritic features to encode the cyclic-noncyclic distinction.
Suspended Affixation further supports this view by showing that in Turkish,
affixes can only be suspended independently of each other if they belong to
different “cycles” (i.e. only if they are not sisters). Evidence from disyllabic
minimality, suffixal scope ambiguity, and Suspended Affixation converge on the
same flat versus hierarchical configurations.

Finally, a methodological conclusion may be drawn from this study. Many
linguists distinguish in practice (some even in theory) between “core” and
“peripheral” linguistic phenomena, designing their theories around the former and
often neglecting the latter. The implicit (sometimes explicit) hope is that the core
phenomena fully characterize the system underlying the language and that no
additional information may be gained from the study of marginal phenomena.
Suspended affixation and derived monosyllabic stems would not generally be
considered core phenomena. However, we have seen that these phenomena
provide crucial insight into the structure of Turkish and, by extension, into the
properties of Natural Language.

ACKNOWLEDGMENTS
I thank the following people for discussing Suspended Affixation with me: Paul
Kay, Jean-Pierre Koenig, Collin Baker, Dan Jurafsky, John McCarthy, David
Perlmutter, Sharon Inkelas, Gün Orgun, Güler Orgun, Günel Orgun, Larry Hyman.
REFERENCES

Cole, Jennifer and John Coleman. 1993. Cyclic phonology with context-free grammars. CLS.
NLLT, 197-228.
Miller, Philip and Aaron Halpern. 1992. English possessives and the syntax of morphological features. FLSM III.
Orgun, Orhan and Sharon Inkelas. 1992. Turkish prosodic minimality. Presented at the 6th International Conference on Turkish Linguistics, Anadolu University, Eskişehir, Turkey.
Pesetsky, David. 1979. Russian morphology and lexical theory. Manuscript, MIT.
1.1. Grammatical categories characterized. Grammatical categories -- the central theoretical concept of linguistics as an explanatory science -- are to be distinguished from categories of relative Saussurean distribution (or "formal categories" of "autonomous" grammar) and from structured denotational and other discursive distinctions (specifically and differentially communicated or signaled semantic and pragmatic content). Grammatical categories are the categories of mapping between such formal categories, definable under structures of morphological and syntactic combinatorics, and such communicable differentiations of denotational and other information as underlie modalized propositional text (or even the textuality of communicative interaction, more ambitiously).

Once we distinguish grammatical categories in principle from formal (distributional) categories, we can see that the structural ingenuity of languages universally is that some grammatical categories are, in fact, transparently coded in isolable paradigms of surface-segmentable (Bloomfield: "lexical") forms, and others are rather complexly coded in morphosyntactic configurations of concurrent paradigmatic forms in particular languages; and some potential grammatical categories, coded transparently or configurationally in some language(s), are not regularly or productively or other-than-collocationally coded in a particular language at all, though the semantic or pragmatic distinction can be communicated with the proper lexical items arranged in phrase-types of a sufficient complexity.

Many of the more fundamental grammatical categories are expressed only in paradigms of configurational cooccurrence of other, more transparent grammatical categories brought together in a formal structure of a certain scope (constituting a particular distributional category-type, e.g., a phrase or clause of a particular type). Such is, for example, the situation with VALENCE, comprehending the coding of a predicate's number-of and configuration-of arguments, with respect to which classes of predicates can be defined. VALENCE is formally expressed only through the entire systematics of clause inflectional morphosyntax under all possible conditions of CLAUSE-LINKAGE in sentence-scope grammatical forms.

The typical statement of a grammatical-categorial regularity has the form of a complex function that maps structures of values of multiple variables (the quantized dimensions of semantic and pragmatic differentiation) into specific formal-distributitional categorial values. There is a sense in which, then, the unique or principal or recurrent (note the degree concept) functional (mapped) associability of certain formal categories with some particular structure of semantic or pragmatic differentiation, no matter how many other variables may be involved, is the basis on which we can label some formal-distributitional category by the associated semantic or pragmatic value. That is, we have linked distribution to denotation through grammatical categorization. CASE-marking, for example, is justifiably
called this--stipulatively labeled as about "case," that is--for three reasons: [1] because it is anchored in universal and particular stipulations involving predicate-argument VALENCEs, indicating the thematic roles (or case-relations) of particular denotata about which a state-of-affairs is being predicated, whatever other semantic and pragmatic variables, in particular formal constructions and/or in particular languages, may go into determinations of these diagnostic formal-distributational regularities; and [2] because any other variables involved in the CASE-marking system are themselves more transparently and paradigmatically coded in formal-distributational structures that do not, inherently, involve the predicate role of a designatum as argument; while [3] the putative CASE-coding is the most transparently and paradigmatically coded formal-distributational structure that maps into this structure of interlocked variables, under our formal account of such a mapping. (We might call these the [1] essentiality, [2] uniqueness, and [3] minimaximality conditions on transparency of coding, as theoretically-derived methodological evaluation criteria on grammatical-categorial analysis.)

1.2. The empirical character of grammatical-categorial description: predictiveness. There are three areas of linguistic function and transformation that begin to give us a sense of the nature of the relationship between typological compatibility-with (or at least "free-ride" non-incompatibility-with) a specific grammatical-categorial universal, and a language's specific conformity-to (or formal exemplification of) it.

First, in a given language that is merely compatible with a typological universal of grammatical categorization, we can generally demonstrate statistical tendencies of formal-distributational coding that, in numerical, if not categorical terms, show a kind of non-structural but nevertheless demonstrable conformity-to the parametric structure of the universal. This means that in a certain sense the universal still inhabits the specific language and its users, though we must differentiate here as sharply as we might structural (competence; langue) and statistical (textual; parole) orders of factuality.

For example, in English case-marking, note that the denotational content space of various semantic-pragmatic types of arguments intersects only marginally with VALENCE in determining distributional form. (The personal deictics have special accusative morphological case-markings for a variety of construction-types; otherwise Noun Phrase constituent-distribution relative to construction-type is the means of coding Nominative vs. Accusative/Dative/etc. case.) Yet, in experiment after experiment that investigated relations of the denotational (and pragmatic) content of Noun Phrases to communicated or likely-perceived argument-role, there are tendencies for Nouns to fall into numerically differentiated types according to what we might term "modal associability with" "Agentive Subject and Patientive Object, falling down the cline dictated by the denotational content space. Textual token-counts as well bear out these conforming tendencies.

Second, intersecting grammatical-categorial dimensions comprise a space of parametric orderings that map into formal-distributational coding regularities, in ways that define modal and extreme intersections of such dimensions relative to a specific language's formal categorial system. At the extremes of particular dimensions, the formal expression should be relatively rarely encountered (in structural terms, not only text-counts), and should be especially formally elaborated as compared with modal and less densely categorial intersections. Thus note that so-called "pure" ergatively-oriented case-marking systems with specifically ergative case-marking
across the entire space of intersecting variables of VALENCE, etc. are vanishingly rare of occurrence. 3

Third, in an area my extended Kiksh example here will illustrate, facts of the way a language both conforms to and is compatible with a set of grammatical-categorial universals determine an **analogical space** in which both the synchronic and diachronic analogical tendencies of users of a language operate. Grammatical change is propagated and spreads in analogical space, that is, the space of susceptibility of spread of innovations through a grammar, once they have a hold in statistical compatibility-space and in structural ("markedness" or conformity-space. The classical theory of so-called analogy, that is, of mechanisms of force or influence of one grammatical-categorial-coding construction type upon another, of course presupposes an analysis of linguistic structure in just such grammatical-categorial terms. 4 We are now able to articulate such a theory of the rise and spread of newly distinctive formal-categorial structure as a filling out of tendencies already present in the way that a particular formal-categorial structure is compatible with and conforming to the "potential" (Mathesius) of a set of grammatical-categorial universals that define a field of analogical pressure.

This can be illustrated with the rise of a distinctive purposive clause construction in Kiksh (Wasco-Wishram) and its particular use of what are otherwise non-cross-referencing "impersonal" inflections as bound anaphors. Let us look to the details of the process. We take up first the grammatical structure in which "impersonals" can be located, and then show the change that can be documented to speakers born in the last quarter of the 19th century.

**2.1. Ergative Impersonals in pronominal inflection.** Kiksh 5 (or Wasco-Wishram Chinookan) inflectional syntax is consistently head-marking (Nichols 1986:57-8) in formal-distributional surface organization. It uses cross-referencing inflectional pronouns (categorized for person, number, and gender agreement) prefixed to the head word to signal the functional syntactic relations of constituents within lexically headed phrases. Thus, note the chart in (1), which lays out the numbered maximal order-classes of an independent-clause finite verb. We can see that up to three cross-referencing pronouns can occur -- in sequence, in the order-classes ergative2, absolutive/nominative3, and dative/locative4, these three, for example, fully inflecting a verb of the 'give' **valence-class**. Of course, not every verb shows all of the three pronominal inflections; most occurring verbs have a combination of one or two in basic form.

(1) Morphological structure of Kiksh verbs:

<table>
<thead>
<tr>
<th>Tense</th>
<th>Ergative</th>
<th>Absolutive</th>
<th>Datative</th>
<th>Post-positional</th>
<th>Directive</th>
<th>/ROOT</th>
<th>Causative/Passive</th>
<th>Conative/ /ROOT</th>
<th>Motive</th>
<th>Aspect/Voice</th>
<th>Deixis</th>
</tr>
</thead>
</table>

[Lexical Stem]

[Stem Derivation]

[Indir. Adjunct]

[Cross-referencing]
Exactly what pronominal forms occur in the various possible inflectional sets of the different valence-classes is, however, a complex function of both a valence-class schema and of what I have termed (Silverstein 1976:124-5) a "global" (as opposed to "local") case-marking schema of noun-phrase denotational content: for a given valence-schema of arguments, determined by verbal valence-class, the categorially-differentiated denotational content of at least two (as opposed to only one) argument-coding verb adjuncts (as well as valence-class) determines the surface inflectional form, here the form of at least one of up to three cross-referencing pronouns. (2) exemplifies this.

(2) (a) i₁-m₂-sh₃-(a)n₄-l₅-u₆/t₇ 'you(sg)₂ just₁ gave₇ them(du)₃ to₅ me₄'₆

(b) i₁-sh₃-(a)m₄-l₅-u₆/t₇ [I₂] just₁ gave₇ them(du)₃ to₅ you(sg)₄

Note in (2a) that the "direct" order of -m₂-sh₃-(a)n₄-l₅- '2sgAgt-3duPat-1sgDat-to' codes all of the prononominals as expected for 'you (sg) [A] ... them (du) [P] to me [D]', while the "inverse" order of (2b) shows -sh₃-(a)m₄-l₅- '1sgAgt-3duPat-2sgDat-to' as the coding for 'I [A] ... them (du) [P] to you (sg) [D]'. Observe that (2b) shows obligatory non-expression of the first person singular 'Agent' when the 'Dative' is second person; the surface morphology shows no segment identifiable as first singular. So the expression of predicate-argument relations maps determinately into the dependent variables of pronominal cross-referencing, within a "space" of two independent variables of grammatical categorization, verbal-predicate valence-class crossing local and global denotational content of argument-adjuncts.

Because of the various factors that contribute to the inflectional schemata of cross-reference, partly "local" and partly "global" in nature, each coherent region of the space of grammatical categories of denotational content seems to have its own particular morphological paradigm of formal case-marking, as shown in (3) [adapted from Silverstein 1976:141]. At the extreme leftmost end of the representation of denotational types is the 'second person singular', which is invariant in form (always -m-) and whose grammatical coding as cross-referencing pronominal is simply a function of its position of occurrence relative to others in the whole set of inflectional pronouns: if in first relative position, it is always "Subject"; if in non-first relative position, it is always non-"Subject"; the 2SG GRAMMATICAL CATEGORY is invariably "accusative" in case-marking compatibility. This is true for 'first person singular' as well, except when a 'second person' serves as a thematically lower argument of a predicate valence-schema, in which case the 1SG has a special, "ergative"-like form, namely zero [cf. (2b) above].

In (3), a linearization of Kiksht NP-category types based on the multidimensional ordering of denotational content space theorized in Silverstein 1976; 1977; 1985, the correspondence of content types and case-marking schemata are laid out in the vertical dimension. As (3) shows, coherently definable regions of the grammatical category space of denotational content -- what can be called 'natural classes' in grammatical-categorial terms -- show similarly coherent case-marking regularities, in conformity with the predictions of how systems range between "accusativity" and "ergativity" in their surface distinctions.
(3) Kiksht splits of case-marking alignments:

```
\begin{array}{cccccccccccc}
| 2sg | 1sg | 1ex | 1ex | inc | inc | 2du | 2pl | lex | lex | 3du | 3pl | 3nt | 3sg | 3sg | imp |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
\end{array}
```


with 2 '0,D' arg.

ACCUSATIVE GLOBAL ERGATIVE ............THREE-WAY A:S:0............ 3-WAY A:O:S LOCAL ERGATIVE ERG

At the extreme rightmost end of the representation of denotational types is the so-called *impersonal* category, represented by the unique form q2-. This representation with fixed subscript '2' indicates that the impersonal form occurs ONLY in ergative2 order-class as a pronominal, thus having a "defective" pronominal paradigm of only one member. In its basic impersonal usage, it does not cross-reference the nominal head of any headed phrasal projection of a dependent argument, as do pronouns agreeing with full, nominally-headed noun phrases. Denotationally, its impersonal character means that while it presupposes that there is/are some quantifiable entity or entities -- generally human and stereotypically agentive -- the descriptive backing for referring to them (cf. Searle 1969:80-91) is purely in terms of the predicature of which the suggested person or persons unknown function(s) as an argument. It is a definite, but nonspecific reference spanning the canonical Donnellan (1971:102) "attributive" reference conditions7 and our English impersonals in they and derivatives, e.g., the paranoid's classic "They are out to get me!" As in English or French or Russian, therefore, as well as in Menomini and other Algonquian languages, the "impersonal" forms in q2- function as 'Voice'-like in discursive implicature, since for a known 'Patient' (normally coded in absolutive3 order-class) or known 'Recipient/Benefactee/Experiencer/Locus' etc. (normally coded in dative4 order-class), a predicature with a definite but nonspecific 'Agent' (coded by q2-) codes as much information about a predictable event or state of affairs as a bipartite or "agentless" 'Passive Voice' category.8 Thus note example (4) from a war narrative told to Edward Sapir (1909:210.12ff.) in 1905 by Louis Simpson, born ca. 1830, and describing events of ca. 1867 in which Wasco and Wishram people were participating in hostilities with Paiute people.

(4) Agha k\textup{\textasciitilde}wapt ga\textup{\textasciitilde}1-q2-nch\textup{\textasciitilde}3-u6-/[k]\textup{\textasciitilde}lhh7-xh. Ga\textup{\textasciitilde}1-nch\textup{\textasciitilde}3-u6-/i7-[i]x iaghailh

'Now then we3-were-taken (=lead). We3-went-(to a) large lake (where wilhala adatilhx i[t]\textup{\textasciitilde}wanxhayuksh. Nawit ilhk'alarmat\#pa ga\textup{\textasciitilde}1-q2- nch\textup{\textasciitilde}3 -were) many Paiutes. Right up-to#a-bridge we3-were-lead-over-to.

-u6-/klhh7-am10-axh. Agha k\textup{\textasciitilde}wapt lhh'up ga\textup{\textasciitilde}1-q2- sh3-nch\textup{\textasciitilde}4-l5-u6-/xh7-axh. Now then shoot-at we3-were-done-to.

Agha k\textup{\textasciitilde}wapt ga\textup{\textasciitilde}1-q2-nch\textup{\textasciitilde}3-u6-/pghna7-iwg-(a)nang-(m)xh ...
Mr Simpson describes a whole series of events in his chronological account with both intransitive and transitive finite verb forms in which the 'First Plural Exclusive' pronominal -nch3,4- occurs in absolutive3 or dative4 order-class. Observe how the statements coded in transitive and ditransitive forms describing things done to Mr Simpson and his Wasco mates are inflected with 'Impersonal' forms in ergative -q2-, which Sapir translated in the instance with bipartite ("agentless") English passive constructions preserved here.

There are two other kinds of occurrences that will locate the impersonal -q2- construction in grammatical-categorial space for the language of Mr Simpson's generation. One is the relation of impersonal forms to the two VOICE systems of finite verbs, and the second is the use of finite impersonal forms in subordinate, non-subject-pivoted relative clauses, making DEVERBATIVE "ADJUNCTIVES." Let us take these in turn.

2. Impersonals in relation to VOICE categories. From a morphological perspective, Kiksht has two distinct systems of distinguishing non-'Active' 'Voice'. One is a suffixal system of syncretic ASPECT x VOICE forms. These are lexically derived from roots so as to make stems of 'changes-of-state', 'accomplishments' and 'achievements' in the Vendlerian-Dowtyan understanding of "Aktionarten"; the predicative here is always conceptualized from the perspective of actional endpoint or its aftermath. Thus note in (5a) the RESULTATIVE PASSIVE in -it10 ("to have become V-ed"; "be V-ed") and in (5b) the TRANSITIONAL PASSIVE in -x8- -it10 ("to get-to-be V-ed"; "become V-ed").

(5) (a) ni1-d3-u6-/chxm7-it10 'they3 long-ago1 became/were10 boil7ed'
(b) ni1-d3-u6-/chxm7-x8-it10 'they3 long-ago1 got-to-be/became8+10 boil7ed'

Each of these forms shows valence-changing (reducing) effects on such 'Direct Transitive' verb stems as -[2-][3-u6-/chxm7- ']2 boil []3 (i.e., those coding a predicate schema 'p(A,P,...)'). At their level or layer of derivation, note, the passive voice suffixes are generally incompatible with full transitive inflectional cross-reference, even where a particular lexical template built from a root is defective and shows only a passive-like derivational form.9 Thus note that forms in -q2- (obligatorily ergative) impersonal inflection do not occur in the morphological space of stems and themes with passive x aspectual suffixation. So there is a complementarity of IMPERSONAL pronominal cross-reference and PASSIVE x TELIC/RESULTATIVE suffixal derivation in Kiksht that is not unlike what we see in even strongly "accusative" and "local" case-marking systems like that of English, where PASSIVE VOICE and NONSPECIFIC subject they also show complementarity.

By contrast, let us consider the two prefixal systems involving marked VOICE, both syntactically-productive and important for seeing the historical trajectory of IMPERSONAL usage. In (6) are shown the morphological formations of the three kinds of cross-referencing REFLEXIVES-RECIPROCATS, marking substantial referential identity (or non-distinctness) of at least two of the arguments of a predicate (cf. English -self anaphors), at the same time as they express the MIDDLE or MEDIOPASSIVE VOICE category.
(6) (a) **direct reflexive**: $i_1$-n$_3$+xh-lgl$_5$-/ga$_7$ 'I$_3$ just$_1$ grabbed$_5+7$ myself/mine'  
(b) **indirect reflexive**: $i_1$-sh$_3$-n$_4$+xh-l$_5$-u$_6$-/t$_7$ 'I$_4$ just$_1$ gave$_7$ myself/my$_4$ two-of-them$_3$'  
(c) **inverse reflexive**: $i_1$-n$_4$+xh-l$_5$-/lha$_7$ 'I$_4$ just$_1$ caught-a-whiff-of$_5+7$ myself$_4$'  

That is to say, in Kiksht the system of MIDDLE/MEDIOPASSIVE VOICE is used for intra-clause REFLEXIVE/RECIPROCAL anaphoric pronominalization, just as, conversely, in many "SAE" languages the formal-distributional reflexive and reciprocal pronoun systems are used to code MEDIOPASSIVE/MIDDLE VOICE and other VALENCE-reducing semantic effects. In each of the Kiksht examples, a form -xh- is postposed to the *rightmost* cross-referencing pronominal position coding a two-plus-argument valence category, (thus representing in order ergative$_2$ 'Agent', absolutive$_3$ 'Patient/Subject', dative$_4$ 'Recipient/Experiencer'), on which the 'Middle' relation of nondistinct argument is definable. What would appear in the unmarked ACTIVE as a more leftwardly pronominal element (or equivalent) is suppressed. Observe the range of types: there are 'Agent-Patient' MIDDLES, as in (6a), also termed "direct reflexives" after the so-called "direct" object; 'Agent-Dative' and 'Subject-Dative' MIDDLES, as in (6b), also termed "indirect reflexives" after the so-called "indirect" object; and 'Theme-Experiencer/Goal' middles, as in (6c), also termed "inverse reflexives" after the "inverse" subject coded by the dative$_4$ order-class pronominal. Of course the inflection of a form with MIDDLE VOICE involving the 'Agent' excludes any inflection with IMPERSONAL markers, just as it excludes other cross-referencing transitive subject ('Agent') pronominals which would be leftward of any others. Mediopassivization plays a central role in the morphology of verbs in subordinate clauses, however, as we will see below.

To the degree that Kiksht is substantially an "ergative" inflectional type of language, we expect it as well to show an ANTIPASSIVE VOICE, as indeed it does. This voice is marked by a prefix -ki/k'i- in the sixth order-class position, that replaces the otherwise occurring 'Directional' prefixes -u$_6$- 'distad' and -t$_6$- 'proximad'. Thus note in (7) that both antipassive forms show this characteristic morphological shift.

(7) **ANTIPASSIVES** of ga$_1$-lk$_2$-d$_3$-u$_6$-/dina$_7$-xh$_{10}$ 'long-ago$_1$ they(coll)$_2$ used-to$_{10}$ kill$_7$ them(pl)$_3$'  

(a) **finite**: ga$_1$-lh$_3$- -k'i$_6$-/dina$_7$-xh$_{10}$ 'long-ago$_1$ they(coll)$_3$ used-to$_{10}$ kill$_7$ [sc., many]'  
(b) **nominal**: it$_3$-lha$_4$-lk'i$_6$-/dina$_7$-xh$_{10}$ 'they-who$_4$ habitually-kill$_{7+10}$ them(pl)$_3$'
As seen in (7a), the antipassive marker, -ki6-, implies that no pronominal occur in a verb's absolutive3 order-class coding a 'Patient' argument; that is, in finite verbs, ANTIPASSIVE excludes cross-referencing expression of a 'Patient', fully in keeping with a universal asymmetry of voice categories across morphosyntactic types. In the finite antipassive form (7a) there is, to be sure, an absolutive3 pronominal -lh3- '[third person] neuter-collective', but it cross-references the nominal 'Agent' argument of the predicate, which is morphologically intransitive.

By contrast, as shown in (7b), in the derived nominalization based on the antipassive verbal inflectional stem, two pronominal-like elements are once again present. But here, the morphology of nominal inflection dictates that the "absolutive3"-like first pronominal is the nominal-stem's '[Person-]Number-Gender' prefix, while the "dative4"-like second pronominal is the cross-reference of the (here, obligatory) nominal 'Possessor'. Compare an English participial their4 habitually-killing them3 (where the subscripts translate the Kiksht thematic relations of the example). Notice that in Kiksht, too, the predicative 'Agent' of such a direct transitive coding is in specifically "adnominal dative4" or "genitive4" coding, while the predicative 'Patient' surfaces in the derived nominalization only as an "adnominal absolutive3," also termed the 'P/N/G' prefix of the noun. As will be seen below, it is the dative4 pronominal position, not the other one present in the nominalization of a verb coding a two-argument predicate, that is the "pivot," or site for anaphoric binding in complex sentences.

The semantic interpretation of a form with productive antipassive (that is, as opposed to its form without such prefix) is an "active intransitive" with genericized, not specifically denoted 'Patient' (by implicature, characterizing whatever category of things are presupposed to be usual and customary targets of an action). Therefore, such forms are frequently used for genericized, that is, habitual, event-implicature as well: "she (habitually) sells (things [sc., sea shells]) by the sea shore"; etc. Note that there is an overwhelming statistical association of ANTIPASSIVE in Kiksht with USITATIVE, CONTINUATIVE-SIMULTANEOUS (present), and ITERATIVE verb-stem formations, confirming this genericizing, generalizing usage. Thus note in (7) the antipassive forms in both (a) finite verbal and (b) deverbal nominalized variants of the construction, each with its aspectual suffixation for USITATIVE in -xh10.

2.3. Impersonals in adjective participial clauses. In the "polysynthetic" and "holophrastic" verbal morphology of Kiksht, as (1) above shows, the various 'Adverbial' types of relations, such as "location in/on/at/etc.," "direction to/into/out of/away from/through/under," "use on/for/with," etc. are coded in the combination of a potentially cross-referencing DATIVE pronominal position at -[4]-, with a particular "postpositional" element -- it is post-posed to the pronominal it governs in the morphology -- at order-class position -5-. Some of these sequences of -[4]-Postp5- constitute part of the morphological template of a particular lexical item, while others are freely added or removed as additional specifiers of a predication, much like facultative "adverbial phrases" in less holophrastic language structures like that of English. Thus note that in (2) above, -[4]-15- in the forms for 'give' is part of the expression of obligatory predicate-argument relations for this ditransitive verb, coding the 'Recipient', while this sequence added to the transitive lexical item -[2]-[3].../chxm7- '[2] boil [3]' that underlies examples in (5) would give an additional specification of 'in [4]'.
Given this morphological possibility, we should observe that there is a kind of finite clause in Kikshit with a characterizing function; it seems to provide a description for some noun in a "higher" clause -- which English accomplishes with a restrictive relative clause, note -- as also denoting whatever is pronominally coded in the dative position of the verbal form constituting the "lower" clause. Observe the following example (8) from Sophia Klickitat, a speaker of Louis Simpson's generation:

(8) [S. Klickitat, quoted in Sapir 1909:230.3]

\[a_1-q_2-a_3-t_6-/h_7-[a]m_{10}-a_{11} \quad a_3-t'iwat\]

'they/one will \_ bring \_ 7-over \_ \_ a-bucket

\[q_2-a_4-l_5-k'i_6-/chx_7-m_7-al_8\]

in5-which4 one2 (habitually)-boils7+8 [things8]'

Observe that the -a3- absolutive3 in agathama 'it will be brought over' cross-references a3-t'iwat 'a/the bucket' as the 'Patient' of the transitive verb -[l]2-[l]3-t6-/ha7- 'bring' (directionality "hither" prvided by -t6-).\(^{10}\) The other verbal form describes the bucket as "the one in which one habitually boils/the one which one habitually uses for boiling/the one in which one is boiling/the one being used for boiling."

The verb form qalk'ichxmal has many of the characteristics of USITATIVEs and CONTINUATIVES that function descriptively in several other, specifically subordinate clause types, including the genericizing ANTIPASSIVE for transitives. Since it is an ANTIPASSIVE form of the verb theme -[l]2-[l]3-[l]4-l5--/chx_7-m_7-- note the presence of the morpheme -k'i_6- -- there is no pronominal expression of 'Patient' and the 'Agent' should be coded in the pronominal of absolutive3 order-class. What we find is that the impersonal pronominal q2- is initial in such a form, looking like an ergative2 pronominal element though having the derived grammatical relation of an "intransitive" subject.\(^{11}\) Note especially its characteristic lack of a TENSE prefix, inasmuch as it describes the denotatum of its dative4 pronominal as "habitually" or "characteristically" being of such-and-such kind in the frame of the other predication. In other respects such a form is like a complete and independent clause, with all the machinery of a definite predication of PRESENT TENSE or SIMULTANEOUS RELATIVE TENSE value about an elsewhere-denoted referent (here, the bucket in question introduced explicitly in the other clause as the 'Patient' of 'bring').

Since the description focuses upon and applies to a locative, instrumental, or other kind of "adverbially"-functioning entity, I borrow Stanley Newman's felicitous term for this class of "adjectives" which occur as deverbative participials in Yokuts (1944:162-7) in strikingly parallel fashion. In Kikshit, it should be stressed, these formations are TENSEless though finite clauses, "adjoined relatives" (Hale 1976) which have various translation-equivalents in languages like English and other "SAE"s depending on the particular "adverbial" role the focused-upon predicate adjunct (not argument) plays in the designated situation. The translation-equivalent in a language like English will shift among in which, on which, with
which, from which, etc., or among while Ving in it/them, while Ving on it/them, etc., depending on the particular complex verb morphology of the -'14-Postp5-sequence in a proper Kiksht lexical construction.

It can be understood now that the so-called "impersonal" pronominal q2- is lowest in rank in the space of denotational content, and is paradigmatically defective insofar as occurring only in ergativeq2 position in normal explicitly 'Tense'd finite forms. It is apparently drawn into the system of cross-clause reference-maintenance through the adjunctive participial formation, where it apparently serves as the "dummy subject" for a clausal descriptor in which reference is maintained to the denotatum of whatever is cross-referenced in the descriptor's dativeq4 order-class. To see the significance of this kind of reference-maintenance in the Kiksht of Mr Simpson's and Mrs Klickitat's generation, we must look more broadly at the overall system of reference-maintenance across clause-like units in Kiksht. Then we will be able to see the nature of the changes obviously taking place in a following generation of speakers, fortunately caught by Sapir in records from 1905.

3. Clause-linkage and sites of referential binding. As a grammatical category, the type of CLAUSE-LINKAGE that connects two potentially autonomous predicables seems to constitute a complex space of multiple semantic and pragmatic connections. Within this space, however, it has become possible to define various grammatical-categorial parameters that make a typology useful to prediction across all morphosyntactic types. The two abstract parameters of interest to us here are [a] LOCAL vs. GLOBAL variants of CROSS-CLAUSE REFERENCE-MAINTENANCE, and [b] (relative scalar) DEGREE OF LINKAGE of predicables. We eschew here a careful development from fundamentals (for which see Silverstein 1976, 1985, 1993; and -- developing Schachter's [1977] concepts of semantic 'Role' and pragmatic 'Reference' -- Foley & Van Valin 1984; Van Valin [ed.] 1993; and numerous specific language studies). We can briefly and intuitively characterize these dimensions, in particular as they apply to Kiksht.

3.1. Forms used in Kiksht LOCAL REFERENCE-MAINTENANCE. There is sometimes practical identity of reference in potential noun phrase (NP) distributional categories that code arguments; these frequently occur in what, to "predicate" (in Searle's [1969] sense) both predicables for these arguments, would have to be two autonomous formal clauses. Cross-clause REFERENCE-MAINTENANCE is the generic category for how language structure marks continuity/discontinuity of reference for denotata of argument-coding noun phrase positions -- call them NP1 and NP2 -- in two such different clauses. Depending on whether the respective clause-internal grammatical-categorial and distributional (formal) properties of both NP1 and NP2, or only of either NP1 or NP2 are directly and transparently coded in some formal distributional material, we have, respectively, a "GLOBAL" system of reference-maintenance marking (one NP each across two clauses) or a "LOCAL" one (a target or "pivot" NP position defined in its clause only). The typically recognized GLOBAL systems are so-called "switch-reference" systems (see Jacobsen 1967; Haiman & Munro [eds.] 1983; Foley & Van Valin 1984:339-67) that signal syncretically for NPs in two linked clauses their RELATIVE REFERENTIAL IDENTITY and ARGUMENT-TYPE, e.g., referentially "same" vs. "different" most topical thematic-role or argument (mistaken as "subject" in much of the literature).
Kiksht, by contrast, has a LOCAL system of reference-maintenance, which is coded in the subordinate, restrictive-relative-like clause describing a denotatum as an 'habitual agent/actor', translatable by an English construction like who/which habitually Vs. Observe that the fully LOCAL system manifests itself only in this particular type of CLAUSE-LINKAGE; the lower clause is specifically and differentially a restrictive descriptor of the denotatum introduced by some NP in a higher-level structure (thus distinguished from "adjoined relatives" such as those discussed above in 2.3.). The descriptor occurs as a deverbalitative nominalization, thus inflected with obligatory 'Person/Number/Gender' pronominal prefix and with adnominal dative, i.e., 'Possessor'-like genitive, for the "pivot" or anaphor bound by the coreferential upper NP. This derived anaphoric site characterizes intransitives no less than transitives, as (9) will make clear.

(9) finite verbal forms and their derived nominalizations:

(a) intransitive: -[]3- /-ga7-la1g 'be flying about';

via indirect mediopassivization:

(nal1-) a4+xl5 /-ga7-la1g 'she4 [who] (recently-was1) flying-about7+8 [for5-self]';

via nominalization:

i3-cha4-l(xl5-) /-ga7-la1g 'she-who4 flies-about7+8'.

(b) transitive: -[]2-[]3- /-pcha7-la1g 'be sewing';

via antipassivation + indirect mediopassiv.: 

(nal1-) a4+xl5-k'i6 /-pcha7-la1g 'she4 [who] (recently-was1)

sewing7+8 [things6] for5';

via nominalization:

i3-cha4-lxh+i5-k'i6 /-pcha7-la1g 'she-who4 sews7+8 [things6] (for5)'.

As the examples in (9) show, and as I have elsewhere described (esp. 1976:146-9; 1993:493-4), the "indirect reflexive" (=MEDIOPASSIVE involving a dative pronominal position) and the ANTIPASSIVE construction are centrally involved in the system of clause-linkage in Kiksht, since they are essential to the derivation of forms with properly linkable pivots (and cf. sec.2.2.). Notice that the intransitive verb -galag in (9a) normally takes an absolutive pronominal inflection, and that the addition of the unmarked (or "dummy") postpositional sequence -[]4-15-, mediopassivized, allows the inflection to be -[]4+xl5-, with proper dative site of anaphoric binding if the finite form is to be used as a relative-clause-like descriptor. The nominalized verb, with or without the -xl5- nominal form of ...xl5...,
uses the dummy 'P/N/G' prefix i_3- '[3]Sg.Masc.', and the derived possessor -cha_4 in the adnominal dative_4 or genitive_4 position. The transitive verb -pchxalal in (9b), too, requires in effect the addition of a -dative_4-postposition_5- sequence to allow the derivation of the finite MEDIOPASSIVE, and in addition uses the ANTI PASSIVE to delete (and, semantically, genericize) the 'Patient'. The deverberative nominalization here, too, shows the genitive_4 pronominal -cha_4- 'her' for the potential pivot when the form ichaxhik'ipchxalal is a descriptor of habitual agency.

Having been introduced to both the finite and nominalized forms involved in local reference-maintenance, we should note that the finite form with dative_4 "subject" can be construed as a descriptor of some denotatum introduced elsewhere, while the nominalized form is always so usable, with all of the characteristics of a restrictive relative. This fits into the general schema of LOCAL systems, as shown in (10). Observe that the order of clause 1 and clause 2 here is irrelevant to the generalization, though in this visual schema it is the second clause, clause 2, that contains the pivot NP position in some derived case-relations of permissible pivothood.

(10) Schematization of LOCAL REFERENCE-MAINTENANCE:

\[\text{\texttt{clause 1}} \quad [\ldots \text{Npi}_j \ldots] \quad \text{\texttt{clause 2}} \quad [\ldots \text{Npi}_k \ldots]\]

- LINKAGE of specifiable type and degree
- represented by inherent denotational content of NP-coding
- represented by anaphoric element bound as NPi (with agreement, case-marking in its clause)

where_1 is the referent(s) index, and j,k represent derived case-relation

As shown in schematization (10), it does not matter what is the grammatical function of any NPi in clause 1 to which the Kiksht forms in (9) -- which would occur in schematic clause 2 -- are bound; the entire burden of potential pivothood rests on the grammaticosemantic properties in clause 2 of whatever NPi can be derived into dative_4/genitive_4 order-class pronominal cross-reference. As clause types, moreover, these fit into both a language-specific, and a more general, universal set of expectations about CLAUSE-LINKAGE, to which I now turn.

3.2. Distribution of clause-types by DEGREE OF LINKAGE. Such LOCAL type of reference-maintenance seems to be particularly associated in
Kiksht with very "tight" or "high-degree" grammaticosemantic linkage of clauses (predicables). Indeed, it is possible to invoke a schema of overall scalar degree of clause-linkage on which various specific semantic categories of inter-relationship of predicables can be arrayed. We can study for each language, as for language in general, the "space" created by distributions of reference-maintenance mechanisms in this cline of linkage types. Indeed, there is an important cross-linguistic empirical regularity that emerges as the structure of this "space," one illustrated here for Kiksht. (It is thus important to see that Kiksht is a language in specific conformity to the space of linkage types, making a distinction as will be seen of three regions of the way that formal reference-maintenance devices are distributed in linkage-space.)

Such a general schema, reproduced from Silverstein 1976:163, is annotated in (11) for the Kiksht specificities to be discussed.

(11) Schematization of DEGREE OF LINKAGE of clauses:

\[
\begin{array}{lcl}
\text{Ergative languages} & \text{possessive} & \text{bound \[-[Gen]_4-\ anaphor} \\
\text{habitual actor} & \text{habitual agent} & \\
\text{relative clause (making definite reference)} & \text{purposive complement (dative infinitive)} & \text{adj.'d, fin. cl.; NP del. in 2nd} \\
\text{desire complement} & \text{indirect discourse complement} & \\
\text{temporal adverbal clause} & \\
\text{if-then disjunction} & \text{parataxis} \\
\text{conjunction} & \\
\text{clause sequence (sequitur)} & \\
\text{clause sequence (non-sequitur)} & \\
\end{array}
\]

For each language with more than one type of reference-maintenance mechanism involved in one or more of these LINKAGE DEGREES, the LOCAL REFERENCE-MAINTENANCE mechanisms of linkage, such as sites of bound anaphora (pivothood) are associated with relatively HIGH-DEGREE LINKAGE, while GLOBAL REFERENCE-MAINTENANCE mechanisms such as "switch-reference" systems are associated with a region below, and no specifically structural reference-maintenance, distinct from the mechanisms of textuality, are associable with the lowermost regions. Of course, no specific language need have both LOCAL and GLOBAL mechanisms; but when they do, they conform to these generalizations, which then can be said to hold over all languages. Observe how the facts of Kiksht conform.

As shown in (11), in Kiksht the 'habitual Agent' and 'habitual Actor' modifier-clauses are the most tightly linked; they have LOCAL REFERENCE-MAINTENANCE by employing the bound-anaphora genitive4 position in the 'habitual' descriptor. At the very extreme, grammaticalized multi-clause constructions cease showing special structural features distinct from what they would show as sentence-scopes in running denotational text (though one must
never confuse the projection of structural types with the existence of multi-sentence
scope segments of text!). For a range of recognizable types of semantic and
pragmatic linkages, Kikst uses merely structural parataxis as its mode of
indication, any REFERENCE-MAINTENANCE being simply a function of how textual
reference is treated, for example in non-repetition of nouns, or of the use of sense-
relations among nouns to employ reference-maintaining hyponyms, etc. In the
mid-range, note, various types of "ADJOINED" FINITE CLAUSES occur, with NP
deletion possible in the second in discourse order (not the same as structural
relation, recall!), and interpretation ranging over many English translation-
equivalents, from 'relatives' to 'temporal adverbials'.

Among such "adjoined" finite clauses are the FINITE DATIVE4
HABITUAL/CONTINUATIVE constructions of sec.3.1, shown in (9) in contrast to
their derivative nominalized forms used with anaphoric binding. Also here are the
ADJUNCTIVE PARTICIPIAL CLAUSES (sec. 2.3.) in q2-[]4-postp5... illustrated in
(8), again with the "pivot"-like target of reference-maintenance in the dative4
pronominal position. It is reasonable to take these constructions as more suggestive
of our anaphorically bound English types when they lack initial Tense1 prefixes and
have suffixal derivational markings of simultaneous or habitual value. In such
cases, we are dealing with the "higher" mid-regions of CLAUSE-LINKAGE,
translatable with restrictive relatives and purposives and the like, it should be noted
and retaining less of the autonomous predicability of coded clauses with full verb
forms.

We see here how "impersonals," though an ergative2-only pronominal
category, are implicated in the overall system of REFERENCE-MAINTENANCE by
occurring in a specific construction at a determinate DEGREE OF LINKAGE of
clauses. We see, furthermore, the way MEDIOPASSIVE and ANTI PASSIVE VOICES
are also implicated in their own ways in this space, creating a complex series of
categorial cooccurrences that constitute the matrix for the linguistic changes I wish
to describe.

4. The analogical space of propagation of change. As in any
linguistic change, there is a complex historical intersection of socio-indexical,
textual (discourse-structural) and grammatical dimensions that jointly constitute the
space in which a change arises and is propagated. To be sure, older work on
grammatical change, culminating in Kuryłowicz's "laws of analogy" (1945/49),
emphasized the structural-grammatical space that generates certain directional forces
on grammatical innovation and spread within the particular configuration of a
system. And contemporarily, the socio-indexical dimensions of phonological
innovation, as demonstrated by Labov (1972) and his associates, have figured
prominently as a more general model of linguistic change as it emerges from the
variation inherent in usage within a linguistic community (this focus culminating a
few years ago in a new variationist-historical journal).

Crucial to this latter project, of course, is the concept of a structured
"phonetic space" in which systemic readjustments of phonological systems take
place. But in transferring the conceptualization of structural change to grammar, we
must see that it is the space of (universal) grammatical-categorial
intersections conforming in a particular system that provides that
system's "space for analogy [=grammatical change]," i.e.,
"analogical space." We can see this very clearly in the particular way that
"impersonals" in Kikstho of the generations born after ca. 1875 have entered the
system of bound anaphora through a combination of sociolinguistic-textual and
grammatical-categorial factors. One factor is the borrowing, under an earlier Chinookan-Sahaptin bilingualism, of various Sahaptin enclitic postpositionals. A second is the analogical filling out of the otherwise "defective" paradigm of the impersonal qa- ergative, perhaps under increasing encroachment of English. And a third factor, the provision of a discursive site for the innovation, is the position of ADJUNCTIVE PARTICIPIALS and related constructions in the Kiksht-specific exemplar of the universal categorial space of CLAUSE-LINKAGE. Let us take these in turn.

4.1. Sahaptin-derived syntactic postpositions in Kiksht. As is diagrammed in (1) and exemplified in (2), (4), (6), etc., the complex morphology of Kiksht verbs provides a sequence of cross-referencing pronominal -[\|4]- followed by its "postpositional\|5" element of varied lexical content -- -gl5- 'for; toward', -gm5- 'next to', -ghl5- 'out of; away from', etc. In the whole Chinookan family, in fact, it is this morphologized sequence that covers many of the adverbial-case-like locative and directional relations that are coded in adpositional phrases in more syntactically structured languages (both head- and dependent-marking, note). The Sahaptin languages, spoken by long-term neighbors of Chinookan speakers, especially those in the easterly riverine areas, appear to be the historical source for borrowed enclitic elements such as those shown in (12).

(12)  \textit{Klikitat (Sahaptin)} \quad \textit{Kiksht (Chinookan)}

\begin{align*}
\#\text{Enmi'} & \ '\text{with;}[\text{made}] \quad > \quad \#\text{E'\text{mni}} \\
& \quad \text{from, out-of} \\
\#\text{pa} & \ '\text{in;}[\text{at}] \quad > \quad \#\text{ba}^{14} \\
\#\text{pama'} & \ '\text{for;}[\text{in order to}] \quad > \quad \#\text{ba'\text{ma}}
\end{align*}

etc.

Borrowing of such elements through increasing asymmetric bilingualism -- Chinookan speakers learning fluent Sahaptin languages but not vice-versa -- that was intensified under reservation conditions from the third quarter of the 19th century, resulted in a completely alternative way of coding these adverbial relations, as Sapir already noted after his 1905 fieldwork (Sapir 1907:541-2 & nn.; Sapir 1911).\textsuperscript{15} These modes have continued to coexist up to the period of moribundity of the linguistic community, with mid-20th century "semi-speakers" (Dorian 1977; 1981:106-7) largely favoring the "analytic syntax" mode as opposed to the sometimes highly lexicalized and idiomatic morphological one.\textsuperscript{16}

Peter McGuff, the fluent trilingual (Kiksht, Klikitat Sahaptin, English [learned or at least perfected at Chemawa boarding school]) whom Sapir engaged as "interpreter" for his 1905 fieldwork, was probably typical of the cohorts of speakers born in the last quarter of the 19th century; he was certainly typical of multilingual Kiksht speakers of whom we have linguistic records down through the last quarter of the 20th. Notice the following McGuff construction (from Sapir 1909:184.21) in (13) that employs the clitic #ba'\text{ma} with an absolute deverbative nominalization.
(13) shd3-axhtau [[lq-i5-lxhulh-#bam+] ish3-lxhlhx

'(dual)3 that [dipnet-fishing # for] (du3)staging'

(cf. wa3-lxhulh-at 'dipnet'; -a3-[l]4+xh-l5-lxhulh- 'to fish with a dipnet; to dipnet')

Since the underlying verb theme, an "inverse" transitive, is inflected with dative mediopassive 'Agent' cross-reference, the absolute form of deverbative nominalization preserves the morpheme sequence ...xh-l5-... as lq-i5-... and the absolute noun gilxhulh 'dipnet-fishing' is in construction with its clitic #ba'ma 'for' as a specifically and differentially "purposive" modifier of the nominal head here, the DUAL NUMBER-GENDER ishElxhlhx 'staging'.17 So it is a staging "for dipnet-fishing" or "for one to dipnet-fish from/with."

Observe the subtle distinction here between this kind of deverbative noun, whether ABSOLUTE or CONSTRUCT, i.e., "possessed" by its genitive-inflected, anaphorically bound subject, and the finite-verbal ADJUNCTIVE PARTICIPIAL (sec. 2.3.) inflected with "impersonal" q2-[l]4-..., with which various kinds of semantically-close "adjoined" purposives are made, as in Mrs Klickitat's example (8) above. In the Kiksht of speakers like Mr Simpson and Mrs Klickitat, only the possessed deverbative nominalization -- not the ABSOLUTE form in q... or k... illustrated in this purposive modifier -- is a structure with a potential syntactic pivot, a site of anaphoric binding by a higher-clause antecedent (see sec.3.2.). And the initial q- of "impersonal" ergative occurs only in the finite clause form of the merely adjoined participial construction. But note the remarkable formal and semantic convergence here in the Kiksht of Mr McGuff's generation, who extensively use the borrowed enclitics for marking specific and differential syntactic relations across clauses.

4.2. Filling out the "defective" IMPERSONAL paradigm. As we have remarked several times, the -q2- impersonal non-cross-referencing pronominal has a "defective" paradigm in that it occurs only in the ergative order-class. I believe that this, too, changed around the time of Mr McGuff's linguistic socialization, probably under the influence of English as an increasingly present and formative verbal matrix for those who, like Mr McGuff, received a school education in it.

Observe that the impersonal in English and many other "SAE" languages is THIRD PLURAL, they/them/their. Certainly the subject of the impersonal passive-equivalent is formally they; similarly the "epicene" and "common gender" REFLEXIVE and ANAPHOR has for some centuries been formed on them- and their-in vernacular usage, despite efforts at prescriptive standardization to the structurally unmarked masculine singular (see Bodine 1975; Barron 1986:190-216; Newman 1992).

The Chinookan family of languages has, by contrast, two grammatical categories that, in some sense, are equivalent to the "SAE" THIRD PERSON PLURAL (see Silverstein 1976:131-4; 1977 for description). There is an individuated or at least potentially enumerable THIRD PERSON PLURAL, with pronominal sign (i)l-, cross-referencing "plurals" of nouns designating persons, large animals, etc. (though there are many seemingly "arbitrary" assignments of plural category in these languages with grammaticalized NUMBER-GENDER). Thus: i3-/a3-kiutan 'horse [m./f.]' pluralizes as it3-kiudan-yu+ksh 'horses', with individuated plural
suffixes in fact. By contrast, there is a THIRD NEUTER-COLLECTIVE, with pronominal sign (i)lh-, that is used for a wide variety of non-individuable multiplicities, for neutralizing the THIRD (SINGULAR) MASCULINE/FEMININE distinction, and for certain other "genericizing" effects on denotation.

Of the two, it is clearly the NEUTER-COLLECTIVE that is implicatively the most similar to IMPERSONAL -q2-, and over the generations has in effect become something of a more paradigmatically integrated substitute for it. Indeed, if one compares Mr Simpson's (born ca. 1830) prose material in Kiksh with that of Charles Cultee (also born ca. 1830) in both Shoalwater (Lower Chinook) and Kathlamet (Upper Chinookan) dictated to Franz Boas circa 1890-1894, the use of -q2- IMPERSONAL in finite verb forms is quite comparable (as illustrated in (4) above). Mr McGuff's material already shows a difference even in this kind of inflectional usage. By the time of my own fieldwork with several of Mr McGuff's slightly younger consociates, I noticed a decided preponderance of -lh- NEUTER-COLLECTIVES in the very grammatical functions to which -q2- had been the exclusive earlier norm. In many people's usage, -q2- more or less did not occur by the 1960s, and always -lh+k2- occurred with non-cross-referencing "impersonal" value in the expected contexts.18 The important point of this spread of -lh- forms into grammaticosemantic functions of -q2- is that the NEUTER-COLLECTIVE has no order-class restrictions on occurrence as a cross-referencing pronominal, and hence the NEUTER-COLLECTIVE has a "full" and completely regular, predictable, and compositional case-form and order-class paradigm, as shown in (14).

(14) Forms of IMPERSONAL and NEUTER-COLLECTIVE codings:

\[
\begin{array}{cccc}
Erg_2 & Abs/Nom_3 & Dat_4 & N[Gen_4] \\
IMPERS. & q2 & & \\
NEUT.-COLL. & lhk_2 & lh_3 & lh_4 & lh_4
\end{array}
\]

As is seen, -lh_3/4- is the basic form; -lh+k- the ergative_2; -lh+a- the genitive_4. As the vertical lines indicate, these forms penetrate into the respective functions indicated by the cells of the table.

4.3. The CLAUSE-LINKAGE dimension of the analogical space.

It is precisely this penetration of the historical NEUTER-COLLECTIVE as bound anaphor under particular conditions of CLAUSE-LINKAGE that is the most revealing of the reality of grammatical categories in the life and growth of languages.

For if we look back at chart (11), we see that there is a scale of degrees of tightness of LINKAGE in which only the uppermost three construction-types are characterized by such "pivoted," LOCAL reference-maintenance that uses bound anaphora at the dative_4 cross-referencing pronominal. If such a type of construction were to spread by grammatical innovation, the organization of this grammatical category space predicts that the spread should work itself out from closer to more remote LINKAGE DEGREE, that is, the space provides the vectorial dimensionality for grammatical change.
The change in question consists of the innovative use of neuter-collective pronominal forms as bound anaphoric markers of local reference-maintenance in all constructions on (11) down to 'purposive' complements, now specifically and differentially indicated by a clause-derivative in construction with bama 'for; in order to'. As we have seen in (8), the inherited adjunctive participial construction, without special linkage-type marking, could serve in an 'adjointed' relative or even purposive/instrumental reading; and the form contains the inherited q2-impersonal in initial position, a cross-referencing -l4-dative4, and is otherwise formed analogously to the actual habitual agentive/actorial forms of the bound anaphoric syntactic type in the uppermost region of the chart (11). What has apparently occurred is the "analogical" spread down the linkage cline (11) of actual bound anaphoric status of the dative4 pronominal position in these forms, now explicitly marked with bama as purposive constructions. And the form of the anaphoric element has been generalized over the paradigm of inherited impersonals, as suggested in (14), by using the inherited neuter-collective form, with 'Impersonal' reading, as the bound anaphoric element.

Thus note the complex sentence in a text (Sapir 1909:190.7-9) from Peter McGuff reproduced in (15).

(15) qidau ga1-q2-n3-t6-\text{xh}7 [bama k'aya i3-lha4-mqt k\text{w}adau it3-lha4-]

------------------S1-----------------

'thus they\text{2-long ago1+6-did-to3-me4}

so-that not l4-am-ill and l4-am-strong,

lhxhiwulx, awachi dauk\text{w}a i3-yulhmaxh g\#i3-lh4-gl5-\text{xhu7-lal8}]

----------------S2-------------------------S3------------------

or-even likewise spirit-power3 (is-)the one-who#[it3] is-preparing7+8-for5 me4'

It will be seen that the 'purposive' marker bama, which I have italicized in (15), introduces a three-clause, doubly-layered disjunction (awachi 'or'), the first two clauses of which are in turn conjoined by 'and' (k\text{w}adau). I have bold-faced the chain of reference-maintenance markers, which all go back to the antecedent -n3- in the "impersonal" construction that opens the sentence as the highest clause. Observe that in each of the lower clauses (S1,2,3) the dative4-position reference-maintenance marker, obviously "bound" to a first person singular antecedent with 'ego'-focused referential value, is a formal neuter-collective that has no other value than to be the bound anaphoric "pivot" maintaining reference to the antecedent.

It does not matter to our point that there might be "external" pressure of a sort derived from the obvious increasing multilingualism of the Kiksh-speaking linguistic community at the period when this change takes place. This is merely a matter of the source of the local, "lexical" (in Bloomfield's sense) form of the new anaphoric element, the historical Kiksh neuter-collective, which, probably
not coincidentally, is the nearest translation equivalent of the English "impersonal" they/them/their.

The critical point is that both the input and outcome of this linguistic change are in conformity with the Kiksht avatar of a universal organization of possible clause-linkage types and reference-maintenance distributions across it, that alone makes a determinate prediction of what system-internal changes of various kinds are possible and, here, actualized by the workings of grammatical "analogy." Grammatical-categorial analysis, and only grammatical-categorial analysis, I maintain, shows what is "law"-like about so-called "analogy," a.k.a. morphosyntactic change. It is therefore an actual linguistic theory with historical consequentiality, unlike much in the literature that claims to be explaining linguistic change by either structure-independent ("functional") reductionism or ("formal") conjuring with mental-organic (re-)settings of distributional parameters.

1All of this was unproblematic in both European and American linguistics down to and including L. Bloomfield. A great transformation took place in the theoretical writings of Zellig Harris and Charles Hockett, who in a way have served as the precursors (or "thesis") to Noam Chomsky's apothecosis (or "antithesis"). For the field was derailed into the (hopeless) trajectory of searching for a complete and consistent "meaning"-less, or formal-categorially "autonomous," approach to linguistic structure as both a methodological and ultimately a theoretical commitment. While this is not the place for a detailed and documented history of science (or history of whatever kind of field linguistics has become), the nature of the intellectual era we live in as linguists should surely become a focus for etiological diagnosis when we contemplate the burgeoning free-market in "formalisms" that now constitutes the fin-de-siecle condition. Formalism is precisely the correct bottom-to-top structural approach, we should hasten to add, within an overall framework of studying grammatical categories, as of course was the point of Saussure's Cours, or of Bloomfield's Language, as charter texts of our field.

2English does not have a specific TRIAL or PAUCAL grammatical category within its general grammatical category of NUMBER, but we can use the measure phrases [three [N...]] and [four or five [N...]] to denote sets of these precise cardinalities. Hopi does not have a grammatical category of TENSE, of which English has two specific exemplars, the PAST and the non-PAST (called "present" in nonstructural accounts based in denotational intuitions), but Hopi uses the combination of a rich set of AKTIONSARTEN (lexical aspect-like categories) and ASPECTS crossed by EPISTEMOLOGICAL STATUS (Whorf's "assertion") deictics to effect, partly by implicature, distinctions equivalent to the more directly-coded English TENSEs. I am, moreover, aware of the controversial nature of Whorf's (1956:57-64;113-5;213;216-7) original contrastive claim about Hopi and English, particularly as it has been misunderstood to be a claim about whether or not Hopi-speaking people are able to extend chronological time points and intervals, or even to perceive duration in chronological temporal terms. Having looked carefully at the mounted counterevidence (see Malotki 1983, for example -- reviewed favorably by Shaull 1985 -- that concentrates on showing that denotationally calendric and clock-temporal lexical phrases have penetrated into contemporary Hopi discursive usage), I see that Whorf's actual claim about multiple grammatical-categorial routes to
potentially same or equivalent extensional differentiations -- nothing more than Boas' or Sapir's or Bloomfield's point about differences of linguistic structure in their (and our) coding-conceptualization of language -- stands exactly as before: Hopi is a "TENSE-less" language.

3Here we may note the so-called "accessibility hierarchy" of Keenan & Comrie (1977; cf. Maxwell 1979; Comrie & Keenan 1979) which, as stated, focuses on cross-linguistic conformities of accessibility to pivoting in the space of intersection of [a] LOCAL REFERENCE-MAINTENANCE systems of various formally-codable argument-roles under [b] specific and differential CLAUSE LINKAGE-TYPE/DEGREE of "restrictive relativization." At the unfavored extremes of intersection of these variables, e.g., restrictive relativization on the head of an adverbial clause, the coding forms, if at all occurring in a language, are very complex; at the favored extremes, e.g., subject of simple-predicate clause, the construction may be formally merged with even tighter linkages like participial adjectives. In the middle range, we have specific and differential robustness of coding.

Observe that Romaine's study of Scots English restrictive relativization (1980; 1982) illustrates the relevance of such clines to sociolinguistic (stylistic) and historical outcomes.

4It should be no surprise, then, that our greatest theorists of grammatical categories were ultimately interested in explanation at a diachronic, or historical, plane, as for example Saussure, Bloomfield, Kurylowicz, for all of whom 'analogy' was a synchronic structure of relationships that constitute a determinate force-field of possibilities for a linguistic system diagnosable through statistical fluctuations and ultimately change of possible forms.

5For a basic Boasian exposition of Chinookan grammar, see Boas 1911; for an early treatment specifically of Kiksh, see Dyk 1933; for a basic treatment of inflectional categories, outline of inflectional syntax, and the nature of the lexical strata that are subject to inflection, see Silverstein 1976, 1977, 1984.

6I have adapted normal English-language font in expected combinations to express segment-types of Kiksh for which special Americanist symbols are generally used. For example, sh expresses the hushing fricative, lh the voiceless lateral fricative, xh the uvular (or "back") as opposed to the velar (or "front") voiceless fricative, gh the voiced uvular stop.

7Note that in such text-sentence-forms as The murderer of Smith will soon be sought, the utterance presupposes that there is an individual who murdered Smith (by implicature on the unmarked singular that this is a single individual) and its utterance as a statement provides a characterization of the entity. Observe that it provides a grammatically formed descriptor, murderer of Smith, backing any possible reference in subject position which may provide, or may not, information to characterize any actual individual.

8In English, etc., where PASSIVE [: ACTIVE] 'Voice' is a well-developed category, the discourse-contextual conditions differ somewhat in terms of "paragraph"-level coreference conditions, bipartite (no expression of 'Agent') vs. tripartite phrasal construction of the clause, etc. And the impersonal usage is in a different register, relative to cultural regularities of genre of discourse, from the passive construction, with a characteristic textual distribution, as Labov and his students have descriptively approached with their numerical text-frequency methods (cf. Weiner & Labov

9 The relation of -it\text{t}_{10} to Kiksh morphological causative formations is only an apparent violation of this, since the suffix combination of ...-m(a)\text{g}...(-it\text{t}_{10}) appears to code CAUSATIVE where no other aspeclal material appears in the derivational base from which the causative formation apparently comes. Observe that such a causative will have the pronominal inflectional schema -[\text{r}_{2}]-[\text{r}_{3}]...-m(a)\text{g}...-it\text{t}_{10}, as though a full, direct transitive formation derived from a resultative-passive-like form *-[\text{r}_{3}]...-it\text{t}_{10}. Where other suffixes are present that already contribute the telic or resultative sense to a non-transitive derivational base, only the -m(a)\text{g}- morpheme appears in the transitive causative formation.

10 It can also be seen that this particular transitive verb is inflected with "impersonal" 'Agentive subject -q_{2}- in this very polite imperative substitute, a finite FUTURE TENSE form with -a_{11}-a_{11}1 circumfix and IMPERSONAL 'Agent' coding.

11 In a more elaborate and theoretically oriented analysis, it could be argued that the surface form is derived from the word-initial, hence "absolute" form of what we encountered earlier in 2.2, as -xh-, the MEDIOPASSIVE/REFLEXIVE marker normally found postposed to an explicit pronominal morpheme. We still find the phonological alternation of #q- with -xh- and of #k- with -x- in the absolutive vs. possessed forms of derived nominalizations from apparently mediopassivized verb forms used as descriptors of habitual agency/action, as described below in 3.

12 Thus note once again that useful typological predictions apply across such purely (self-styled) "formalist" divisions as 'configurational' and 'non-configurational' syntactic types, as of course they must to be empirically non-vacuous. Also, note that no consistent purely "formal" typology (i.e., no strictly Saussurean-distributionalist, or "algebraic" one in Jakobson & Halle's [1956:15] terminology) will in principle be able to capture these predictive regularities. Wherever they might seem to, they are actually bringing grammatical-categorial considerations -- or even semantic ones tout court -- into the picture.

13 Among the many complex and at least partly morphologized formal alternations of the Postpositional\text{5} elements, one involves the (finite) verbal vs. nominal status of the theme of which it forms a part. Here, for example, verbal Postposition\text{5} -l5- 'to, in, for' occurs in corresponding nominal themes as -l5- following common, though morphologically conditioned alternations between l\text{n} and i in this set of morphological elements. The nominal form xhi- in this nominal-stem-initial position (marked with \text{l} preceding the stem) is, furthermore, the CONSTRUCT ("dependent" or "possessed") STEM marker, the ABSOLUTE (non-"possessed") STEM marker being corresponding qi-: thus the ABSOLUTE noun of 'habitual action', i-lqik'ipchxalal 'sewing'.

14 Observe that all the Chinookan dialects in both Lower Chinook (Shoalwater, Clatsop) and Upper Chinookan (Kathlamet, Kiksh) have a verb root -/b\text{a}7 'go out' used also as a subsidiary motion root in ninth position class, e.g., -/b\text{a}7\text{h}- /b\text{a}7\text{-'} 'crawl out'; its antonym is -/p\text{a}7 [Lower Chinook -/p\text{a}7] 'go in'. Whatever the possibility of remoter connection, the enclitic #ba in Kiksh seems to be of independent historical origin in borrowing at the time horizon with which we are here concerned.
In both discussions, Sapir specifically notes and illustrates the use of these postpositionals as clause subordinators occurring as enclitics to morphological verbs while coding specific LINKAGE-types.

The fieldwork of Robert E. Moore -- see, for example, a preliminary report in Moore 1988 -- on contemporary linguistic culture among people of Kiksht heritage language has also been confirming this trend, obvious in my own fieldnotes from 1966-1974 research on Yakima Reservation (Washington) and Warm Springs Reservation (Oregon). Among contemporary people, to be sure, English has been a pervasive factor of their languagescape throughout their lives.

This is a platform on which to stand over rapids, built by cantilevering out from the riverbank, its projecting edge resting on two posts anchored below.

In the discourse of Mr Simpson's niece (brother's daughter) that I recorded in the late 1960s and early 1970s, the Kiksht forms built on -lh- THIRD NEUTER-COLLECTIVE are especially regular in elicitation protocols when she offers translation equivalents for English somebody, going from Kiksht to English or from English to Kiksht. Several examples are offered at Silverstein 1978:248-9 of her and others' using -lh+k2- and -lh3- "IMPERSOLN" paraphrases in Kiksht of EVIDENTIAL PASSIVE forms, illustrating this kind of usage and its equivalence to -q2- forms.

References


*My sincere thanks to the BLS21 organizers for their invitation, to Susan Ervin-Tripp for sound advice that has influenced the visual version, and to Robert Moore for a very useful reading of the draft that was presented on 19 February 1995.
Local and global functions of a borrowed/native pair of discourse markers in a Yucatec Maya narrative
Julie Solomon
Stanford University

1. Introduction.

Five hundred years of contact with Spanish has resulted in the introduction of many Spanish loanwords into the native languages of the Americas, including Yucatec Maya (hereafter referred to as YM), which is spoken by several hundred thousand people in the Yucatán Peninsula, Mexico and in northwestern Belize. Among these borrowings has been a profusion of discourse markers (hereafter referred to as DM's), such as entoneses\(^2\) 'then', 'so', which have joined the ranks of the pre-existing Maya discourse markers, such as \(ká\) '(and) then', 'when'.

Entoneses and \(ká\) are semantically overlapping linguistic items that are frequently employed in narratives; yet, like most other YM DM's, their textual functions have not been investigated systematically. In this paper, after providing a definition of DM's, defining my approach to their analysis, and reviewing previous work on the uses of entoneses and \(ká\), and I analyze the functions of these two items in a spoken YM narrative I recorded in 1994. In particular, I demonstrate that as DM's, they have complementary global functions: while entoneses serves to set orientation apart from action and mark succession BETWEEN narrative episodes, \(ká\) foregrounds successive actions (and key states) WITHIN narrative episodes. In addition, \(ká\) also functions locally as a subordinator and as a coordinator that marks the temporal relationship between successive or overlapping events. This finding has two important implications: it indicates how past analyses of these Maya forms must be revised and expanded, and it also demonstrates how the analysis of DM functions can be significantly refined by distinguishing not only between local and global relations, as Schiffrin (1987) does, but also among degrees of global relations.

II. Definitions and Approaches
A. Definition of and approach to discourse markers. My approach to DM analysis takes as its starting point aspects of Schiffrin's seminal work on DM's (1987). First, I follow Schiffrin (1987) in defining DM's as "sequentially dependent elements which bracket units of talk" (31)\(^3\) that "look simultaneously forward and backward" in the text (37). Schiffrin (1987; see also 1992) discusses connectives and other items that can function as DM's in terms of their local and global uses on various planes of discourse, including semantic, pragmatic, and cognitive planes. In this paper I focus on the functions of entoneses and \(ká\) in IDEATIONAL STRUCTURE, a semantic plane of discourse that is made up of propositions, or ideas (Schiffrin 1987:25-6, borrowing from Halliday 1977). Among the criticisms that have been leveled against Schiffrin's model of discourse structure is Redeker's claim that "haziness surrounds the notion of ideational structure" (1991: 1162); in particular, across her work, Schiffrin is not consistent in how she defines the local and global relationships between the units that comprise this plane. By extrapolating from Schiffrin (1987:24, 247 and 1992:789, ft. 8), I define LOCAL RELATIONSHIPS within ideational structure as those that exist between two contiguous propositions without considering the organization of the texts in which they may be embedded. I define GLOBAL RELATIONSHIPS in ideational structure, meanwhile, as (1) those found between propositions that are not adjacent or (2) those found between contiguous propositions when one
considers the organization of the texts (e.g. narratives) in which they are embedded. Finally, I adopt Schiffrin's position (1987) that by definition a connective has a DM function at ideational structure if and only if it is functioning globally, in addition to any local function it may also have.

To make these definitions more concrete, let us consider examples of local and global uses of English then on the ideational plane:

1. Q: Did you visit John when you went to New York?
   A: I visited him then.  (Schiffrin 1987: 248-9)

(2a) There's a cop up the next block.
(2b) There used- there was a cop over here.
(2c) He's retired.
(2d) Then uh, there a a [sic.] couple uniform cops around here somewhere.
     (Schiffrin 1992: 771)

The then in (1) is not a discourse marker, for it functions only locally to indicate temporal overlap between the events in the question and the answer. The then in (2d) is a discourse marker, as it is functioning globally to mark a return to the list (begun in 2a) after a descriptive digression (in 2c). Note that I follow Schiffrin in assuming that a token can be multifunctional (see examples below; see Fraser 1990:389 for an opposing viewpoint), but I recognize that in practice it is often difficult to differentiate an ambiguous token from a multifunctional one.  

B. Narrative structure and genre. I have drawn the principal YM data for this paper from narrative, which Labov (1972:359-60) defines as "one method of recapitulating past experience by matching a verbal sequence of clauses to the sequence of events which (it is inferred) actually occurred." In his terms, a fully-formed narrative may include an ABSTRACT, ORIENTATION, COMPLICATING ACTION, EVALUATION, RESOLUTION, and a CODA (ibid.:363). In particular, the orientation give descriptive information about the setting and characters, and the complicating action units comprise the main body of the narrative; each describes an event. I will be focusing on a single narrative text, a more or less factual account of the contemporary settling of the narrator's native village, Xcocomil, Yucatán, by a group that included the narrator's father. It exhibits the characteristics of YM formal narrative speech, including repetition and parallelisms (see Burns 1973:145-64, Mudd 1979). James Fox (personal communication) classifies this text as a member of the ñuchben tzikbal 'myth' subgenre (see Burns 1973), since the narrator defines the time frame as ñuchih 'long ago' and the characters as ñabwolosô'ob 'forefathers' (see YM text in Sect. IV, lines 1, 2, 5, 21), and he weaves into the story an account of the origin of current local political strife.

III. Entonces and ká: Functions in Spanish and Maya
   Before examining how entonces and ká are functioning in this narrative, it is useful to review what is already known about their uses in YM generally. In addition, since entonces has been borrowed from Spanish, a language with which YM is still in contact, I will present some data from contemporary Spanish in order to suggest some functions that we might find in the YM data.  

A. Functions of entonces and entonces. Cortés Rodríguez (1991:89-97) classified by function all of the tokens of entonces in a corpus of spoken Peninsular Spanish.
Most of his categories reflect, in Schiffrin's terms, local resultative and anaphoric uses and global textual functions at ideational structure. By analyzing tokens of *entonces* in a Yucatec Spanish oral narrative produced by the same native speaker of Yucatec Maya whose Maya story I will discuss later, I determined that it has similar functions in this contemporary Spanish dialect; some examples are in (3a-b)\(^6\):

(3a) *Y esa sascabera que donde sacamos sascab y empezamos a meter pura piedra*
and that quarry that where remove-1P/pret clay and begin-1P/pret to put/in-inf pure stone
*Adentro ... Entonces se llenó, ni asi ... y entonces la maquinaria allí ah limpió*
inside so refl/3 fill-3S/pret not thusly and then the machinery there uh clean-3S/pret
*esa pedacita.*
that piece-dim

'And that quarry where we took out clay and we began to put stones in .... So it got filled up, not even like this ... and then the machine there uh it cleared that little piece [of land].'

b. *Este, ya que dejas tu escuela, ya entras con los ejidatarios ... a trabajar ....*
uh already that leave-2Spres pp2S school already enter-2Spres with the landowners to work-inf
*Entonces pues la escuela ... hay un, sascabera allí s- allí lo sacamos sascab hicimos*
So well the school exist-pres a quarry there fs there acc3S remove-1Ppret clay make-1Ppret
*una cueva muy grande hah?*
a cave very big yeah

'Uh, once you finish school, you enter in with the communal land owners ... to work ....
So well the school there c- uh there's a quarry there w- there we took out clay we made a big cave, right?'

In (3a) the first *entonces* is functioning locally as a resultative linking propositions of cause and effect, while the second is functioning both locally and globally to connect sequential events and mark an action unit in a narrative, respectively. The *entonces* in (3b), meanwhile, is functioning globally as a DM to mark the transition from background discussion to narrative events; note the corresponding use of the present and preterite verb tenses in the description and action, respectively.

No in-depth study of Mayan *entonces* can be found in the literature, but various references to its functions suggest that it does NOT have the wide range of uses of its Spanish counterpart. Based on a review of cross-linguistic, principally narrative Maya data, Brody (1987:510-1) classifies *entonces* as a borrowed conjunction that can also function as a DM; such items, she claims, always occur clause-initially. In Tojolabal Maya, in particular, borrowed conjunctions occur most commonly at topic shifts and episode boundaries, and often they occur in a phrase that combines a number of native and borrowed DM's (Brody 1989: 18-9). In these doublets or triplets of DM's, which are common across Mayan languages, the borrowed Spanish form usually comes first; further, the sequence often fills the function of highlighting and emphasis (Brody 1987:511-2, but see Brody 1989:24-5 on sequences in Tojolabal; for examples from various Mesoamerican languages, see Stoltz and Stoltz 1995 and the references in Brody 1987:512).

Virtually no work has been done, meanwhile, on the use of borrowed Spanish DM's in YM, specifically. Ligorré Perramón (1990:145-6) merely states that he found *entonces* to be one of three markers borrowed "with a frequency markedly superior to that of the rest of the loans of this type" (translation mine) in the YM narratives he collected. Blair (1964:45) called *entonces* and other YM discourse markers "particle stems" that had yet to be further subclassified. Burns (1973:166) says that some borrowed and native DM's (including *entonces*) are "grammatically similar and semantically 'empty'." Finally, Lucy (1989:16) makes
reference to the role of *entones* in "draw[ing] out implications" in Yucatec Maya speech, but he does not provide any examples.

**B. Functions of ká.** Several authors have provided analyses of YM ká, however, including Andrade (1941), Blair (1964), and Blair and Vermont-Salas (1965-7, vols. I-II, hereafter referred to by volume; note that all examples from these authors are slightly adapted orthographically; in this and subsequent sections, ch, ch', tz, tz', and x will be employed in place of their IPA equivalents ċ, ć, č, ć', and š, respectively; glosses are mine and contain the abbreviations listed in ft. 6). The uses of ká fall into two general categories: subordinating and coordinating. Subordinating ká functions as a complementizer, as seen in (4a), when it follows a verb of wish, command, or preference; the verb it subordinates take the suffix -Vk (V= echo vowel) if transitive or -ő if intransitive (Blair and Vermont-Salas II:568). In a related construction, subordinating ká functions like the English optative modal may, as in (4b) (see ibid.:566-7). In addition, subordinating ká can function like the English adverbial wh-phrase when to mark temporally overlapping past events (see ibid.: 687-8, Andrade 1941: 373). When the clause containing ká precedes the clause containing the temporally overlapping event, the construction is of the form (le) ká + completive verb + e'; when the ká clause follows the clause with the temporally overlapping event, the form is (le) ká + completive verb (see 4-d).

(4a). *In kátat ká talake' ex a k'ahóolt in tănaho'.*

erg1S want-incomp that come-sub-abs2P erg2 know-sub erg1S household-deic  
'T want [for] you to come to meet my household.' (Blair and Vermont-Salas I:323)

b. *Ká séeb utzahakahеч.*

may fast good-become-sub-abs2S  
'May you get well soon.' (Blair and Vermont-Salas II:560)

c. *Ká h k'úl' [juben xóok yáax t u káahil Káantamayeke', yan ten syéete 'ányos.*

when comp deliver/pass/comp-abs1S study first at erg3 town-relab Cantamayec-top exist pron1S seven years  
'When I was first put in school in the town of Cantamayec, I was seven.' (Blair and Vermont-Salas II:607)

d. *Tu'ux h binech (le) ká luk'each Káantamayeck?*

where comp go/comp-abs2S (the) when leave/comp-abs2S Cantamayec  
'Where did you go when you left Cantamayec?' (ibid.:687)

This last subordinating use of ká is closely related to its conjunctive use, which is to link successive clausal events. Andrade (1941: 370-1) concluded, based on the data he examined, that the conjunction ká cannot precede the first statement of a discourse or the first sentence spoken in answer to a question, it cannot be used before a negative, it is always clause initial, and the verb that immediately follows it must be in the completive or in the inceptive (*hoop* 'began to'+ *incompletive verb*) aspect, as in (5a,b). As (5b) shows, ká + *hoop* can contract to káp (Lucy 1993: 122, ft. 32, citing William Hanks, personal communication).

(5a). *ká h máaanen t ho' ká t in wilah Hwáaan ká ...*

then comp move/comp-abs1S to Mérida then comp erg1S see-comp Juan then  
'(and) then I moved to Merida (and) then I saw Juan (and) then .... '  
(Blair and Vermont-Salas II:624)
b. 

*pues káp u bin u kam t u t'án*

well then.Begin/comp erg3 go erg3 raise-sub to erg3 speech

'well [then he began to raise his voice' (Lucy 1993:110, line 117)

In contrast with Andrade's claim that conjunctive ká co-occurs only with completive and inceptive verbs, however, the *Diccionario Maya* (1991:Maya-Spanish, 276), which draws upon sources from the 16th century to the present, gives examples of conjunctive ká (spelled ka') co-occurring with other verbal aspects. Finally, Burns (1973:127) considers conjunctive ká to be "a defining feature of narrative," where it is "used as an utterance-beginning term. When it occurs in this position ... it seems to signify only that the utterances are part of a narrative ..." (127). After examining the YM text in Section IV, we will be able to evaluate the claims of Burns (1973) and others with respect to the features and functions of YM *entoneses* and ká (for comparisons with the closely related Itzá Maya, see Hofling 1987:485).

IV. The Yucatec Maya Narrative

The narrator of the text to be analyzed is a male corn farmer who was in his mid-forties when I recorded the text in his house as part of an interview (March, 1994). He is bilingual and literate in Maya and a nonstandard variety of Yucatec Spanish. The transcription was initially done by a Spanish-Yucatec bilingual college student in Yucatán, but it was revised by several additional native and nonnative Maya speakers (including myself). A phonetically transcribed excerpt is presented below; line numbers refer to the complete transcript, and ellipsis marks places where parts of the original text have been omitted. Phonological material that is present in underlying representation but has been deleted in speech is in small brackets; editorial comments are italicized and in small brackets. The text is split into lines by breath pauses (following Lucy 1993:103), and commas mark pauses internal to lines. *Entoneses* and ká are in boldface, and tokens of these forms that are DM's are underlined; other items I will be mentioning are in small capital letters.

**[INITIAL ORIENTATION--1]**

1. o sea le: tyéempôh 'túchih
   or be-sub the time happen-comp
   In other words, a: long time ago,
2. le:, 'dabwélósô'obo'seny- senyorô' [ob]. [Spanish:] son los abwéelos
   the grandfather-P-P-deic sit-P. [Spanish:] be-P pres the grandfathers
   the forefathers gentle- gentlemen, they are the forefathers
3. *dése Tz'înup ...*
   from Dzitunp
   from Dzitunp ...

**[EXTENDED ORIENTATION--2]**

6. *ká hóop' u::, 'u tâal u yéemlóo[b]*
   then begin/comp erg3 erg 3 come erg3 descend-incomp-P
   Then the::y, they began to come down,
7. *ká lük'ôob Tz'înup tún tâalô'ob t'ôon*
   then leave/comp-P Dzitunp prog3 come/incomp hunt/incomp
   then they left Dzitunp to come hunting,
8. *tún tâalô[ob] t'ôon beyô' ká h k'uchô'ob waye'*
   prog3 come/incomp-P hunt/incomp thus-deic then [or: when] comp arrive/comp-P here-deic
   they were coming hunting thus, then they arrived here.
   [or:] they were coming hunting thus when they arrived here.
9. chéen làa'[b] kàah ...
just ruined town
Just a ruined town ...

[END OF ORIENTATION]

14. entóonse
s o
So,

[COMPLICATING ACTION: HUNTING EPISODE--3]

15. ká h tāalóobè', 'um p'ée pàartée' ká h táap' Chan Chichimila'
when comp come/comp-P-top one cl/thing part-top then comp hit/comp Chan Chichimila'
when they came, as for one group, then they came across Chan Chichimila'.

16. káp u sëegr u tāalóob[bi] ts'don, ká t u káaxt[ah]o'obi' u la' káaha[l]i'e' pero bey hats'utz
then/begin/comp erg3 continue/incomp erg3 come/incomp-P hunt/incomp then comp erg3
find-[comp]-P-immediately erg3 other town-top but thus pretty
Then they began to continue coming to hunt, then they found right away [what was] another
town, but [it was] a pretty one.

17. ká h tāa[l]i mìn hay tuuli'[li], ká t u káahsah ká u ts'ón mehen nàa'yi'[li]
then comp came/comp some how/many cl/animate-relab then comp erg3 begin-cause-comp
that erg3 hunt-sub little house-relab
Then came some people, then they began to cause [to be made], in order for hunting [so that
one hunt], little huts.

18. káp u sëegr r u binó'o[b] letió'o[b] ...
then/begin/comp erg3 continue/incomp erg3 go/incomp-P pron3-P
Then they began to continue to come ...

21. bey u tibbatik le:: ñabwelosó'ob ñučha'
thus erg3 say/it-incomp the grandfather-P-P happen-comp-deic
thus say the:: forefathers of here long ago.

[COMPLICATING ACTION: FARMING EPISODE--4]

22. entóonse PWES: káp u bin u káahskol'[ob] bey waya'
then well then/begin/comp erg3 go/incomp erg3 settle-incomp-P thus here-deic
Then well, then they began to go settle thus this place;

23. DESPWÉESE', k u binó'o[b], k u ká'a binó'o[b] Ts'itinup
after-top incompg erg3 go/incomp-P incompg erg3 two go/incomp-P Dzitup
after, they go, they return to Dzitup.

24. k u tāaló'o[bi], tāak u meent la- káp u men[i]k u kółol'[ob]
incomp erg3 come/incomp-P want erg3 make is then/begin/comp erg3 make-incomp erg3 farm-
P
They come, they want to make? then they began to make their [corn] farms.

25. k u mëe[n]t[i]k u kółol'[ob] bey waya' PWES
incomp erg3 make-incomp erg3 farm-P thus here-deic well
So they make their farms here thus,

26. káp u chan p'áat[a]l'o'[ob]
then/begin/comp erg3 little stay-incomp-P
then they began to stay a little bit,

27. chéen 'u[m] p'ë semañane' ...
just one cl/thing week-top
just a week,

[COMPLICATING ACTION: SECOND HUNTING EPISODE--5]

30. entóonse ká t u sëegr u binó'o[b] letió'o[b] te' ts'donó'[ob] utz t u yichó'ob u ts'ón[i]ko'[ob]
le káax[li] k'éj[k]'eno'
So then comp [or: prog] erg3 continue/comp [or: continue/incomp] erg3 go/incomp-P pron3-P
there hunt/incomp-P good to erg3 eye-P hunt-incomp-P the mountain-relab pig-deic
So then they continued going hunting, they liked hunting the mountain pigs there.
[or:] So then they were continuing to go hunting, they liked hunting the mountain pigs there.
31. k u k’uchulo[’o]b tak Ichmul kàah...
    incompl erg3 arrive-incomp-P to Ichmul town
    They arrive as far as the town of Ichmul...

[CLIMAX: SETTLING OF XCOCMIL--6]

40. PWES: bey u tåalo[’o]b’o bey u tåalo[’o]b’o’ låakah t u yilah, u màasi:[l]
    so thus erg3 come/incomp-P-deic thus erg3 come/incomp-P-deic all comp erg3 see-comp erg3
    more-relab
    So, thus they come thus they come the rest saw it, more from:,

41. T’situnp[il]o’obe’, ká h ‘éemo’[ob], ká ya’abhi[l] máax waya’
    Dzitunp-[relab]-P-top then comp descend/comp-P then much-relab who here-deic
    Dzitunp, then they came down, then [there were] many people here.

42. DESPWSEE’ ká t u tukul t u bino[’o]b kàahah Chik’intz’ono’ot
    after-top then prog erg3 think-incomp prog erg3 go/incomp-P town Chikindzonot
    After, then they were thinking of going to the town of Chikintzonot,

43. PERO Chik’intz’ono’ote’ ma’ xan p’äato’obi, ká ka’ ‘éemo’[ob] waye’, way Xkokmil
    but Chikindzonot-top not also stay/comp-P-neg then two descend/comp-P here-deic here
    Xcocmil
    but as for Chikindzonot, they didn’t stay there, then they returned down here, here in Xcocmil.

[SHIFT TO ANOTHER TOPIC: CHIKINDZONOT--7]

44. entônoses, u màasi[l] T’situnp[il]O[’o]b TUNE’ ká h ‘éemo’obe’ dirèekto’ Chik’intz’ono’t
    binò’[ob]
    so erg3 more-relab Dzitunp-[relab]-P finally-top when comp descend/comp-P-top straight
    Chikindzonot go/comp-P
    Then, finally more from Dzitunp, when they headed down, straight to Chikindzonot they went,

45. PEERÓOH PWES Chik’intz’ono’ote’
    but well Chikindzonot-top
    But well as for Chikindzonot,

46. hach séeb ‘úuch u: ‘u káaksak ‘úucho’ nohoch kàah nohoch kàah...
    very quickly happen/comp erg3 erg3 populate-cause-sub happen/comp-deic big town big town
    very quickly long ago it became populated long ago [it was a] big town big town ...

V. Local and global textual functions of YM discourse markers

I now turn to an examination of entônoses and ká in this narrative to demonstrate that the ideational functions of these items are in complementary distribution: entônoses is a DM that globally marks the transition from one episode to the next, while ká is a connective that functions locally to link contiguous propositions and a DM that globally joins units within episodes. First, note that all tokens of entônoses and ká correspond with semantic (propositional) and syntactic (clausal) boundaries; in particular, all are clause initial. While all four examples of entônoses also correspond with line beginnings (which mark breath pauses), only seven of the ká’s are in line-initial position (another three are preceded on their lines only by Spanish DM’s). 7 In order to analyze the functions of the DM’s, I have divided the text into seven subsections by topic; these are numbered in the text above. The first two topics are the initial and extended orientation, respectively, the third topic is hunting, the fourth is farming, and the fifth is hunting. The sixth topic, the climax of the narrative, addresses the settling of Xcocmil, and the seventh topic is a new narrative about the settling of Chikindzonot. 8

The first topic, lines 1-5, comprises the initial orientation of the narrative; here the time frame, physical setting, and characters are introduced. There are no active verbs in this section, and no tokens of ká or entônoses. The second section of the narrative, the extended orientation, includes lines 6-13. Lines 6-8 are action units with completive or inceptive verbs (hóop u tåal ‘they began to come’,
*luk'ó'ob* 'they left', and *k'uchó'ob* 'they arrived', respectively), and all three contain *ká*. The *ká* in line 6 is functioning globally to mark the transition from background description to action units. The *ká* in 7, meanwhile, is functioning both locally to conjoin successive events and globally to indicate a successive action unit. Finally, the *ká* in line 8 is ambiguous: it could have local and global coordinating functions, like the *ká* in line 7, or it could be functioning only locally as a subordinating *ká* akin to English *when* (thus it is half underlined). These three lines, which tell how the forefathers came to the general area that is now Xcocomil and neighboring towns, are clearly part of the orientation of the narrative, as they are followed by a long description (lines 9-13) which orients the listener to the physical setting for all of the topics that follow. As in the initial orientation clauses, there are no active verbs and no *ká* or *entonces* tokens.

The extended orientation ends in line 13, and the principal complicating action starts in line 15; in line 14, globally marking the transition from description to action, is the first token of *entonces*. Tokens of *ká* are found in lines 15-18; they all precede completive or inceptive verbs (*ká* + *hóop* contracts to *káp*). The first *ká* in line 15 is a subordinator that locally marks temporal overlap between events in its clause and the following clause. The second *ká* in line 15 is functioning both locally to coordinate overlapping events and globally as a DM to mark a key action unit. The two tokens of *ká* in line 16 and the first two in line 17, meanwhile, have this same global function, as well as the local function of marking successive events. The third *ká* in line 17 is functioning locally as a complementizer.

The fourth topic of the narrative begins in line 22. *Entonces* (and *pwes*, see below) marks the transition. The *káp* 's in lines 22, 24, and 26 globally mark and connect the main action units and foreground them in relationship to the intervening material; the *káp* 's in lines 24 and 26 also locally mark events in temporal succession. The shift to the fifth topic occurs in line 30, where hunting is taken up again. Line 30 begins with *entonces*, and *ká* globally marks the initial complicating action; note that the verbal construction *t u séeegr* in line 30 is ambiguous between the completive 'they continued' and the progressive 'they were continuing'.

The sixth topic, which begins in line 40, is the climax of the narrative, for it treats how the forefathers ended up in Xcocomil, specifically. The transition between the descriptive discussion of hunting in Ichmul and the action-laden section on the settling of Xcocomil is not made by *entonces*, but by another borrowed discourse marker, *pwes* (again, see below). The first *ká* in line 41 and the one in line 43 function both locally to connect successive events and globally to mark the main action units; they both precede completive aspect verbs. The second *ká* in line 41 and the *ká* in line 42 are different, however, for they do not precede completive or inceptive verbs; instead, the token in 41 is in a clause that has no verb at all, and the one in 42 precedes a progressive aspect verb (*t u tukul* 'they were thinking'). In addition, both of these *ká* 's can be given resultative interpretations, and they both introduce key states in the chain of events leading to the settling of Xcocomil: in line 41, as a result of the settlers' coming down, there were many people in the area; in line 42, as a result of the crowding, people thought of going to Chikindzonot. In short, conjunctive *ká* can function as a resultative, even in verbless clauses, to foreground the material it introduces.

This sixth section includes the resolution of the narrative, for it completes the story of how Xcocomil came to be settled. The narrator shifts to a different but not unrelated topic in the seventh section, namely, how the neighboring Chikindzonot became a big town. The DM *entonces* globally marks this shift in line 44. Meanwhile, the *ká* in line 44 functions locally to mark temporarily
overlapping linguistic events; note the completive verbs éemó'ob 'they descended' and bíno'o[b] 'they went' in the temporally overlapping clauses.

In summary, the text in question contains four tokens of *entoneses*, all of which are functioning globally as DM's: the first marked the transition from orientation to complicating action; the second, third, and fourth marked episode shifts. *Entoneses* did not mark the shift to a very important episode, that containing the narrative climax, but another borrowed DM, *pwes*, served this function. Most tokens of *ká*, meanwhile, served local (non-DM) functions as subordinators or as coordinators of temporally overlapping or sequential linguistic events, and two tokens of *ká* appeared to be functioning locally as resultatives. All of the locally functioning coordinating *ká*'s and the two resultative *ká*'s were simultaneously working as DM's, for they filled the global narrative function of foregrounding important complicating action units and states within a given topic. Finally, there were several *ká*'s (in lines 6, 22, and 30) that appeared to be serving only this global function of marking action units within topics. Thus, in short, the ideational functions of *entoneses* and *ká* in this text are in complementary distribution: the former globally marks the transition from one narrative episode to the next, while the latter locally connects propositions and globally marks event and key state units within episodes.

As I referred to above, there are also a number of other items in this text that function as connectives and/or DM's, including Spanish-derived *pwes* 'well', 'so' (lines 22, 25, 40, 45), *peroh* 'but' (lines 43, 45), and *despwées* 'after' (lines 23, 42), and native Maya *tun* 'then', 'finally' (line 44). Line initial *pwes* occurs in a triplet with *entoneses* and *ká* to begin the only episode in which farming is mentioned (line 22); it occurs alone to begin the climactic sixth episode (line 40); and it appears in a doublet with *peroh* in the beginning of the new narrative about Chikindzonot (line 45). Line initial *pwes* thus seems to be functioning globally to highlight certain episodes or topics. *Despwées*, meanwhile, appears in sequence with *ká* in line 42, where it indicates an important shift in action internal to the episode. Finally, it is significant that the short climactic episode on the settling of Xcocmil has a particularly high concentration of DM's, including *pwes* (line 40), *ká* (two tokens in line 41, one each in lines 42 and 43), *despwées* (line 42), and *peroh* (line 43). Like the repetitions in Rickford's (1986:305) creole data, these items have the effect of "letting the individual events sink into consciousness slowly"; this is a DM function not in the semantic ideational structure, but in the pragmatic participation framework (see Schiffrin 1987:26-8 citing Goffman 1981, see also Halliday 1977). Ultimately, of course, a comprehensive account of the functions of YM DM's must look not only at ideational functions, but also at interactive and cognitive functions; it must consider sociolinguistic factors, such as the age, sex, and level of bilingualism of the narrator; and it must include data from across textual genres (e.g. see Solomon, in preparation, on YM DM's across narrative subgenres).

VI. Implications of the analyses

The analyses presented above allow us to evaluate critically not only previous claims in the literature about YM *entoneses* and *ká*, but also Schiffrin's approach to DM analysis, in particular her two-way distinction between local and global functions. First, contrary to the position of Burns (1973:12), conjunctive *ká* is not just a indicator of narrative genre; it always serves local and/or global linking functions in the text. Second, while the data support Andrade's assertion that conjunctive *ká* is always clause initial, they contradict his generalization that this *ká*
is restricted to clauses with inceptive or completive verbs (1941:370-1), for conjunctive *kd* can co-occur with progressive verbs and even in verbless clauses. Further, Brody's claim (1987:512) that sequences of DM's often function to highlight or emphasize does not hold for YM *entones* and *kd* (or for sequences involving *pwes* or *despwês* with *kd*); as was shown above, even when they are in sequence, these items fill complementary local and global discourse functions. Finally, Brody's position and mine on the function of DM sequences yield different predictions about the order of the linguistic items that compose them: although Brody (ibid.:511) observed that across Mayan languages, Spanish-derived DM's usually precede other DM's in sequences, her claim that these sequences are used for emphasis does not restrict the order of the components, while my analysis of YM *entones* and *kd* predicts that they when they appear in sequence, *entones* will be first. Further empirical data will show whether this prediction holds.

Finally, the YM data show that Schiffrin's distinction between local and global ideational relationships (1987, 1992) can be carried further, for a more precise analysis of DM functions requires reference to a cline of global relationships. In narrative texts, these global relationships range from connections between small chunks of text, such as complicating action units, to relationships of larger scope in which episodes or topics are linked. This more fine-grained distinction represents an important addition to current taxonomies of DM functions (e.g. see those of Schiffrin 1987, Fraser 1988, Redeker 1991).

VII. Directions for future research
The implications of work on YM DM's in fact reach beyond the study of the YM language and DM analysis to theoretical issues in borrowing and grammaticalization. As Brody discusses (1987:507-8), work on Mayan conjunctions and discourse markers has shown, contrary to previous belief, that (1) such forms can indeed be borrowed from one language into another and (2) borrowing often does not fill a linguistic gap (see ibid. for references). An issue for future research is the extent to which borrowed and native DM's have complementary functions, in YM and cross-linguistically, and what the origins of these complementary functions might be. For example, with respect to YM, in her discussion of loanwords in narrative couplets (note that there is no treatment of DM's), Mudd (1979: 58) claims that "the stylistic contribution of the [Spanish] loan words arises from the fact that they 'stand out' against a background of Yucatec words; they differ phonologically from the numerically predominant Yucatec words in both content and patterning." *Entones* stands out from surrounding Maya material because of its length (and *pwes* and *despwês* stand out because of their *pw* cluster). Evidence that these are desired features in linguistic items that mark more global relationships comes from a survey of other narratives (e.g. see Burns 1973): when *entones* or another loanword is not used to mark a topic or episode shift, the native word *ha'alibe* 'well', which is marked because of its length, is often used instead. Another possible origin of the contrasting functions of these DM's is that historically, such forms first come into a language at the highest (or most global) level of discourse (see Stolz and Stolz 1995:47). A cross-linguistic historical investigation of the grammaticalization of DM's would provide data to support or refute this theory. In short, approaching the study of DM's from discourse, sociolinguistic, and historical perspectives simultaneously promises exciting new data for a diverse range of linguistic theories and issues.
Notes

1. I conducted the fieldwork for this project in 1994 with funding from an NSF Graduate Fellowship and a Pre-Dissertation Field Grant from the Center for Latin American Studies at Stanford. My thanks go to the following people for their input: "Don Moises," Patricia Martínez Huchim, Santos Pedro Miz Serralto, Elizabeth Traugott, James Fox, John Rickford, Tom Wasow, John Lucy, and the participants in the Stanford/Berkeley Discourse Markers Interest Group and BLS 23. The content of this paper remains solely the responsibility of the author, however.

2. I will use the Spanish orthographic form of this term (entonces) when I am referring specifically to the Spanish use of this item; otherwise I will use the IPA spelling (entoneses).

3. Or units of writing, I would add, but the focus of this paper is talk.

4. An analysis of intonational features appears to eliminate the ambiguity in many English examples; an analysis of YM intonation, which I do not attempt in this paper, might do the same for some of the Maya examples (see also ft. 7).

5. I will leave the precise nature of the diachronic and synchronic relationships between Spanish entoneses and YM entones as an issue for future research.

6. Throughout this paper I will use the following abbreviations in glosses: 1= 1st person, 2= 2nd person, 3= 3rd person, abs= absolute, acc= accusative, cause= causative, cl= classifier, comp= completing, deict= deictic marker, dim= diminutive, erg= ergative, fs= false start, incomp= incomplete, inetr= interjection, neg= negative, n= plural, pass= passive, perf= perfect, pp= possessive pronoun, pres= present, pret= preterite, prog= progressive, pron= independent pronoun, refl= reflexive, relab= relational/abstractive suffix, s= singular, sfut= subordinate and subordinating future, to= top= topicalizer, /links features of a single morpheme (note: used only when the several glosses that must be listed for a single morpheme are unclear without it). - indicates a morpheme boundary.

7. A logical question at this point is whether a line division based on clauses (e.g. see Labov 1972: 361-2), pause length (e.g. see Burns 1973, following Tedlock 1972) or intonation pattern (e.g. see Chafe 1980) might be more telling. I prepared versions of the text with line breaks based on independent clauses and on pause length, but neither provided a significantly more informative picture of the DM's. Line breaks based on intonation might indeed prove useful, but given that YM is a tone language, and given that virtually no work has been done on YM intonation, I felt that such an undertaking was best left for a future project.

8. During the discussion following the presentation of this paper at BLS 23, Michael Silverstein suggested that employing a more complex textual architecture might help to provide an integrated account of the functions of not only entoneses and ká, but also the other borrowed and native discourse markers in the text. I will leave such an analysis, however, for future work.

9. Note that in YM, borrowed pwes is not always clause initial (e.g., see line 25 of the YM text in Sect. IV); however, an analysis of the functions of pwes is beyond the scope of this paper.

10. Although throughout this paper I have referred to ká as a native YM word, it is possible that it was in fact itself an earlier (pre-Conquest) borrowing from Nahuatl.

References


Solomon, Julie. In preparation. Marcadores de discurso dentro y a través de distintos géneros literarios de literatura oral de los mayas yucatecos contemporáneos. Paper to be presented at the Tercer Congreso Internacional de Mayistas, Chetumal, Mexico.


ON THE STATUS OF UNIVERSAL ASSOCIATION CONVENTIONS:
EVIDENCE FROM MIXTECO

BERNARD TRANEL
University of California, Irvine

1. INTRODUCTION. Two recent studies (Archangeli & Pulleyblank 1994: Chapter 4; Hyman & Ngunga 1994) have independently questioned the existence of the universal association convention (Goldsmith 1976) automatically linking together free elements from different autosegmental tiers, as in the case of floating tones and free tone-bearing units illustrated in (1a) (see also Odden 1995: 459-460). The gist of Archangeli & Pulleyblank's argument is that there is no reason to privilege the linking of free elements, because it actually exhibits the same sort of parametric variation as spreading. In complementary fashion, Hyman & Ngunga argue that floating high tones in the Bantu language Ciyaqo are never automatically associated to a free mora; rather, they are either linked by means of language-specific morphological and phonological rules or else stray erased, even if free tone-bearing units are available. Thus, instead of automatic association for free elements, both studies defend association by language-specific rules, just as the once presumed automatic tonal contouring and spreading illustrated in (1b-c) have been argued to be governed by language-specific options rather than universal principles (Clements & Ford 1979; Pulleyblank 1986).

(1) a. Automatic association of free elements:

<table>
<thead>
<tr>
<th>μ μ</th>
<th>μ μ</th>
</tr>
</thead>
<tbody>
<tr>
<td>T T</td>
<td>T T</td>
</tr>
</tbody>
</table>

b. Automatic contouring:

<table>
<thead>
<tr>
<th>μ μ</th>
<th>μ μ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>/</td>
</tr>
<tr>
<td>T T</td>
<td>T T</td>
</tr>
</tbody>
</table>

C. Automatic spreading:

<table>
<thead>
<tr>
<th>μ μ</th>
<th>μ μ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>/</td>
</tr>
<tr>
<td>T T</td>
<td>T T</td>
</tr>
</tbody>
</table>

In this paper, I present concurring evidence from Mixteco, a tone language of southern Mexico studied by Pike in the 1930's and 40's (see References). I will argue that, as in Ciyaqo, the purported universal convention governing the association of free elements must in fact not apply in this language, despite excellent opportunities. This situation seems to indicate that Universal Grammar should not include such a convention. Alternatively, if this convention is nevertheless held to be part of Universal Grammar, then the grammar of Mixteco must somehow encode that it plays second fiddle to all other tone-assignment rules in the language, including the process assigning a default mid tone to toneless vowels; one encoding possibility for this property would be to assign to the so-called convention the status of a parameter, which would happen to be set to 'off' in Mixteco.

2. BACKGROUND ON MIXTECO. Lexical words in Mixteco are bimoraic, typically of the shapes shown in (2).

(2) a. CVV
b. CVCV

Each mora can bear one, and only one, of three level tones: High (H), Mid (M), or Low (L). Thus, as diagrammed in (3), contour tones do not occur.

(3) No contour tones: 

<table>
<thead>
<tr>
<th>μ</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
</tr>
<tr>
<td>\ /</td>
</tr>
<tr>
<td>T T</td>
</tr>
</tbody>
</table>
The distribution of the three level tones over the two moras of a lexical word is almost free. Out of the nine logically possible tonal patterns tabulated in (4), there are only two restrictions, having to do with the italicized MH and LL patterns.

(4)  
HH    MH    LH
HM    MM    LM
HL    ML    LL

One restriction, explicitly noted by Pike (1944: 124; 1948: 57) and diagrammed in (5), is that the LL pattern does not occur at all (by contrast, the other two double patterns, HH and MM, are attested, as in sáná 'turkey' and bina 'today').

(5)  LL constraint:  * L  
       / \  
      μ μ

The second tonal gap, not mentioned by Pike, but of significance for our purposes, concerns the MH pattern. This pattern apparently fails to occur over words of the shape CVV and CV?V, when the two vowels are the same (i.e. in the case of CVV words, when the word contains a long vowel). The constraint in (6) conveniently conflates the two contexts (this formulation assumes two separate root nodes to express vowel length (Selkirk 1990) and a Supralaryngeal node dominating Place and Height (Goad 1993), and it takes advantage of the lack of a Supralaryngeal node for laryngeal consonants (cf. Steriade 1987). Nothing in my analysis will actually hinge on the specific formulation of this constraint; only the fact that the restriction exists in some form is important.

(6)  MH constraint:  * M H  
       | |  
      μ μ  
       | |  
      R R  
       / \  
Supralaryngeal

3. FLOATING HIGH TONES AND THEIR ASSOCIATION. The phenomenon of interest for this paper is that morphemes in Mixteco may include a final floating high tone, which manifests itself exclusively by, roughly speaking, replacing the first tone on the next morpheme, provided no pause intervenes. The minimal pair in (7) shows the effect of one of these so-called 'perturbing' words, in contrast with a 'non-perturbing' word (Pike 1948: 80). In (7a), the floating high tone of the perturbing word for 'to eat' replaces the initial low tone of the word for 'child', while in (7b), the homophonous but non-perturbing word for 'to go away' has no effect on the word for 'child' (I represent floating high tones with a parenthesized H).

(7)  a. kee (H) 'to eat' + súči 'child' —> kee súči 'the child will eat'
    b. kee 'to go away' + súči 'child' —> kee súči 'the child will go away'

Floating high tones also occur as part of what Pike (1948: 82) terms 'zero words' or 'ghost words', that is, morphemes without any independent phonetic realization. Thus, the continuative morpheme is a prefixed floating high tone functioning like the final floating high tone illustrated in (7a). In the example in (8a), we can see it replace the initial low tone in the verb for 'to sew'. (8b) shows the underlying LM tonal pattern of this verb surfacing.

(8)  a. Continuative:  (H) + kiku 'to sew' + ná 'I' —> kiku-ná 'I am sewing'
    b. Potential:  kiku 'to sew' + ná 'I' —> kiku-ná 'I will sew'
A floating high tone is also part of the consonantal causative morpheme $s$, as shown by the examples in (9), which reprise the verbs used in (7), but this time as targets of a floating high tone (Pike 1948: 90-91). We see that the floating high tone of the causative docks into initial moraic position in the two verbs for 'to eat' and 'to go away', changing their surface tonal patterns from MM to HM.

(9) a. s (H) 'causative' + kee (H) 'to eat' $\rightarrow$ skèe 'to feed'
    b. s (H) 'causative' + kee 'to go away' $\rightarrow$ skèe 'to cause to go away'

According to Pike (1948: 80), lexical words with the tonal patterns HL, MH, and LM 'never cause any perturbations'. Translated into our terms, lexical words with these tonal patterns can never include a final floating high tone. I list in (10) the apparently proscribed underlying patterns. These gaps remain to be explained.

(10) Restrictions on the occurrence of final floating high tones:
*HL(H), *MH(H), *LM(H)

Pike (1944: 123; 1946: 23-24, note 2; 1948: 79-81) also draws the conclusion that lexical words with the tonal patterns given in (11) may not be perturbed, since they remain unchanged when preceded by an otherwise perturbing morpheme.

(11) Unchanged when preceded by a floating high tone:
    a. HH, HM, HL
    b. MH

As I will now argue, Pike's restriction is actually unnecessary; the tonal patterns in (11) can be shown to be fully regular with respect to the attachment of a preceding floating high tone. For the three tonal patterns in (11a), which all begin in a high tone, a preceding floating high tone can simply be taken to attach to a mora that is already high-toned, resulting in no perceptible change.8 These patterns are thus not exceptions to perturbation by a preceding floating high tone; it is merely that the perturbation in these cases is of necessity invisible. With respect to floating high tone association, words with an initial high tone are therefore entirely parallel to words with an initial low tone, whose behavior is again illustrated in (12), with underlying LH and LM examples repeated from (7a) and (8a) above.9

(12) kee (H) 'to eat' + sùčí 'child' $\rightarrow$ kee sùčí 'the child will eat' [= 7a]
    (H) + kiku 'to sew' + ná 'I' $\rightarrow$ kiku-ná 'I am sewing' [= 8a]

The parallelism is diagrammed in (13).

(13) a. $\mu$ $\mu$ $\mu$ $\mu$
    \| \| \| \|
    (H) L T H L T

    b. $\mu$ $\mu$ $\mu$
    \| \| \|
    (H) H T H T

In both cases, the floating high tone can be described as attaching to the first tone-bearing unit in the following morpheme. The association of the floating high tone to a low-tone vowel causes the low tone to delink, since contour tones are barred in Mixteco (see Constraint 3 above), and the attachment of the floating high tone to an already high-tone vowel simply results in a high-tone vowel.

The case of the unchanging tonal pattern MH in (11b) presents a more challenging problem, which brings us more generally to the group of lexical words beginning in a mid tone. Given the just observed effect of a floating high tone on a following CVCV word with an initial low or high tone, CVCV words with an initial mid tone might be expected to surface with an initial high tone when preceded by a floating high tone. Indeed, quite generally, words with a MM tonal pattern fulfill this expectation, as diagrammed in (14).
The examples in (15), repeated from (9) above, illustrate this type of change with CVV words, and the examples in (16) with CVCV words. In (16a), the underlying MM pattern of the word for 'today' is changed to HM under the influence of the final floating high tone of the word for 'mountain' (Pike 1948: 77), and in (16b), the word for 'mountain' is itself subject to a change from MM to HM because of a preceding floating high tone (Pike 1946a: 23, note 2).

(15) a. s (H) 'causative' + kee (H) 'to eat' —> skée 'to feed'
b. s (H) 'causative' + kee 'to go away' —> skée 'to cause to go away'

(16) a. kî'n-ná + ñuku (H) + bina —> kî'n-ná ñuku bina
b. hà (H) 'that' + ñuku (H) 'mountain' —> hà-ñuku 'that mountain'

Yet, as Pike noted, words with a MH tonal pattern (e.g. kùći 'pig') are apparently not affected by a preceding floating high tone, that is, they don't change to a HH pattern. The absence of this change does not seem reducible to an OCP effect, since as illustrated in (7a) above, words with a LH pattern do change to HH when preceded by a floating high tone, thereby demonstrating that a floating high tone can link to a vowel immediately preceding a high-tone syllable in the same morpheme. The answer to the puzzle comes from the behavior of CVCV words with a ML pattern. Preceded by a floating high tone, these words do not change their tonal pattern to HL, but rather to MH, as illustrated in (17) with the words for 'branch' (Pike 1946a: 24, note 3), 'priest' (Pike 1945a: 131), and 'to scrape' (Pike 1948: 92).10

(17) a. hà (H) 'that' + ñuku 'branch' —> hà-ñuku 'that branch'
b. nuù (H) 'to' + sutú 'priest' —> nuù sutú 'to the priest'
c. mà (H) + tuhi (H)11 + rí —> mà-tuhí-rí 'I will not scrape'

'T'

These examples show two important things. First, in a ML pattern, the mid tone is transparent to the association of a floating high tone. Secondly, if we generalize this transparency property to the MH pattern, then this pattern becomes regular with respect to the association of a floating high tone. As shown in (18), in both ML and MH cases, the floating high tone bypasses the italicized initial mid-tone vowel and anchors itself to the word's second low-tone or high-tone vowel, with exactly the same effects that it had when it docked onto an initial low-tone or high-tone vowel (compare (18) with (13) above).

(18) a. µ µ —> µ µ
    (H) M L —> M H L

b. µ µ —> µ µ
    (H) M H —> M H

What needs to be captured formally now is first, the transparency of the mid tones in ML and MH patterns, and secondly, the process whereby a floating high tone gets anchored. With regard to the transparency question, I suggest as an answer that mid-tone vowels are underlyingly toneless, in other words that the mid tone is the default tone in Mixteco, as is commonly proposed for other languages with the three level tones High, Low, and Mid (see for instance Pulleyblank 1986).12
(19) M-Default: \[ \mu \rightarrow \mu \rightarrow M \]

Regarding the anchoring of the floating high tone, I suggest the two rules sketched in (20).

(20) (H)-Association:

a. \[ \mu \rightarrow /\mu \]
   \[
   \begin{array}{c}
   (H) \\
   T \\
   H \\
   T
   \end{array}
   \]

b. \[ \mu \rightarrow \mu \rightarrow H \]
   \[
   \begin{array}{c}
   (H) \\
   H
   \end{array}
   \]

Part (a) of (20) states that a final floating high tone targets the first toned mora in the next morpheme and anchors to it. As already mentioned, everything else in this situation results from independent operations. Thus, if \( T \) is a low tone, it is automatically delinked, since Mixteco does not allow contour tones (see Constraint 3 above), and if \( T \) is a high tone, the outcome is simply a high-tone vowel. Part (b) of (20) states that in case the floating high tone does not find a toned mora in the next morpheme, then it anchors to the first toneless mora in that morpheme.\(^{13}\) If neither (20a) nor (20b) can apply, that is, if the floating high tone occurs before a pause, then nothing may save it: it remains unanchored and is consequently not realized phonetically.

To summarize so far, Rule (20a) accounts for all of the patterns in (21a), and rule (20b) accounts for the single pattern in (21b).

(21) a. \[ HH \rightarrow HH, HM \rightarrow HM, HL \rightarrow HL \]
   \[ MH \rightarrow MH, ML \rightarrow MH \]
   \[ LH \rightarrow HH, LM \rightarrow HM \]
   \[ MM \rightarrow HM \]
   b. \[ MM \rightarrow HM, ML \rightarrow HL \]

Rule (20b) might appear to account for relatively little, but in fact it accounts for more than meets the eye in (21b). The analysis so far has essentially ignored both CVV and CV?V words where the vowels share the same melody, that is, words where Constraint (6) banning the MH tonal pattern is in effect. When preceded by a floating high tone, such words behave as schematized in (22).

(22) CVV and CV?V words with a single vowel melody

a. \[ HH \rightarrow HH, HM \rightarrow HM, HL \rightarrow HL \]
   \[ LH \rightarrow HH, LM \rightarrow HM \]
   b. \[ MM \rightarrow HM, ML \rightarrow HL \]

I provide in (23) examples for the cases where the first mora of the word does not carry an underlying high tone, that is, when the effects of a preceding floating high tone are visible. I show here only the presence of a preceding floating high tone, without specifying its morphological source. For each tonal pattern, CVV words are illustrated in (i) and CV?V words in (ii).

(23) a. \[ LH \rightarrow HH \]
   (i) (H) + káán 'yellow' \[ \rightarrow káán \quad (Pike 1945b, 1948)\(^{15}\) \]
   (ii) (H) + áñéé 'to look at' \[ \rightarrow áñéé \quad (Pike 1948: 94) \]
   \[ LM \rightarrow HM \]
   (i) (H) + čii 'beneath' \[ \rightarrow čii \quad (Pike 1948: 94) \]
   (ii) (H) + háá 'to give' \[ \rightarrow háá \quad (Pike 1945a: 131) \]
b. MM → HM
   (i) (H) + kee 'to go away' → kée (Pike 1948: 78-79)\(^6\)
   (ii) (H) + be'xì 'house' → bëxì (Pike 1948: 79, 80)
ML → HL
   (i) (H) + kóò 'snake' → kóò (Pike 1948: 79)
   (ii) (H) + ka'xàn 'to speak' → káxàn (Pike 1944: 124)

A comparison of the tables in (21) and (22) reveals two differences, highlighted by the italics on the tonal patterns beginning in a mid tone. First, CVV and CV?V words with a single vowel melody don’t show instantiations of an unchanging MH pattern; this gap is simply due to the lack of an underlying MH pattern to begin with (see Constraint 6 above). Secondly, when underlyingly ML, these word types derive a HL pattern instead of a MH pattern, thus revealing another application of Rule (20b). What needs to be explained now is why Rule (20a) does not apply in these cases, since it is supposed to take precedence over Rule (20b).

As illustrated in (24), CVV words with a long vowel contrast minimally with CVV words containing two different vowels, the latter behaving like CVCV words and changing their ML pattern to MH.

(24) a. ML → HL (H) + kóò 'snake' → kóò [repeated from 23]
   b. ML → MH (H) + žáù 'hole' → žáù [repeated from note 10]

This contrast can be explained by Constraint (6), which bans the MH pattern on CVV words with a long vowel. Because of the constraint, the floating high tone cannot successfully aim for what would otherwise be its normal target, i.e. the first toned mora in the next morpheme (Rule 20a). It must therefore settle for the first available toneless mora (Rule 20b), and the outcome is a HL pattern rather than a MH pattern. The latter pattern does obtain with CVV words containing different vowels, because these words are not subject to Constraint (6).

The situation is slightly different with CV?V words (Pike 1944: 124; 1948: 81). Here, as shown in (25), words with an underlying ML pattern change to HL, whether or not they contain identical vowels.

(25) a. ML → HL (H) + ka'xàn 'to speak' → káxàn [repeated from 23]
   b. ML → HL (H) + ta'nù 'to beat' → tá'nù (Pike 1945b: 219, 221)

Furthermore, as shown in (26), words of the shape CV?CV also behave in this fashion.

(26) ML → HL (H) + ka'nù 'to walk' → há'nù (Pike 1945a: 135)\(^7\)

The descriptive generalization here is that toneless vowels are not transparent to (H)-Association exactly when a glottal stop intervenes.\(^8\) I will therefore assume that a word-medial glottal stop constitutes a barrier to (H)-Association (Rule 20a), perhaps because its featural specification and that of a high tone share a tier, and that linking a high tone across a glottal stop would result in a line-crossing violation.\(^9\) Under this view, Constraint (27) would be nothing more than an instantiation of Goldsmith's line-crossing constraint (1976).

(27) High tone - Glottal stop constraint: No crossing over

At any rate, the effect observed in (25a), (25b), and (26) is that a floating high tone is unable to target the toned vowel in the second syllable of the next morpheme (Rule 20a), and must again settle for the toneless vowel in the first syllable (Rule 20b); hence the change from ML to HL instead of MH.\(^10\)

In summary, the main elements of the proposed analysis for the association of floating high tones in Mixteco can be itemized as follows:
(28) Summary of analysis

(i) Mid-tone vowels are underlyingly toneless (the mid tone is the default tone).
(ii) A floating high tone's domain does not extend beyond the next morpheme.
(iii) Its preferred target is the first lexically toned vowel (Rule 20a).
(iv) If a toned vowel is unavailable, it targets the first toneless vowel (Rule 20b).
(v) It is otherwise unrealized because unanchored.
(vi) The availability of a toneless vowel as an anchor is determined by three factors:
   a. Whether the following morpheme has a lexically tone.
   b. Whether the output would violate a restriction on tonal patterns (Constraint 6).
   c. Whether the association would create a line-crossing violation (Constraint 27).

4. THEORETICAL IMPLICATIONS. The view that mid-tone vowels are lexically toneless provides an appealing explanation for their transparency to the association of floating high tones. It also constitutes an attractive alternative to Goldsmith's proposal (1990: 24-26) that a floating high tone metathesizes with a following morpheme-initial mid tone in CVCV words where the medial consonant is not a glottal stop. However, the analysis leads to serious descriptive problems if there is a universal convention automatically associating floating tones to free vowels.

Under perturbation by a preceding floating high tone, we saw earlier (see (8a) above) that words with a LM tonal pattern change to HM. This is again illustrated in (29a-b) with the word for 'puddle' (Pike 1948: 79). As already mentioned, the linking of the floating high tone to the low-tone vowel (Rule 20a) causes the low tone to delink, since contour tones are banned in the language (Constraint 3).

(29) a. [mini → b. mini ⎯→ c. *mini

   (H) | L | (20a) | H | UAC | H | L

   But if the second vowel in the word is lexically toneless, the universal association convention (UAC) predicts that the delinked low tone should attach to it, incorrectly yielding the HL pattern in (29c).

Another problematic case arises with perturbing words whose last vowel is mid toned. Consider for instance the word for 'mountain', źuku (H), a noun with mid tones which was shown earlier (see (16a) above) to have a final floating high tone. Under our analysis, its lexical representation is as in (30a).

(30) a. źuku b. *źuku c. *źuku d. źuku

   (H) | | H | H | M

Assuming that the UAC applies whenever it can, one would expect such words not to exist in the language, for if they did, they would be immediately restructured by the automatic linking of the floating tone to one of the two free vowels, as in (30b) or (30c). Alternatively, one could assume on a principled basis that the UAC is preempted by more specific language-particular rules such as (20a) and (20b). This approach would correctly allow the floating high tone to link to a following morpheme, as exemplified in (16a) above. But the problem would remain when our example in (30a) is used before a pause. Again, the UAC predicts that the floating high tone should link to one of the two free vowels. It does not: M-Default (19) applies instead, yielding the phonetic representation in (30d).

In both (29) and (30), then, M-Default must preempt the universal convention linking floating tones to free vowels; but there is no rational basis for such precedence relation, since default rules are by definition principles of last resort. One possibility of course would be to assume that a language-specific rule deletes floating tones in Mixteco just in case other language-specific rules fail to assign
them to a mora; however, since the lack of phonetic realization of such floating tones can be more simply attributed to their unanchored status, their language-specific deletion looks like a devious way to bar the application of the UAC.

To conclude, the purported universal convention automatically associating a floating tone to a free tone-bearing unit must not apply in Mixteco, despite excellent opportunities. Unless one resorts to the suspicious language-specific deletion of floating tones, the implication of this situation for phonological theory is a follows. Either (i) the convention in question is not in fact part of Universal Grammar, as argued independently by Archangeli & Pulleyblank (1994) and by Hyman & Ngunga (1994), or (ii) it must be assigned the status of a parameter, which would happen to be set to 'off' in Mixteco.23

APPENDIX I: ON THE MH GAP

Words of the general shape CVV fall into two groups, depending on whether the sequence VV corresponds to a single melody (i.e. a long vowel) or two distinct melodies. Vocalically bimelodic CVV words are apparently rather rare. In combing through Pike (1948: Chapter VII), I found only half a dozen such items, compared to more than 40 vocalically monomelodic CVV words. In addition to lacking the LL pattern, as do all lexical words, words with a long vowel do not occur with the MH pattern. I checked all of Pike's writings and Mixteco texts listed in the references and could not locate a single counterexample to this generalization. It is unlikely that the absence of the MH pattern on long vowels in the available data is due to chance, since I found in Pike (1948: Chapter VII) at least two examples for each of the other seven tonal patterns; additionally, I did find the MH pattern as one possible tonal pattern on three of the few words with contiguous but melodically different vowels (siá 'to loosen'; ʔau 'cave, hole'; tiú 'uncle').24

In parallel with the case of CVV words, the great majority of CV?V words have identical vowels in both syllables; thus, in Pike (1948: Chapter VII), out of more than 20 words of the shape CV?V, I found only three with different vowels (Adâ'í 'to be poor', bâ'uí 'coyote', ñâ'ã 'sacred personage'). For the CV?V words containing identical vowels, I found at least one example for each logically possible tonal pattern, except for the generally banned LL pattern and the MH pattern also absent in words with a long vowel. Although I found no occurrence of a MH pattern on CV?V word with different vowels either, I surmise that this gap, at least as an underlying gap, is not real, but rather due to the paucity of available tokens. As examples (24b) and (25b) above illustrate, what can be established for certain from the available data is that as opposed to CVV words with different vowels, CV?V words with different vowels do not allow the derived MH pattern.

It is worth noting that as formulated, Constraint (6) does not differentiate between the two laryngeals [h] and [ʔ].25 Therefore the same kind of split with words containing a medial [h] is predicted, namely, no MH pattern if the two vowels are the same, but no such restriction if they are different. I was unable to find any CVhV example with the same vowel in both syllables,26 so the first half of this prediction remains to be tested. Regarding the second half of the prediction, one of the two CVhV words with different vowels which I found in Pike (1948: Chapter VII) does occur with the MH pattern (tuhi 'to scrape').27

It should finally be noted that in general, by contrast with CV?V words, CV?CV words with a medial consonant other than [ʔ] appear to have different vowels in their two syllables more often than not, an expected asymmetry assuming chance distribution of vowel quality. The MH pattern is clearly attested in either case, in
particular when the two vowels happen to be identical: e.g. źińi [from źinę 'insides'], bini [from kuni 'to know'], kiśi [from kusú 'to sleep'], haća 'to throw away', toađó 'much', kutū [from kutū 'nose'], sūtú [from sūtū 'priest'], źukū [from źukū 'brush'].

APPENDIX II: ON CVV WORDS AND RELATED MATTERS

CVV words with a long vowel and a MM tonal pattern, that is, in our terms, CVV words with an underlyingly toneless long vowel, may optionally change to HH, instead of HM, when preceded by a floating high tone (e.g. kée ~ kée; Pike 1944: 123-124, 1945b: 220, 1948: 80-81). I assume that Rule (20b), which links a floating high tone to a following toneless mora, may iterate and link the high tone to the next mora, just in case the two moras form a toneless long vowel. Pike's descriptions indicate that the extra linking is strictly limited to this context (a floating high tone and a toneless long vowel). Thus, he does not mention that the phenomenon occurs with any other closely related configurations, such as CVV words with a long vowel and an underlying LM pattern, or CV'T words with the same vowel and either a MM or LM pattern. Also, CVV words with a long vowel and an underlying HM pattern do not appear to have an optional pronunciation with a HH pattern.

Pike (1945a: 135; 1948: 80, note 5, 86) does mention two examples, shown in (i) below, where źuń 'rock' surfaces as źuń, and which Goldsmith (1990: 25) interprets as illustrations of the possible total propagation of a preceding floating high tone onto a CVV word with a long vowel and an underlying ML pattern (although ađežu 'food' demonstrably has a final floating high tone, as shown in (ii) below, there is actually no independent evidence that ađi, a form not found by Pike (1945a: 135) anywhere else, does).

(i) a. ađi - źuń 'gizzard'
    b. ađežu źuń 'rock-like food', i.e. 'thick or solid food'

In fact, the derived HH pattern observed here on the word for 'rock' has nothing to do with the presence of a preceding floating high tone of the sort discussed in this paper. Rather it is a tonal template commonly used in Mixteco to derive noun or verb modifiers (Pike 1944: 117, 135; 1945a: 133, 135; 1948: 82-87), especially when the modifier takes on a figurative meaning, as is clearly the case in the phrases in (i). Thus, (ib) can be contrasted with (ii), where the word for 'rock' has a literal meaning and where the final floating high tone of the word for 'food' links to the word for 'rock' in predicted fashion (Rule 20b).

(ii) ađežu (H) 'food' + źuń 'rock' —> ađežu źuń 'food made out of rocks'

As illustrated in (iii), the HH tone at work in (i) seems to apply regardless of the segmental or tonal make-up of words.

(iii) kąągo 'crooked' —> kąągó 'crookedness'
    śini 'head' —> śini 'in vague relationship to the head'
    sùći 'child' —> sùći 'young'
    tázán 'comrade' —> tázán 'togetherness'

On the basis of the examples in (iv) also involving the word for 'rock', Goldsmith (1990: 25) additionally argues that a floating high tone may link if and only if it is preceded by an anchored tone. He thus explains why the word for 'rock', itself endowed with a final floating high tone, as shown in (iva) by its action on the underlying enclitic -de 'his', fails to be perturbing in its źuń garb, as shown in (ivb) (Pike 1948: 80, note 5): the floating high tone is preceded by a delinked low tone.
(iv) a. źuù - dé 'his rock'
   b. ẑdi - źúù - de 'his gizzard'

But it is not true that a floating high tone needs to be preceded by an anchored tone in order to be able to associate to a mora, witness the continutive morpheme illustrated in (8a) above, which is not in the least incapacitated when phrase-initial. The explanation for why źúù does not perturb the enclitic -de is more likely related to the fact that mid-tone enclitics are quite generally not affected by a preceding otherwise perturbing word with a HH (or HL) tonal pattern (Pike 1948: 91). Note also that one could not argue that the HH template replaces all of a word's underlying tonal pattern, including a final floating high tone (for example, that źúù (H) is changed to źúù rather than źúù (H)), since, as shown in (vi), the word sučí (H) continues to be perturbing when endowed with such a pattern (Pike 1948: 87).

(v) Underlying morphemes: źani 'brother', sučí (H) 'child', źn 'that'
   a. źani + sučí (H) + źn  →  źani sučí-ún 'that younger brother'
   b. źani + sučí (H) + źn  →  źani sučí-ún 'the brother (of) that child'

(vi) a. [źni + sučí (H)] + [sučí (H) + sučí (H) + źn]
   'the younger brother' (of) 'that young child'
   b. źni sučí sučí sučí-ún
   c. źni sučí // sučí sučí-ún

In contrast to (vb), (va) illustrates the use of the HH template on the word for 'child' used as an adjectival modifier for the word for 'brother'. Since the word for 'that' is high toned underlyingly, whether a floating high tone actually precedes it or not is immaterial. The larger construction in (via) is analogous to (vb), with its two main constituents each analogous to (va). Its two possible pronunciations in (vib-c) show that even with its HH pattern, the first occurrence of the word for 'child' is perturbing in the usual way when no pause intervenes.

For the sake of completeness, I note finally in (viia) an isolated case where the continutive morpheme changes a verb's ML pattern to HH. Pike described this change as a 'unique perturbation'; it is actually also found in the verb's reduced form in (viib), although not in its cognate in (viic), which undergoes the expected perturbation to MH (Pike 1944: 134, 137).

(vii) a. (H) + ẑačí 'to say'  →  ẑačí
    b. (H) + ẑaà 'to say'  →  ẑaà
    c. (H) + kačí 'to inform, to say'  →  kačí

NOTES

1 There are in addition a few words of the shape CV2CV, which, as we will see, exhibit some of the special properties of CV2V words (see (26) below). Interjections (e.g. bidáà 'indeed'; Pike 1948: 88) and Spanish borrowings (e.g. fábóor 'favor'; Pike 1948: 93) may also have a more complex structure. Finally, [s] can be found word-initially before voiceless stops (the first two examples might be bimorphemic, with [s] the causative prefix; cf. (9) below): e.g. stážan 'to light', 'to show', 'to insult' (Pike 1945a: 129), skááda 'to toss' (Pike 1945a: 133), stáà 'tortilla' (Pike 1945a: 137), stoo 'uncle' (Pike 1945b: 220).

2 Following Pike's practice in his IJAL articles, I represent high tones with an acute accent and low tones with a grave accent, mid tones being left unmarked (Pike 1948 marks these with a macron).

3 For details, see Appendix I.

4 CVV and CV2V words containing identical vowel melodies share another distinctive property. Reduction to CV is frequent in them, especially in proclisis.
Such shortening is very rare in other word types, and when it does occur, it is the first syllable, rather than the second, which tends to be lost (Pike 1944: 128, 131, 132; 1945b: 223; 1948: 10, note 18. For examples of the latter type of abbreviation, see Pike 1945a: 133; 1945b: 223). There is another parallel among the words affected by Constraint (6), albeit this time with the added participation of analogous words with different vowel melodies. Pike (1944: 115; 1948: 10, note 18) observes that vowel nasality affects both vowels of a word if 'the medial consonant is zero, or [h], or [ʔ]'. In other words, laryngeal consonants are transparent to nasality. Caveat: my classification of Mixteco [h] as a laryngeal consonant is based on the generally accepted phonetic value for the symbol 'h', but it might actually represent a velar approximant or fricative in Pike's notation: '[h] varies from little to considerable friction on the velum' (Pike 1944: 115).

5 I will not seek to account for the behavior of enclitics here, as a number of them seem to exhibit idiosyncratic properties in terms of the ways in which they are perturbed and in turn cause perturbations (Pike 1948: 89-92).

6 *sčj* actually has a final floating high tone underlyingly (Pike 1948: 87), but I suppress it here, since it has no phonetic consequences, being phrase-final (but see (vi) in Appendix II).

7 The continuative morpheme may also have a palatalizing influence (e.g. (H) + kaka 'to walk' yields hika), which Pike (1944: 123; 1948: 94, note 10) transcribed with a small raised 'y' clearly prefiguring today's notion of floating element. Pike (1944: 130) pointed out the existence of another floating morpheme (which he transcribed with a small raised 'n'), one with a coronal and nasal influence on verb stems and the semantic value of implying 'resultant duration of action'. For instance, thus prefixed, the verb kête 'to enter' becomes ndete 'to enter and remain'. Although the floating continuative and resultant morphemes occupy different preverbal slots, they may under the right circumstances have cumulative effects on a following verb stem, yielding in this case ndete 'to enter continuously or repeatedly and stay there hiding'. Surprisingly, Pike did not use a similar notation to express the concept of the floating high tone, resorting instead to the abstract symbols '(' (in Pike 1944) or '(b)' (in Pike 1948) to mark the 'raising influence' of tonally perturbing morphemes.

8 A more neutral way of putting it would be to say that the surfacing of the tonal patterns in (11a) concurrently satisfies the realization of the underlying tonal patterns of the words themselves and the realization of the floating high tone.

9 Recall that there are no lexical words with the LL tonal pattern (see Constraint 5 above).

10 As shown in (i), CVV words with different vowels behave in the same way.

   (i) ... (H) + žau 'hole' → žau (Pike 1944: 116; 1948: 81)

11 The word for 'to scrape' itself has a final floating high tone that affects the following clitic, changing its low tone to high.

12 Viewing mid-tone vowels as underlingly toneless might seem to breathe new life into an OCP account of why the MH pattern remains unchanged following a floating high tone. One could now argue that in the case of the change from LH to HH (e.g. súčj → súčć), the delinked low tone acts as an insulator on the tonal tier between the two anchored high tones, whereas in the case of MH remaining MH (e.g. kučj), there would be no intermediate tonal element between the two high tones. While this account would make it possible to extend the range of cases where the target of a floating high tone is the initial vowel of the next morpheme, it would still leave unresolved the case of CVCV words with a ML pattern, where the second
vowel is targeted (ML → MH). What makes an OCP account additionally doubtful is that words with the HH or LH tonal patterns may be perturbing, that is, include a floating high tone right after another (anchored) high tone; such configurations indicate that in Mixteco, floating tones must apparently be ignored in the computation of OCP violations.

13 The following example (Pike 1948: 82) shows that floating high tones only have access to the first morpheme on their right (the floating high tone here is the continuative morpheme; see (8a) above).

(i) (H) + na + kiku + ná → ná - kiku - ná, *na - kiku - ná 're' 'to sew' 'I' 'I am re-sewing' In (i), (20b) applies rather than (20a), even though on a purely phonological basis, one might have expected the floating high tone to skip the toneless vowel of na and anchor to the first vowel of kiku. The reason it does not skip the toneless vowel is that it does not have scope beyond the first morpheme to its right.

14 Again, recall that there is no LL tonal pattern (see Constraint 5 above).

15 The final 'n' here and elsewhere (see other examples in (23) and (25), in notes 19 and 26, and in Appendix II) denotes nasality on the preceding two vowels (see note 4 above and Pike 1944: 115; 1945a: 129, note 1; 1948: 10, note 18).

Although reproducing the canonical form expected from Pike’s descriptions, the actual example used here (k"áán → k"áán) may be questionable. It relies on an analysis of the word for 'yellow' which is different than that proposed in Pike (1948: 86). Following Pike (1945b: 222), I have assumed that the underlying form for this adjective is k"áán and that the form k"áán found in kete k"áán 'a yellow animal' (Pike 1948: 86) is due to the fact that kete contains a final floating high tone (Pike 1944: 129, 132). The LH tonal pattern apparently occurs very rarely on Mixteco long vowels; in all of Pike’s publications listed in the references, I was able to find it in only one other item, the onomatopoeia pään (Pike 1948: 93).

16 On the alternative pronunciation kee and related matters, see Appendix II.

17 Pike does not actually provide a translation for the verb ka'mù itself, but its combination with the form hitè 'shin bone' is given the meaning 'to walk with the knees bending slightly, or giving under one, at each step'. Note that the continuative morpheme responsible for the tone change here (cf. (8a) above) exhibits an additional, spirantization, effect on the verb, with the initial [k] turning into [h] (see note 7 above and Pike 1944: 123).

18 Recall that the other laryngeal consonant [h] (but see caveat in note 4 above) does not prohibit a floating high tone from attaching to a following low-tone vowel (see example 17c above).

19 As shown in (i) (Pike 1948: 78), an initial glottal stop does not prevent the attachment of a preceding floating high tone.

(i) kee (H) 'to eat' + ẑísò 'rabbit' → kee ẑísò 'the rabbit will eat'

I presume that as opposed to medial glottal stops, initial glottal stops are not contrastive and therefore not present underlingly. They simply occur on the surface as a default onset consonant. Cf. the borrowing ¿orá from Spanish hora [ora] 'hour' (Pike 1944: 116) and also the fact that in Pike’s transcriptions, an initial glottal stop is not consistently present at least in the case of the morpheme for 'one': øen vs. œen; see for instance Pike (1945a: 133; 1948: 80 vs. 1944: 118; 1945a: 131; 1945b: 219; 1946a: 22).

20 According to the present analysis, there is then a double reason why CV?V words with identical vowels change from ML to HL rather than MH when preceded
by a floating high tone: Constraint (6) and Constraint (27). Regarding the fate of an underlying MH pattern in CV2V words with different vowels and in CV1CV words (should such a tonal pattern turn out to exist on these words; see Appendix I), the analysis predicts that a preceding floating high tone should trigger a change to HH.

21 Which vowel would receive the floating tone would depend on whether the left-to-right or the right-to-left option is selected for the universal association convention.


23 I view neither of these implications as a positive outcome for phonology: (i) the UAC under fire has served insightful purposes in countless analyses, and (ii) the 'parameterization of universal conventions' is a contradiction in terms indicating that there are in fact no universal conventions. However, since my discussion of the Mixteco data has been conducted within a rule-based framework (with a sprinkle of constraints added on), these implications only affect this particular theoretical approach. I provide in Tranel 1995 an alternative analysis within Optimality Theory which yields a preferable account of the data without parallel theoretical drawbacks.

24 siá also occurs as síá, in a rare case of free alternation (Pike 1948: 27); šáuí is derived from underlying šáuí (see note 10 and example (24b) above); and tíú is also attested as tíú (Pike 1948: 94), with the general direct address HL tonal pattern (Pike 1948: 87).

25 See caveat in note 4 above regarding the classification of [h] as a laryngeal consonant.

26 The word túñun 'word' (Pike 1945b: 221; 1946a: 22) occurs in a related usage as túñun in a song (Pike 1946b: 131), but still without illustrating the MH pattern.

27 This MH pattern is actually derived from underlying ML (see example (17c) above). The other word is kahí 'to eat', also attested with the (derived) pattern HM in Pike (1944: 119).

28 Underlying forms are provided here within the square brackets if I know them to differ from the phonetic forms used as illustrations. These data were culled from Pike (1944; 1945a,b; 1948), including the Mixteco texts found in these references.

REFERENCES


HYMAN, LARRY M., and ARMINDO NGUNGA. 1994. On the non-universality of
tonal association 'conventions': Evidence from Ciyao. Phonology 11.25-68.
KIPARSKY, PAUL. 1973. 'Elsewhere' in phonology. A Festschrift for Morris
Halle, ed. by Stephen R. Anderson and Paul Kiparsky, 93-106. New York:
Holt, Rinehart and Winston.
KOUTSODAS, ANDREAS, GERALD SANDERS, and CRAIG NOLL. 1974. On the
Theory, ed. by John A. Goldsmith, 444-475. Oxford and Cambridge, MA:
Blackwell.
PIKE, KENNETH L. 1944. Analysis of a Mixteco text. International Journal of
PIKE, KENNETH L. 1945a. Tone puns in Mixteco. International Journal of
American Linguistics 11.129-139.
PIKE, KENNETH L. 1945b. Mock Spanish of a Mixteco Indian. International
PIKE, KENNETH L. 1946a. Another Mixteco tone pun. International Journal of
PIKE, KENNETH L. 1946b. The Flea: Melody types and perturbations in a Mixteco
song. Tlalocan 2.128-133.
PIKE, KENNETH L. 1948. Tone Languages. Ann Arbor: University of Michigan
Press.
SELKIRK, ELISABETH. 1990. A two-root theory of length. University of
STERIADE, DONCA. 1987. Locality conditions and feature geometry. Proceedings
of the 17th Meeting of the North East Linguistic Society, 595-618. Amherst:
Graduate Linguistic Student Association, Department of Linguistics, University
of Massachusetts.
Phonologie on Current Trends in Phonology, Royaumont, France, June 19-21,
1995.
VERBAL PREFIXATION IN MALAY:
RECONFIGURING PARADIGMATIC RELATIONS
Lionel Wee
UC Berkeley

1 INTRODUCTION. Malay has four verbal prefixes, which are assumed to have the following properties:

(i) they mark voice
(ii) they mark volitionality
(iii) they form a paradigmatic set

However, problems arise when one tries to be explicit about the ways in which these various assumptions are related. This goal of this paper is to show how these problems can be dealt with, and the assumptions retained.

2 VOICE AND VOLITIONALITY. In (1), we see three of the four prefixes. These three are usually assumed to form a paradigmatic set of voice markers (Benjamin 1993; Hassan 1974).

(1)

a. meN-1: active voice
   Ali meN-pukul John
   Ali meN-hit John
   Ali hit John

b. di-: passive voice
   John di-pukul (oleh Ali)
   John di-hit (by Ali)
   John was hit (by Ali)

c. ber-: middle voice
   Siti ber-dandan
   Siti ber-dress up
   Siti dressed up (herself)

In (2), we see the fourth prefix, ter-, which marks non-volitionality (Wouk 1980; Winstedt 1927).

(2) ter-: non-volitionality
   Ali ter-pukul John
   Ali ter-hit John
   Ali unintentionally hit John
The presence of ter- raises the following question: what is the relationship of ter- to the paradigm? It is claimed that all four prefixes form a paradigmatic set (Benjamin 1993; Hassan 1974). The main reason for this claim is that all four prefixes are mutually exclusive. The alternative would be to simply stipulate that Malay has at most a single slot for a prefix. This is surely undesirable since there is no independent evidence (that I am aware of) that Malay has any kind of slot morphology.

But the assumption that the four prefixes form a paradigmatic set raises a problem: there is no apparent reason why a marker of voice and a marker of non-volitionality cannot co-occur. So, how do we rule out a construction like the one shown in (3), for example, where the prefixes meN- and ter- are both attached to a verb?

(3) *meN-ter-VERB

Before attempting to deal with these problems, I will briefly examine two recent attempts to be more explicit about the relationships among the prefixes. We will see that these attempts embody most of the problems already mentioned.


(4) meN- 'active (transitive) voice' or 'actor-focus'
    di-   'passive (transitive) voice' or 'patient-focus'
    ber- 'active (intransitive)' or 'middle voice'

The problem here is that the ter- prefix is not shown in (4), even though Benjamin clearly makes the claim that "(w)hatever the source and possible earlier meanings of ter-, it has ended up in the modern formal language as a component element of the paradigmatic set consisting of the 'voice/valency'..." (p363).

And as we have already noted, even if Benjamin were to treat ter- as a non-volitionality marker, this still doesn't account for its inability to co-occur with the other prefixes.


(5) meN- 'active voice (volitive)'
    di- 'passive voice (volitive)'
    ber- 'middle voice'
    ter- 'active/passive voice (non-volitive)'

Notice that in Hassan's version, the notion of volitionality plays a prominent role. Notice also that ter- is presented as marking both active as well as passive voice, in addition to marking non-volitionality. This active/passive nature of ter- is based on a comparison of (6) with (2).

(6) Kaca itu ter-pijak oleh Ali
    glass the ter-step by Ali
    The glass was stepped on by Ali

Unfortunately, Hassan's version is still problematic for the following reasons:
(i) Since all the other prefixes are marked for both voice and volitionality, why is ber- not given a volitionality value?

(ii) ter- is given two voice specifications; so why are we not given a five-way distinction instead? That is, what prevents Hassan from giving us the five-way paradigm shown in (7)?

(7) meN- 'active voice (volitive)'
    di- 'passive voice (volitive)'
    ber- 'middle voice'
    ter- 'active voice (non-volitive)'
    ter- 'passive voice (non-volitive)'

(iii) There are empirical problems with Hassan’s characterization of meN- as being 'volitive' in nature; meN- actually seems to be neutral or unspecified for volitionality. This is shown in (8), where a meN- construction has no problems co-occurring with either an adverbal of intentionality or unintentionality.

(8) a. Ali meN-pukul John dengan sengaja
    Ali meN-hit John with intention
    Ali intentionally hit John

b. Ali meN-pukul John dengan tidak sengaja
    Ali meN-hit John with NEG intention
    Ali unintentionally hit John

I suggest that the problems for both Benjamin and Hassan result from their subscribing to the same set of unquestioned assumptions. They both seem to be assuming that:

(i) since there are four phonologically distinct prefixes, there must therefore be a four-way contrast, and

(ii) this contrast is primarily one of voice.

The result is that they first start with a voice contrast involving the prefixes meN-, di-, and ber-. Then, they run into difficulty trying to fit in the fourth prefix, ter-.

What I propose to do is to reject these assumptions, and instead, begin by taking seriously the fact that ter- marks non-volitionality. I suggest that the verbal paradigm can be reconfigured into a three-way volitionality contrast, shown in (9). Voice contrasts are actually internal to each volitionality category.

(9) unspecified volitional non-volitional
    Vol [-] Vol [+] Vol [-]

Before proceeding, let me be clear that I use the notations 'vol [-], vol [+], vol [-]' only as convenient shorthand. They are NOT meant to indicate anything like discrete features within any kind of formal framework. As we shall see, a proper
understanding of what is going in the Malay verbal system (as with any other language) requires an appreciation of subtle semantic/pragmatic factors that the speakers make use of in their coding choices. The reification of features, on the other hand, runs the risk of disembodying the grammar from its speakers, and tends to reduce understanding to the mere formulation of algorithms (see Lakoff 1987; Langacker 1987, among others, for more discussion on the errors of an 'objectivist' view of grammar).

5 RECONFIGURING THE PARADIGM. The evidence for the volitionality categories comes from two sources:

(i) the lexical semantics of the stems that the prefixes attach to, and

(ii) the possibility of adding adverbials that explicitly indicate the volitionality of the agent.

(10) shows that the prefix meN- can attach to verbs that indicate either volitional or non-volitional behaviour, while ber- and ter- are restricted to volitional and non-volitional verbs respectively.

(10) Vol [ ] Vol [+ ] Vol [-]

meN-tari 'dance'
meN-curi 'steal'
meN-kantuk 'fall asleep'
meN-tangis 'cry'
ber-kerja 'work'
ber-bual 'discuss'
ber-baring 'lie down'
ber-lari 'run'
ter-gelincir 'slip'
ter-kejut 'startled'
ter-jatuh 'fall'
ter-cengang 'astonished'

Consider what happens when adverbials are added. We already saw in (8) that meN- can easily co-occur with adverbials of intentionality or unintentionality, which follows from the fact that meN- is actually vol [ ]. For convenience, (8) is repeated here as (11).

(11) meN-: vol [ ]

a. Ali meN-pukul John dengan sengaja
   Ali meN-hit John with intention
   *Ali intentionally hit John

b. Ali meN-pukul John dengan tidak sengaja
   Ali meN-hit John with NEG intention
   *Ali unintentionally hit John

What is particularly interesting is the way these adverbials interact with the prefixes ber- and ter-. As (12) shows, because ber- is already vol [+], an adverbial of intentionality is considered redundant (12a), while an adverbial of unintentionality is contradictory (12b).
(12) ber-: vol [+]

a. ?Ali ber-cukur dengan sengaja
   Ali ber-shave with intention
   *Ali intentionally shaved
   [redundant]

b. *Ali ber-cukur dengan tidak sengaja
   Ali ber-shave with NEG intention
   *Ali unintentionally shaved
   [contradictory]

Exactly the converse happens with ter-, allowing us to conclude that ter- is indeed vol [-]. Thus, (13a) is contradictory since the adverbial indicates intentionality, while (13b) is redundant since the adverbial is one of unintentionality.

(13) ter-: vol [-]

a. *Ali ter-tembak orang itu dengan sengaja
   Ali ter-shoot person the with intention
   *Ali intentionally shot that person
   [contradictory]

b. ?Ali ter-tembak orang itu dengan tidak sengaja
   Ali ter-shoot person the with NEG intention
   *Ali unintentionally shot that person
   [redundant]

What about the prefix di-? (14) shows that the subject of a di- construction, John, can either have been a volitional or non-volitional patient. Thus, di-, like meN-, is unspecified for volitionality.

(14) John di-pukul (oleh Ali)
    John di-hit (by Ali)
    John was hit (by Ali)

This means that both meN- and di- are members of the same category where volitionality is unspecified, so that internal to this category is an active/passive voice contrast.

We can therefore ask if there are any voice contrasts internal to the other categories also. Recall from (5) that Hassan recognizes that ter- has both active and passive variants. This means that within the category of vol [-], there is an active/passive distinction. In fact, the same can be said for the category of vol [+], represented by ber-. Thus, compare (15b) with (1c), which is repeated here as (15a).

(15) Siti ber-dandan
    Siti ber-dress up
    Siti dressed up (herself)

b. Kain itu sudah ber-jahit
    cloth the already ber-sew
    The cloth has already been sewn
The reconfigured verbal paradigm is shown in (16).

(16) The Verbal Paradigm In Malay (Reconfigured)

\[
\begin{align*}
\text{Vol} [\ ] & \quad \text{Vol} [+]^2 & \quad \text{Vol} [-] \\
\text{meN-: active} & \quad \text{ber-: active} & \quad \text{ter-: active} \\
\text{di-: passive} & \quad \text{ber-: passive} & \quad \text{ter-: passive}
\end{align*}
\]

(17) abstracts out a schema for the Malay verbal prefix, showing that each prefix represents a volitionality category, and is also marked for voice.

(17) Schema For The Malay Verbal Prefix

\[
\begin{align*}
\text{Vol} [X] & \quad (\text{where 'X' is unspecified, volitional, or non-volitional}) \\
\text{active} & \\
\text{passive}
\end{align*}
\]

The analysis proposed here results in a paradigm structure that successfully integrates all four prefixes. Volitionality emerges as the crucial feature that distinguishes the three categories. The voice distinction that was assumed to be the primary paradigmatic contrast turns out to be a secondary feature internal to each category.

6 A POSSIBLE OBJECTION. One might object to the paradigm in (16) on the grounds that it still doesn't account for the mutual exclusivity of the prefixes. The objection might be as follows: Given that meN- belongs to a category vol [ ], what prevents this category from getting a volitionality value from a prefix like ter-? After all, we know from (11) that a meN- construction can be 'further specified' via an adverbial. What's to prevent meN- from receiving this further specification from the prefix ter- instead of an adverbial? In other words, the skeptic will claim that unless this objection can be dealt with, we still have no grounds for ruling out a construction like the one in (3).³

In fact, there are two possible ways to deal with this objection. One way would be to simply concede that, yes, there is no apparent reason why the vol [ ] of meN- cannot receive further specification from ter-. The inability of meN- and ter- to co-occur will then not be due to the volitionality values, but to the voice specifications. It is important to bear in mind that the prefixes mark both volitionality and voice. Since meN- only marks active voice, and ter- either active or passive, we can rule out (18a) on the grounds that the voice values conflict. We can also rule out (18b) on the grounds that having two phonologically distinct markers for the same semantic category \textit{within the same word} amounts to a case of multiple exponence.

(18)

\[
\begin{align*}
\text{a.} & \quad *\text{meN-ter-VERB} \\
& \quad \text{active-passive-VERB} \\
\text{b.} & \quad *\text{meN-ter-VERB} \\
& \quad \text{active-active-VERB}
\end{align*}
\]
As Peterson (1994:98) notes, multiple exponence is a marked phenomenon; "(i)n the unmarked case, ME (multiple exponence, LW) of a feature will not be required." The crucial point is that we are not compelled to justify cases where multiple exponence is absent, rather it is claims of multiple exponence that need to be carefully examined. And in fact, it appears that the most uncontroversial cases of multiple exponence tend to be restricted to negation. For example, in Luganda, the verb stem contains two phonologically distinct markers of negation, one involving segmental material and one involving tone.

Bill Weigel (p.c.) points out that there is good reason why negation should be prone to multiple exponence; the presence/absence of negation makes a difference to the content of a clause. If a marker of negation has undergone grammaticalization to a point where only its tone remains, then speakers may feel it necessary to bring in a new (and phonologically more substantial) marker to aid in restoring the salience of negation. On the other hand, voice, and in particular, active voice, is usually unmarked since it is usually inferable from the valence of a verb. In the light of all this, we can rule out multiple exponence of active voice markers.4

A second and stronger way to deal with the objection would be to point out that the entire objection is based on a highly questionable assumption in the first place. It assumes that there is no significant difference between encoding non-volitionality in the form of an adverbial, and encoding it in the form of the prefix ter-. This kind of assumption only arises if we make the mistake of treating parts of a grammar as being equivalent simply because they appear to have the same truth-values or bear similar 'propositional content'. This kind of mistake is reminiscent of attempts to treat the passive as a purely syntactic variation on the active. But as numerous works have shown us (DeLancey 1981; Langacker 1987; Rice 1987; Shibatani 1985; Van Oosten 1986), the passive is a marked coding choice, carrying a variety of semantic/pragmatic nuances; it presents an event from the perspective of the terminal phase, de-focuses the agent, and thus indicates that the prototypical overlapping of agent and topic does not apply.

Coming back to the case at hand, as Talmy points out, in understanding the relationship between form and meaning, it is necessary to pay attention to salience, defined as "the degree to which a component of meaning, due to its type of linguistic representation, emerges into the foreground of attention or, on the contrary, forms part of the semantic background where it attracts little direct attention ... a semantic element is backgronded by expression in the main verb root or in any closed-class element (including a satellite - hence, anywhere in the verb complex). Elsewhere, it is foregrounded" (1985:122).

The following examples are Talmy's:

(19)

a. Last year I went to Hawaii by plane
b. Last year I flew to Hawaii

Talmy notes that the sentences in (19) are "virtually equivalent in the total information that they convey, but they differ in that the fact of transit by air is
pivotal" in (19a) by virtue of the adverbial, "whereas it is an incidental piece of background information" in (19b) "where it is conflated within a verb." Likewise, in the following sentences, the notion of non-volitionality is either backgrounded within the verb (20a), or foregrounded as an adverbial (20b). In (20c), the speaker does not commit himself/herself to the volitionality of John's action.

(20)

a. John ter-pukul Ali
   John ter-hit Ali
   John unintentionally hit Ali [non-volitionality backgrounded]

b. John meN-pukul Ali dengan tidak sengaja
   John meN-hit Ali with NEG intention
   John unintentionally hit Ali [non-volitionality foregrounded]

c. John meN-pukul Ali
   John meN-hit Ali
   John hit Ali [volitionality unspecified]

The speaker therefore has to decide if s/he wants to background the volitionality of the action or not. Notice that if the speaker wants to foreground the volitionality, this cannot be done with a prefix; it must be done with an adverbial. And as we have already seen, the only prefix that will co-occur easily with an adverbial is meN-. Thus, choosing meN- either allows the speaker to remain uncommitted as to the volitionality of the action, or to foreground it. Choosing ter- only allows it to be backgrounded.

This actually leads to the interesting possibility (discussed further in Wee 1994b) that the three volitionality categories do not form a flat structure. Rather, they might be hierarchially related as shown below.

(21)    volitionality
       unspecified  backgrounded
       vol [-]  vol [+]  vol [-]

Thus, the contrast between the volitionality categories is not simply a formal or structural one. They reflect choices that Malay speakers can and must make in deciding how to code the volitionality of an action.

7 SOME DIACHRONIC CONSIDERATIONS. I want to end on a diachronic note, by speculating on how the system of prefixes might have developed. According to Teeuw (1959:145), there was no evidence of ter- in Old Malay. There were, however, prefixes such as mam-, ni-, and mar-. These prefixes are respectively assumed to be cognate with the prefixes meN-, di-, and ber- (Coedes 1930; Hopper 1979). The situation is summarized in (22).
(22) Verbal prefixes in Old Malay

meN-  (< manN-)
di-   (< ni-)
ber-  (< mar-)

So, while we can be fairly certain that among the four prefixes, ter- entered the language last, it is unclear in what order the other three prefixes entered the language.

As a hypothesis, let's simply assume that mar- entered the language after maN- and ni- . This is shown in (23).

(23) One Possible Order In Which The Prefixes Entered The Language

Stage one:    maN-, ni-
Stage two:    mar-
Stage three:  ter-

We can now speculate: it is possible that at stage one, there is no volitionality distinction. Both maN- and ni- represent active and passive variants.

At stage two, mar- enters the language to mark volitional actions. It therefore forms a volitionality category distinct from maN- and ni-, which are now seen as members of a volitionality unspecified category. This gives the vol [ ] category an internal voice contrast. By analogy then, mar- develops active/passive variants as well.

When ter- turns up at stage three, the schema is fairly well 'entrenched' (to use a term from Ronald Langacker). Because of conflicting volitionality values, ter- is unable to combine with ber-. As such, in order to meet the requirements of the schema, ter-, too, develops active/passive variants.

The result is what we see today: a tripartite volitionality system involving four phonologically distinct prefixes.

NOTES

1. The prefix ends in an underspecified nasal which assimilates to the place of the initial consonant of the stem. If the stem begins with a vowel, the nasal appears as a velar. There are some complications. If the initial consonant of the stem is voiceless, it gets deleted. And if the stem begins with an /s/, the nasal appears as a palatal. This is a historical reflex if we assume that the /s/ was originally a palatal stop or fricative (Adelaar 1992:106). In this paper, I will present the prefix separate from the stem in all my examples so that with a stem like tangis 'cry', instead of writing menangis, the form will be meN-tangis. The reason for this is clarity of exposition since it will allow the stem in each case to be easily identified.

2. The fact that ber- marks active and passive voices should not be seen as a rejection of the claim that it is a middle. As Kemmer (1993) has shown, the middle
is actually a complex category defined in conceptual terms, namely, a low degree of event elaboration. In effect, middles display limited transitivity levels, where transitivity is understood as a gradient phenomenon (Hopper and Thompson 1980; Rice 1987). I discuss elsewhere in detail the interaction of volitionality and transitivity in the Malay verbal system (see Wee 1994b).

3. I thank Paul Kay for bringing this to my attention.

4. Jespersen (1924:333ff) makes a similar point about negation, though he does not explicitly contrast it with other grammatical categories like voice.

5. Not surprisingly, disagreements exist. There are claims that di- actually developed from a preposition di that eventually came to replace ni-. Also, it's been speculated that mar- was originally a Batak borrowing rather than a cognate of ber-. See Adelaar 1992 and references therein for a useful discussion of these various positions.

SELECTED REFERENCES


Understand in Conceptual Semantics

Rebecca Wheeler
Independent Scholar

O. INTRODUCTION. This paper presents a semantic analysis of the polysemous verb understand (Wheeler 1989, 1993) which (i) specifies the meaning of each sense, (ii) explains the relationship between the senses, and (iii) shows why these and no others comprise the meanings of understand. Further, this analysis explains aspects of the syntax and pragmatics of the English verb understand.

Understand poses a particular and even strident methodological challenge for lexical semantics. Denoting an interior state, this verb of cognition is not backed by a chorus of observational grounding to help motivate the verb’s sense statement. The linguist can not look out in the world to see, hear, or touch an instance of understanding to suggest aspects of its meaning. Instead, the analysis of the meaning of such verbs proceeds (as it should) a cappella.

To diagnose the semantics of understand, I rely on syntactic patterning as evidence of semantic content. In particular, I turn to its syntactic alternations (Levin 1993). In variation on Levin’s theme, instead of examining alternations in which understand itself participates, I examine the syntactic alternations in which the direct objects (and adjuncts) of the senses of understand participate. These alternations inform us of the semantics accruing to the position of object argument to understand. I This information will be shown to constitute the heft of the lexical conceptual structure (LCS) of understand.

I represent my findings within the framework of Conceptual Semantics (Jackendoff 1983, 1990). In doing so, I make no claim as to the innateness or the universality of the semantic components proposed.

I begin by summarizing earlier work delineating three senses of the English verb understand (Wheeler 1989, 1993, ms).

1. THE VERB UNDERSTAND -- ITS SENSES. Understand is an interior state verb; specifically, a verb of cognition. Like 75% of the more than 240 monolexemic verbs of cognition in English, understand is a two-place, transitive predicate. It occurs productively in the following syntactic frames:

(1) a. NP1 V NP2
    b. NP1 V wh- + S
    c. NP1 V S'
    d. NP1 V NP2 + adjunct/complement

Examples (2) - (5) illustrate these patterns, respectively, and also list typical object arguments.

(2) Marie understands NP2.
   the problem, the situation, the issue, the idea, the news, the factor involved, the importance of dealing with the way the system was structured, the wisdom of the decision, the patterns that form in their own behavior, the new organization and its requirements, the psychology of the project members, the company's customs, each business well, the market,
the instructions, the scorecard at her company, the invention's nuances and history

(3) They understand \( wh + S \)
how to win the game, why it wasn’t managed better,
what is wrong with the old one, how vital good management is
to their company

(4) a. They understand that their progress in the company depends on the results they achieve, not on their eloquence in meetings.

b. Understand that big machines, separated equipment, and long conveyor systems disconnect people, obscure opportunities for merging processes, and result in divided accountability!

c. I understand that ever more techno-business is moving into Utah.

(5) a. By poetical colors, the neoclassicist understands words, elegant phrases, figures of speech [Webster's Third New International Dictionary].

b. I understand the phrase literally.

c. I understand her suggestion as a complaint.

d. I understand her suggestion to be sarcastic.

In Wheeler 1989 and 1993, I explored whether understand shows restrictions in the syntactic environments in which each subcategorial frame could occur. What proved distinctive was the ability of understand to occur under main verb negation, in imperative form, or intransitively. Accordingly, I delineated use classes reflecting these distinctions. Accruing to each class is one or more characteristic lexical substitutes for understand.

Class I (COMPREHEND):
Subcategorization: \[ \underline{\text{NP/wh- interrogative}} \]
Syntactic distribution: \( \sqrt{\text{neg; } *\text{imperative}^3; \sqrt{\text{intransitive}}} \)
Lexical substitutes: see, get it, grasp, comprehend

Class II (REALIZE):
Subcategorization: \[ (that) \ S \]
Syntactic distribution: \( \sqrt{\text{neg; } \sqrt{\text{imperative; } \sqrt{\text{intransitive}}} \}
Lexical substitutes: realize

Class III (READ + MANNER):
Subcategorization: \[ \underline{\text{NP X; } \{X=adjunct/complement}\}} \]
Syntactic distribution: \( *\text{neg; } \sqrt{\text{imperative; } *\text{intransitive}} \}
Lexical substitutes: take, interpret, read, view

Thus, I have proposed an operational definition of sense: each sense of a lexeme will exhibit a characteristic pairing of

i) subcategorization frame and

ii) syntactic distribution potential (ability to occur intransitively, under main verb negation, in imperative mood, etc.).

The senses so identified, will manifest distinct lexical substitutes.
2. THE ALTERNATIONS OF UNDERSTAND. In treating the semantics of understand, I begin with the COMPREHEND reading, the most frequent sense of this verb. This reading occurs in the frame [NP1 V NP2/wh - S] [e.g. She understands the problem/the news/the situation, etc.). I will demonstrate that the NP2 argument to understand exhibits syntactic distribution characteristic of arguments selected by verbs of material composition. This patterning is key to the semantics of the COMPREHEND sense of understand.

Verbs such as consist, constitute, comprise, form, and make occur in the frame [NP1 V NP2] and denote aspects of material composition. Jackendoff observes that ‘sentences describing composition always involve a relationship between a whole and its parts’ (1990:120). Verbs of material composition differ according to whether they select the whole or parts as NP1. For example, consist and constitute are mirror-images of each other: consist selects as NP1 an argument denoting a whole, and as NP2, an argument denoting parts to that whole [see (6)a]. The part/whole relations are reversed for constitute/comprise/form/make as seen in (6)b. These verbs select as NP1 an argument denoting parts and as NP2, an argument denoting the whole.

(6) a. The triangle consists of three lines.
   b. Three lines constitute/comprise/form/make a triangle.

The NP2 argument of understand occurs as argument to verbs of material composition. In particular, it occurs in the position of arguments denoting a whole in a part/whole relationship [see (7) - (8)]. Thus, the object argument to understand occurs as NP2 to constitute, or comprise, and as NP1 to consist:

(7) a. Storing important documents constitutes/comprises the purpose of the table.
    b. His father’s coming home constitutes/comprises the news.
    c. That the wires don’t connect constitutes/comprises the problem.

(8) a. The purpose of the table consists in storing important documents.
    b. The problem consists of the wires not connecting.

I conclude, therefore that the COMPREHEND sense of understand selects as direct object, a whole in a part/whole relationship. As such, the conceptual structure of this reading will reflect the semantics of predicates of material composition.

3. THE CONCEPTUAL SEMANTICS OF MATERIAL COMPOSITION

3.1 THE SEMANTICS OF VERBS OF MATERIAL COMPOSITION. Prefatory to treating the semantics of understand, I step back and describe the semantics of material composition within Conceptual Semantics.

Jackendoff integrates the field of material composition

by subsuming it with Identification under a supercategory that might go under the name Character ... Predicates of Character tell about the object itself: what category it belongs to and what properties it has (Identification), and what it is made of (Composition) (1990:118).
Stative predicates of composition are analyzed as encoding 'a variety of the function BE, with a Theme and a "reference object."' (Jackendoff 1990:119). As standard in Conceptual Semantics, the function is annotated with a semantic field feature, here Comp, signifying Composition. Further, the field feature supports a diacritic reflecting whether the verb of material composition selects an NP1 denoting a whole in a part-whole relationship, \(\text{BE}_{\text{Comp}^+}\), or a part in a part-whole relationship, \(\text{BE}_{\text{Comp}^-}\). The resulting field structure is represented as in Figure 1 (adapted from Jackendoff 1990:118).

![Figure 1](image)

The Lexical Conceptual Structures (LCS) corresponding to verbs of material composition in (6)a-b are given in the bottom lines of (9)a-b (Jackendoff 1990:119-120).

(9)  

a.  [NP1 consist NP2]: NP1 denotes whole; NP2 denotes parts  
The triangle consists of three lines.  
\[-\text{DIR}  
\text{Sit BE}_{\text{Comp}^+}([\text{TRIANGLE}], [\text{Place AT[3 LINES]]})\]

b.  [NP1 comprise NP2]: NP1 denotes parts; NP2 denotes whole  
The three lines constitute/comprise/form/make a triangle.  
\[-\text{DIR}  
\text{Sit BE}_{\text{Comp}^-}([\text{3 LINES}], [\text{Place AT[TRIANGLE]]})\]

The LCS of (9)a states that consist encodes an undirected situation (a state); in particular, consist encodes a relation of composition, \(\text{BE}_{\text{Comp}^+}\), selecting as first argument a whole and as second argument, parts of that whole. The LCS of (9)b states that comprise encodes an undirected situation (a state), specifically, a relation of composition, \(\text{BE}_{\text{Comp}^-}\), and selects as first argument the parts to the whole named in second argument position.

3.2 Object argument to understand encodes semantics of material composition. Extending this analysis to understand and reflecting the syntactic and semantic patterning described in (6) - (8), I propose that the COMPREHEND reading encodes \(\text{BE}_{\text{Comp}^+}\) as seen in (10)b.
(10) a. I understand the purpose of the table.
    b. -DIR
       \([\text{Sit BE}_{\text{comp}}^+(\text{[THE PURPOSE OF THE TABLE]}, \text{[Place AT[Y]}]))\]

Example (11) renders the lexical entry for understand as in *I understand the purpose of the table*:

(11) understand
    V
    ___<\{NP\}{wh- + S}\rangle_k
    -DIR
    __\text{PERCEIVE} (\text{[NP]}, \text{[Sit BE}_{\text{comp}}^+(\text{[Sit/Thing PURPOSE OF THE TABLE]}_k, \text{[Place AT[Y]}]))\}

This lexical entry states that understand is subcategorized for two arguments, subject (unspecified, by convention) and object, \(<\{NP\}{wh- + S}\rangle_k\). The Lexical Conceptual Structure (LCS) says that the being denoted by NP, the subject of understand, perceives an undirected situation (a state); specifically, a relation of composition holding between a whole and parts. The function, PERCEIVE is adapted from Miller and Johnson-Laird (1976:115). Since understand’s subcategorized object is indexed to the first argument of the function \(\text{BE}_{\text{comp}}^+(\{X\}, \text{[AT[Y]}]))\), this LCS specifies that understand selects a direct object denoting a whole in a part/whole relation.

In effect, understand selects as semantic object the function CONSIST\((X, Y)\) where \(X\) is indexed to understand’s subcategorized direct object. Since the parts are not indexed to any syntactic structure in the lexical entry, these are encoded as semantically implicit. Thus, NP\(_1\) asserts they perceive the whole. The LCS (i) presupposes the existence of the parts [see (12)] and (ii) entails that NP\(_1\) knows the identity of those parts [see (13)]. Note that in (13), *Storing important documents* fills the NP\(_1\) argument position to constitute, the position corresponding to parts in a part/whole relationship. An asterisk in (12) - (13) marks semantic infelicity.

Existence of parts is presupposed
(12) a. I understand the purpose of the table.
    b. I don’t understand the purpose of the table.
    c. *There is nothing that constitutes the purpose of the table.

Presupposition: Something constitutes the purpose of the table.

Knowledge of identity of parts is entailed
(13) A: I understand the purpose of the table.
    B: What is it? What \textit{is} the purpose of the table?
    A: (i) Storing important documents \textit{constitutes} the purpose of the table.
        (ii) *I don’t know (what constitutes the purpose of the table).
        (iii) *Something constitutes the purpose of the table.

Entailment: NP\(_1\) knows the identity of the parts.

This analysis is of consequence for our understanding of how meaning is distributed across lexical items in a sentence. The predicate PERCEIVE is fairly unelaborated semantically. Instead, the real semantics of understand resides in the semantics of its subcategorized object argument.
There is more to be said about the semantics of the COMPREHEND reading. The LCS for understand proposed thus far would incorrectly predict that (14) receives an interpretation as in (15) corresponding to the LCS of (16).

(14) She understands the book.

(15) She perceives that the book consists of parts (i.e. pages, cover, spine).

(16) understand
    V
    _<the book>_k
    -DIR
    [Sit PERCEIVE ([SHE i ], [Sit BEComp+
    ([Sit/Thing THE BOOK], [Place AT [PAGES, COVER, SPINE]]))]}

However, the actual interpretation of (14) is not (15) but as illustrated in (17).

(17) a. I understand what the author is saying.
    b. I understand the import of the book.
    c. I understand the point of the book.
    d. I understand why the book was written.


(18) Reagan thinks bananas.

(19) a. What is Kissinger’s favorite fruit?
    b. Reagan thinks (that Kissinger’s favorite fruit is) bananas.

By itself, example (18) is ungrammatical: the lexical entry of the verb think specifies a sentential object argument but occurs in (18) with a noun phrase argument. However, in the context of (19)a, (18) receives an interpretation as suggested in (19)b which satisfies the lexical constraints on think. (Sells 1985).

When an utterance violates the lexical specifications of a verb, hearers seek to construe an interpretation which does conform with those specifications. The lexical item coerces and a hearer construes a conforming interpretation. Thus, we can follow the trail of hearer construal, to induce what lexical specifications must be guiding that construal.

Returning to example (14), we see the telltale signs of semantic coercion: Example (14) manifests a noun phrase object denoting a concrete entity, but the interpretation in (17) shows either a wh - interrogative or an NP headed by abstract nouns such as import or point. If the hearer has construed an interpretation distinct from (15) and (16), then the LCS as written does not yet fully reflect our semantic competence of understand. Accordingly, I refine the ontological specifications on the X argument of BEComp + as in (20) to reflect the relevant structure of the construed utterance meaning:
(20) \[ \textit{Sit PERCEIVE ([NP]}, \textit{Sit BE}^{\text{Comp}} ([\textit{Sit X} \text{ Mat X}_k \text{ Place AT[Y]})]) \]

The conceptual category \([-\text{Mat}]\) (short for non-Material Entity) represents an extension of the category Material Entity proposed in Jackendoff 1992. As such, it captures the generalization that \textit{understand} selects as object an abstract entity (-Mat) or a situation (\textit{wh} - S).

4.0 EXPLAINING THE RANGE OF UNDERSTAND’S SENSES. The relationship between the senses of \textit{understand} lies in this verb’s structuring of the Character field. I have shown the COMPREHEND sense to encode the function \(\text{BE}^{\text{comp}}\). I will now demonstrate that the remaining senses of \textit{understand} encode the remainder of the subcategories of the semantic field of Character (see Figure 1).

4.1 THE REALIZE SENSE. The COMPREHEND reading of \textit{understand} selects an object argument encoding a whole in a part/whole relationship. We have seen that verbs of material composition (\textit{consist vs. comprise}) differ by whether they select a whole or a part as subject. Since two classes of verbs contrast in this fashion, it is reasonable to inquire whether distinct senses of a single verb could contrast in the same way. I argue that they do and that it is precisely this contrast which adheres between the COMPREHEND and the REALIZE readings of \textit{understand}.

In the REALIZE sense, \textit{understand} occurs in the frame [NP1 V that + S]. \textit{Understand} is instantiated in this frame in (21) [see also (22)].

(21) Why should it be difficult to agree on objectives? Doesn’t everyone \textit{understand} that \textit{CARE} is in business to help poor people overseas? That the Salvation Army helps the homeless? That the Girl Scout program fosters goodwill and socially constructive attitudes? The problem is that these broad statements of purpose are not objectives at all. They are only concepts. And they certainly do not encourage the measurement of progress. (The Harvard Business Review Jan.-Feb., 1987:14, emphasis added).

(22) Doesn’t everyone understand
a. that \textit{CARE} is in business to help poor people overseas?
b. That the Salvation Army helps the homeless?
c. That the Girl Scout program fosters goodwill and socially constructive attitudes?

As seen in (21), the sentential objects identified in (22)a-c bear the reference of the NPs \textit{objectives}, \textit{broad statements of purpose}, and \textit{concepts} [see (23)].

(23) a. Why should it be difficult to agree on objectives?
b. ... these \textit{broad statements of purpose} are not \textit{objectives} at all.
c. They are only \textit{concepts}. (emphasis added)

Example (23)c may be glossed as in (24)b.

Recall that \textit{comprise} and \textit{constitute} select a subject argument denoting parts in a part-whole relationship [see (24)c]. Objects of the REALIZE reading occur as subject to \textit{constitute} or \textit{comprise} and therefore encode parts in a part-whole relationship [see (24)b].
(24) a. That CARE is in business to help poor people overseas, that the Salvation Army helps the homeless, that the Girl Scout program fosters goodwill and socially constructive attitudes constitute concepts.

b. That CARE is in business ... , that the Salvation Army helps the homeless, that the Girl Scout program fosters goodwill ... , these concepts/objectives/broad statements of purpose.

c. Three lines constitute/comprise a triangle.

Accordingly, I conclude that the REALIZE sense of understand selects as object a part in a part/whole relationship and I propose (25) as the LCS of the REALIZE reading of understand.

(25) REALIZE SENSE
understand
V
\[ [\text{that} + S]_k \]
\[-\text{DIR} \quad -\text{DIR} \]
\[ [\text{sit}\Perceive([\text{NP}]), [\text{sit}\ BE_{\text{Comp}} ([\text{sit}\ X]_k \ [\text{Place} \text{AT}[Y]])] ] \]

The LCS states that the being denoted by NP1, the subject of understand, perceives an undirected situation (a state); specifically, a composition relationship holding between parts and a whole. In particular, since the first argument of the function \( BE_{\text{Comp}} \) is indexed to the subcategorized direct object, this LCS states that understand selects as direct object the part (or parts) in a part/whole relation.

Parallel to the lexical entry for the COMPREHEND reading, since the whole is not indexed to any syntactic structure in the lexical entry, it is encoded as semantically implicit. But what comprises the implicit whole? First, the short answer. Example (24)b shows that concepts, objectives, and statements of purpose denote the whole in the relevant part/whole relationship. Of course, these NPs are precisely those that the COMPREHEND sense selects as object argument, NPs denoting a whole in a part-whole relationship.

Issues of prototypicality or canonicity suggest a longer answer. The standard way of referring to part/whole relationship is by naming the whole. For example, given an assembled jigsaw puzzle, it would odd to say The picture on all those pieces is interesting instead of The picture on that puzzle is interesting.

While the REALIZE and COMPREHEND readings of understand both select \( BE_{\text{Comp}} \) as semantic object, it is the COMPREHEND reading which invokes this relationship in a prototypical or canonical fashion. The COMPREHEND reading invokes the part/whole relationship in the usual way, by naming the whole and presupposing the parts. But the REALIZE reading invokes this relationship non-prototypically, by naming the parts, and presupposing the whole. This non-prototypical mode of invoking the part/whole relationship is part of what makes the REALIZE reading appear less clearly associated with a part/whole relationship than the COMPREHEND reading.

I suspect another factor is also at work. It appears that the implicit semantic argument is less crucial to the REALIZE reading than it is to the COMPREHEND reading. Thus, if someone asserts I understand the problem, it is critical that they actually know the identity of the components of that problem, else they don’t understand it. However, if someone asserts I understand that the mortgage is due
on the first, the real crux of the matter lies in the part as specified. I explore possible explanations for this in Section 5.

In any case, the central point remains; the COMPREHEND and REALIZE senses of understand select an object argument encoding a relationship of composition, and that these senses vary by whether it is the whole or the part which is realized in object position.

4.2 THE [READ + (MANNER)] SENSE. This sense of understand, manifest in examples such as (5), occurs in the frame [NP1 V NP2 + adjunct/complement].

The syntax and semantics of identificational predicates provide the key to the syntax and semantics of this sense of understand. Identificational predicates (e.g. is, become, turn into, etc.) occur in the frame [NP1 V NP2/Adj]. NP2 and Adj indicate what category NP1 belongs to or what properties it manifests [see (26) - (27)].

(26) Elise is a pianist.
   -DIR
   [Sit BEIdent ([Thing Token ELISE], [Place ATIdent ([Thing Type PIANIST])])]

(27) The light is red.
   -DIR
   [Sit BEIdent ([LIGHT], [ATIdent ([Property RED])])]

The NP2 argument to understand in (5) occurs as NP1 to the identificational predicate be in (28). Filling out the distributional parallel, understand’s adjunct in (5) occurs as NP2 or Adj. in (28). The semantics parallel these syntactic patterns: just as NP2/Adj to identificational predicates in (26) - (27) encode the type or property of the entity denoted by NP1, so the adjunct to understand in (5) occurring as NP2/Adj to identificational be in (28) also encodes the type or property of the entity denoted by NP1 [see (28)].

(28) a. The phrase is literal.
   b. Her suggestion is a complaint.
   c. Her suggestion is sarcastic.

Therefore, I conclude that understand in this reading selects an NP2 and adjunct encoding an identificational relation. I rename this as the IDENTIFICATION reading of understand. 4

Accordingly, omitting certain coding complexities, I propose (29) as the analysis for the IDENTIFICATION sense.

(29) IDENTIFICATION SENSE
understand
   V
   NP_k {as NP/Adj}_m {to be NP/Adj}_m
   -DIR
   -DIR
   [sit PERCEIVE ([NP]), [Sit BEIdent ([+/-Mat X]_k, Place ATIdent ([Type/Property Y]_m)])]

This lexical entry says that the IDENTIFICATION reading of understand occurs with an NP direct object, followed by either as + NP or as + Adj or by to be +
NP or to be + Adj. The LCS states that the being denoted by NP1, the subject, perceives a non-directed situation (a state); specifically, a relation of identity holding between X and Y such that Y identifies the [TYPE] or a [PROPERTY] of X.

5. EXPLAINING THE SYNTAX AND PRAGMATICS OF UNDERSTAND. Aspects of the syntax and pragmatics of understand follow directly from the semantic analysis proposed, specifically from the semantic component PERCEIVE and from BEComp.

It is in the nature of perceiving and part/whole relationships that one may perceive a whole, without also perceiving the parts comprising that whole. As such, we would expect the COMPREHEND reading to be felicitous under negation. Indeed, that is just what we find. To assert I don’t understand the problem is to assert perception of the whole (problem) while denying knowledge of identity of the parts.

Relevant to the near non-occurrence of COMPREHEND in imperative, Miller & Johnson-Laird find that PERCEIVE ‘contains neither a causal nor an intentional component’. Even if the whole is named, one can not command perception of constitutive parts. Thus, if some interlocutor does not already perceive that particular parts comprise a whole, perception of that relationship can not be compelled by simply naming the whole. Accordingly, the COMPREHEND reading is infelicitous in imperative form.

In the REALIZE reading, understand is specified to select a sentential complement. This object argument corresponds, in the LCS to the parts constituent of the function BEComp. The fact that a hearer can far more readily reason to some whole when presented the parts, than they can reason to parts when presented a whole, may well explain why the REALIZE sense does occur in imperative form. For example, the REALIZE reading [e.g. Understand that the mortgage is due on the first!] selects as direct object a sentential complement corresponding to the parts constituent of the LCS. The hearer, presented with parts, can readily induce a whole; the whole may be quite general (e.g. requirement, problem, issue, etc.), but it nonetheless completes the composition circuit specified by the LCS. Accordingly, the REALIZE sense of understand is felicitous in imperative. This pattern of reasoning may also explain why the implicit whole is less important to the REALIZE sense than the implicit parts to the COMPREHEND sense.

The pragmatics of PERCEIVE explains aspects of the pragmatics of understand. For example, since my spouse and I have long used a Krupp II Primo cappuccino maker, it would be odd for me to assert to him of this cappuccino maker, I understand how to make cappuccino. The semantics of PERCEIVE involves attending to and forming a judgment of a thing (Miller and Johnson-Laird 1977). However, judgment forming is only appropriate when a judgment is missing or faulty. So, if our existing judgment of a thing is not amiss, it is pragmatically bizarre to judge it again. This explains the pragmatic oddity of I understand how to make cappuccino in this context.

Of course, if confronted with a different machine, or a different audience, the pragmatics are back in order -- it is contextually appropriate form a judgment; accordingly, PERCEIVE is again pragmatically appropriate, and I understand how to make cappuccino is pragmatically felicitous.

6. SUMMARY & CONCLUSION. I have explained the relation between the senses of understand and have motivated why just these senses and no others
comprise the meanings of that verb. Three-ways polysemous, each sense of the verb encodes one subcategory of the semantic field of Character: the COMPREHEND reading encodes BE_{comp}^*; the REALIZE reading, BE_{comp}^*; and the IDENTIFICATION reading encodes either BE_{Ident (Type)} or BE_{Ident (Property)}.

To assert that someone understands something is to assert that they perceive the character of that thing in precisely the ways embodied in the subcategories of the semantic field, Character. This analysis explains the selectional restrictions on that verb, the relationship between its three senses, and aspects of its syntactic and pragmatic patterning.

Finally, this analysis provides evidence on how meaning is distributed across lexical items. In the case of the English verb understand, I have shown that the heft of its meaning resides in the semantics accruing to the position of its subcategorized object argument.

FOOTNOTES

1While responding to a question Thomas Shannon asked, I realized that evidence from syntactic alternations (in which the object argument participates) tells us about the semantics of that subcategorized argument position and not about the semantics of any particular object argument per se. While there will, of course, be a relationship between these, the two are nonetheless distinct.

2Although Wheeler 1993 identified four senses of understand, this finding was revised in Wheeler (ms.) where I demonstrated that the previously identified HEARSAY reading (I understand your aunt is coming) is a pragmatic variant of the REALIZE sense.

3While understand does occur in examples such as Understand the problem before you proceed! this use of the verb is ill-formed under main-verb negation and in intransitive form. That is, the affirmative imperative use does not exhibit the full syntactic distribution characteristic of the COMPREHEND reading: it is not paired with a negative imperative (Don't understand the problem before you begin!) and does not occur in intransitive form (Understand!), suggesting that if the imperative exemplifies the COMPREHEND reading of understand, it is a peripheral instance.

Further, the use of understand in Understand the problem! is similar to that of know in Know the answer!. Like understand, know is a non-volitional, stative predicate, and as such is analyzed as typically ill formed in imperative. Its imperative occurrence is non-prototypic and accordingly, does not vitiate the broader generalization that know is hostile to imperative. I suggest the same holds for understand. While these uses are readily interpretable, I do not, at this point, have an account of how hearers garner an interpretation.

4The NP2 argument to the IDENTIFICATION reading is often one typical of the COMPREHEND reading of understand as in (i). (Thomas Shannon reminded me of this point).

(i) She understands the phrase/comment as literal/sarcastic.

However, as seen in (ii), while the given NP2 does not meet the semantic constraints of the COMPREHEND sense, it is nonetheless felicitous as NP2 to the IDENTIFICATION reading [as in (iii)].
(ii)  
  a. *She understands the table.
     b. *She understands the book. (under the reading sans semantic coercion)

(iii)  
  a. She understands the book to be a gift.
     b. She understands the table to be red/the dividing line.

However the felicity of the utterance characteristically varies with the complement/adjunct structures represented [see (iv)].

(iii)  
  a. ?She understands the book as a gift.
     b. ??She understands the table as red/the dividing line.

There is clearly more to be said about the relationship between the COMPREHEND and the IDENTIFICATION readings of understand.

REFERENCES

     __. 1992. Parts and Boundaries. Lexical and Conceptual Semantics, ed. by Beth
          Levin and Steven Pinker, 9-46. Cambridge, MA: Blackwell.
     __. ms. The conceptual structure of intending and volitional action.
     10:172-188.
LEVIN, BETH. 1994. English verb classes and alternations. Chicago: University
     of Chicago Press.
     Cambridge, MA: Belknap.
     University Linguistic Club.
     17.4:409-441.
SELLS, PETER. 1985. Lectures on contemporary syntactic theories. Center for the
WHEELER, REBECCA S. 1989. The lexical entry of the English verb understand.
     Chicago: University of Chicago dissertation.
     __. 1993. Sense and subsense: The lexical semantics of the English verb
     __. ms. On the notions of part/whole and identification in the semantics of
          understand.

Rebecca Wheeler
Independent Scholar
rebwhlr@cc.usu.edu
Rule Ordering, and Constraint Interaction in OT

Young-mee Yu Cho
Stanford University

1. From the very start of generative phonology and throughout much of its history, except for a period in the 70's when there was vigorous debate on whether language-specific rule ordering is needed (Koutsoudas, Sanders and Noll 1974), it has been widely assumed that phonological rules are extrinsically ordered (Halle 1962, SPE 1968, Bromberger and Halle 1989).[1] In fact, rule ordering was one of the most powerful tools in phonological analysis in derivational analyses. Furthermore, numerous cases have been reported of dialects or historical stages of a language that contain the same underlying representations and the same rules, but differ simply by virtue of the ordering of the rules (Kiparsky 1971, Halle 1962).

The interaction between /t/-voicing and diphthong raising in Canadian English is one of the most widely cited examples supporting the rule-based nature of phonology. This example has been cited repeatedly in the literature (Joos 1942, Halle 1962, SPE 1968, Chambers 1973, Bromberger and Halle 1989). The situation is summarized in (1). The low vowel nucleus in the [ay] and [aw] diphthongs raises before voiceless consonants, and the intervocalic /t/-/d/ contrast is neutralized in favor of [d] (instead of a flap as in American English). Dialect A distinguishes the words, writer and rider, while in dialect B they are homophonous. The difference has been explained in terms of the difference in the order of the two rules. It is argued that reversing the order of rule application of Dialect A results in Dialect B. To use the ordering terminology, in Dialect B the rules apply in bleeding order; in dialect A they apply in counterbleeding order.

(1) Dialect A (counterbleeding)

<table>
<thead>
<tr>
<th>/rayt/</th>
<th>/rayd/</th>
<th>/rayt-ær/</th>
<th>/rayd-ær/</th>
<th>UR</th>
</tr>
</thead>
<tbody>
<tr>
<td>rayt</td>
<td>______</td>
<td>raytær</td>
<td>______</td>
<td>Raising</td>
</tr>
<tr>
<td>______</td>
<td>______</td>
<td>raydær</td>
<td>______</td>
<td>Voicing</td>
</tr>
</tbody>
</table>
Dialect B (bleeding)

/rayt/     /rayd/     /rayt-ør/ /rayd-ør/     UR
rʌyt         ______      raydør     ______     Voicing
                      ______      ______     Raising

An exactly parallel case has been reported for American English, as shown in (2). In this case, the two rules involved are Vowel Shortening before a voiceless consonant and Flapping of /t/ and /d/.

(2) American English Flapping and Vowel Shortening

Dialect 1: Shortening counterbleeds Flapping

/pæt+ɪn/     /pæd+ɪn/     Shortening
pætin         [pærin]     Flapping
[ pærɪn ]

Dialect 2: Flapping bleeds Shortening

/pæt +ɪn/     /pæd+ɪn/     Flapping
pærin         pærin     Shortening

In Optimality Theory (McCarthy and Prince 1993, Prince and Smolensky 1993), the grammar evaluates candidate outputs in parallel against a hierarchy of ranked, violable constraints. Devoid of rules and rule ordering, a powerful device of operational theories, Optimality Theory has to somehow reflect dialectal variation by relying solely on constraint interaction. Before we provide such an account for English, we need to spell out two additional assumptions about the substantive aspect of well-formedness constraints, in addition to the usual ones.

The first concerns the markedness considerations. According to Kiparsky (1994), constraints cannot specify unmarked feature values, and for every constraint that refers to a phonological category, there is an otherwise identical constraint that refers specifically to the marked member of that category (e.g. Fill-lab, Fill-place, Spread-lab,
Spread-place, etc.) In addition, a specific constraint is active only if it precedes the corresponding general constraint by Panini’s Theorem (Prince and Smolensky 1993). Since no constraints refer to an unmarked feature, it is impossible to have a constraint on an unmarked feature ranked above the constraint on the corresponding marked feature.

Second, we assume, following McCarthy (1994), that there is a distinction between two relations obtaining between a feature F and a segment S (or between a segment S and a prosodic constituent P). Sponsorship obtains when a feature is associated with the segment in the input, while parsing refers to the association in the output.

(3) Two relations between a feature F and a segment S
Sponsorship: S sponsors F iff F is associated with S in input.
Parsing: S parses F iff F is associated with S in output.

Now, we attempt an optimality analysis of the American English facts in (4). Flapping ensures a foot-internal ambisyllabic alveolar stop to be realized as a flap.

(4) An OT account of Vowel Shortening and Flapping

\[ \begin{align*}
\text{Constraint 1:} & & \text{Vowel Shortening:} & & \tilde{\nu} C[-\text{voice}] \\
\text{Constraint 2:} & & \text{Flapping:} & & \\
\end{align*} \]

Dialect 1 (patting [pærɪn], padding, [pærɪn])

<table>
<thead>
<tr>
<th>patting</th>
<th>VS 1</th>
<th>Flapping</th>
</tr>
</thead>
<tbody>
<tr>
<td>pætɪn</td>
<td>*!</td>
<td>*</td>
</tr>
<tr>
<td>pærɪn</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>pætɪn</td>
<td></td>
<td>*!</td>
</tr>
<tr>
<td>pætɪn</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- pærɪn
\[
\begin{array}{|c|c|c|}
\hline
\text{padding} & \text{VS 1} & \text{Flapping} \\
\hline
\text{pædin̄} & * & ! \\
\text{├→ pærin̄} & & \\
\text{pædin̄} & * & ! \\
\text{├→ pædin̄} & & ! \\
\hline
\end{array}
\]

Dialect 2 (patting [pærin̄], padding [pærin̄])

\[
\begin{array}{|c|c|c|}
\hline
\text{patting} & \text{VS 2} & \text{Flapping} \\
\hline
\text{pætin̄} & * & * \\
\text{├→ pærin̄} & & \\
\text{pætin̄} & * & * \\
\text{├→ pærin̄} & & * \\
\hline
\end{array}
\]

(5) two versions of Vowel Shortening

a) f-Sponsor (vs1)

\[\begin{array}{c}
\text{VC} \\
\text{├→ [voice]} \text{ in input}
\end{array}\]

b) f-Parser (vs2)

\[\begin{array}{c}
\text{VC} \\
\text{├→ [voice]} \text{ in output}
\end{array}\]

In Dialect 1, a short vowel results when followed by a consonant that sponsors the feature [-voice] (i.e., an underlyingly voiceless consonant), shown in (5a). It makes no difference whether the underlying [-voice] is parsed in the output or not. The flapping process wipes out the voicing specification on the surface. There is some question as to whether Dialect B in Canadian English and Dialect 2 in American English actually exist (Kaye 1990), but if they do, the dialectal difference can be obtained by activating a slightly different constraint, Vowel Shortening 2 in (5b); i.e. a short vowel occurs only with a [-voice]-parsed consonant. When the sponsoring constraint (VS1) is active, the optimal candidate reflects the underlying distinction (in this case, the voiced /d/ vs. voiceless /t/). On the other hand, when the parsing constraint (VS2) is active, the optimal candidate is true to the surface forms; i.e., the constraint relies on the phonetic voicing of the flap.
The upshot of this short exercise is that there is no direct translation of the powerful argument for extrinsic rule ordering within a phonological theory based on defeasible ranked constraints. The two constraints, VS and Flapping do not interact as rules do. Rather, the variation depends on the nature of the constraints (in this case, the sponsorship vs. the parsing).

2. For the past two decades or so, determination of the ordering between rules which potentially interact was an important aspect of phonological analyses. The ordering relationship between rules is stated in terms of the potential effect (both positive and negative) that the application of one rule has on the application of another.

(6) Rule Ordering Terminology (Kiparsky 1968)

1. Feeding: Rule A feeds Rule B if A applies before B and creates places in which B can apply.
2. Bleeding: Rule A bleeds Rule B if A applies before B and prevents B from applying.
3. Counterfeeding (= failure to feed): If Rule A would have fed Rule B but is ordered too late to do so, it counterfeeds B.
4. Counterbleeding (= failure to bleed): If Rule A would have bled Rule B but is ordered too late to do so, it counterbleeds B.

It is clear that there is no direct correlation between rule ordering in the old theory and constraint interaction such as constraint ranking in OT. In the old theory, differences among languages and dialects resulted from languages having different rules, and in some cases different ordering relations among the same rules. Now Optimality Theory articulates a more restricted position where all of the constraints are universally present, and the differences between languages and dialects arise solely from the difference in the rankings of the constraints.

In derivational theories, the sequential order of rules was supported by the observation that a rule may require information which is supplied by another rule. For instance, in the typical feeding relation, information supplied by the
first rule is essential to the application of the second rule. One way of interpreting this in OT is to assume that the information relevant in activating constraints is already present underlyingly and that the two constraints independently exert their power, thus yielding an output that satisfies both of the constraints. This conceptual difference between these two types of theories yields quite different predictions. For instance, well-known cases of neutralization and assimilation have been accounted for in terms of a feeding relationship between the processes, as exemplified by the Korean example in (7).

(7) Korean Neutralization and Assimilation

Feeding Order (Neutralization precedes Assimilation):
/kas+pota/ → katpota → [kappota] ‘rather than green mustard’
/kot+pota/ → [koppota] ‘rather than soon’

Hyphothetical Counter-feeding Order (Assimilation first):
/kas+pota/ → [katpota] (Only Neutralization applies.)
/kot+pota/ → [koppota] (Assimilation of underlying dentals only)

The fricative [s] neutralizes to the dental stop [t] in coda position, and undergoes assimilation when a marked consonant such as a labial or a velar follows. Derivational theories generate two equally plausible grammars depending on the order of the rules, but the fact is that feeding order is the only order that surfaces in language after language. For example, in the Sanskrit examples in (8), neutralization of laryngeal features appears to apply before assimilation such as palatalization and lateralization.

(8) Sanskrit Neutralization and Assimilation

/tad ca/ → tat ca → [tacca] ‘and this’
/tad lebhe/ → tat lebhe → [tallebhe] ‘I obtain this’

In contrast, an optimality account provides only the single possibility in which both Neutralization and Assimilation apply, namely the result of a feeding order. I
assume that assimilation is enforced by a constraint which prohibits multiple specification of a feature.

(9) Korean Neutralization and Assimilation

\[
\begin{array}{cc}
\text{Neutralization} & \text{Assimilation} \\
* C_o & \text{C} \\
\mid & \text{C} \\
[+\text{cont}] & \text{Place}
\end{array}
\]

<table>
<thead>
<tr>
<th>/kas+pota/</th>
<th>Neutralization</th>
<th>Assimilation</th>
</tr>
</thead>
<tbody>
<tr>
<td>kaspota</td>
<td>* !</td>
<td>* !</td>
</tr>
<tr>
<td>ktpota</td>
<td></td>
<td></td>
</tr>
<tr>
<td>→ kappota</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In this account, the processes of Neutralization and Assimilation are handled by two independent constraints, and there is no obvious way to obtain the 'counterfeeding effect'. One might wonder if it is possible to have the Assimilation constraint sensitive to the underlying distinction between the dental stop /t/ and the fricative /s/, limiting the scope of Assimilation only to a segment that sponsors the features of the dental stop. However, this is in violation of Kiparsky's markedness frame in which constraints cannot specify unmarked feature values alone. [2] There is no way to refer to dental stops to the exclusion of the coronal fricative or other marked consonants.

3. McCarthy (1994) presents two processes in the Najdi dialect of Bedouin Arabic. In non-final unstressed open syllables, short /i/ deletes (Syncope) and short /a/ is changed to [i] (Raising) but the [i]'s from /a/ do not delete.

In a derivational account, the most obvious thing to do would be to order Syncope before Vowel Raising so that the output of VR does not have a chance to undergo the first rule, a typical counterfeeding order.
(10) Arabic chain shift (McCarthy 1994)

<table>
<thead>
<tr>
<th>Counterfeeding order</th>
<th>Feeding order</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i \rightarrow \emptyset$</td>
<td>$a \rightarrow i$</td>
</tr>
<tr>
<td>$a \rightarrow i$</td>
<td>$i \rightarrow \emptyset$</td>
</tr>
<tr>
<td>(* $a \rightarrow i \rightarrow \emptyset$)</td>
<td>$a \rightarrow i \rightarrow \emptyset$</td>
</tr>
</tbody>
</table>

Kiparsky (1994) provides a slightly revised version of McCarthy's analysis. The crucial modification involves markedness considerations. In both account, some constraints shown in (11) are crucially sponsoring constraints. We have seen earlier that when the output is sensitive to underlying distinctions, sponsoring constraints are required.

(11)

Parse-V (A [low]-sponsoring vowel must be parsed.)

\[
\text{low}
\]

Parse-V (A [place]-sponsoring vowel must be parsed.)

\[
\text{place}
\]

(12) Kiparsky's improvement on Arabic chain shift

\[
*V] \sigma, \quad \text{Parse V} \quad >> \quad \text{Parse-place} \quad >> \quad \text{Parse V}
\]

<table>
<thead>
<tr>
<th>place</th>
<th>low</th>
<th>place</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>$V$</th>
<th>low</th>
</tr>
</thead>
</table>

$*$

<table>
<thead>
<tr>
<th>$a \rightarrow i$</th>
<th>$&lt;\text{low}&gt;$</th>
<th>$*$</th>
</tr>
</thead>
</table>

---

#### Table

<table>
<thead>
<tr>
<th>*V]*σ</th>
<th>Parse V</th>
<th>Parse-place</th>
<th>Parse V</th>
</tr>
</thead>
<tbody>
<tr>
<td>place</td>
<td>low</td>
<td>place</td>
<td></td>
</tr>
</tbody>
</table>
As illustrated by the first three candidates in (12), Parse-V-low is ranked higher than its corresponding general constraint Parse-V-place, and the vowel segment of the marked vowel [a] needs to be parsed, though its terminal feature [low] does not surface in the output due to the Neutralization constraint, which prohibits any place features in the weak position. The unmarked vowel, /i/, however, does not surface due to the lower ranked Parse-V-place because there is no Parse-V-high constraint.

In Arabic the unmarked vowel /i/ deletes, and the marked vowel /a/ changes to the unmarked vowel /i/, the state of affairs generated by the counterfeeding order between Syncope and Raising. The feeding order would predict a situation where both marked and unmarked vowels delete, as found in Yokuts, Icelandic and Piro. An optimality account can derive such a situation just as easily, by reranking of the same constraints, in particular by reversing the ranking between Parse-V-place and Parse-pl (Kiparsky 1994). Parse-V-low is dominated and thus has no effect.

We have observed in Korean that the only order possible between neutralization and assimilation is the feeding relation while in the vowel syncope and raising case of Arabic, both orders are typologically possible. In derivational theories, there seems to be no principled way to account for such contrasts. In OT, however, the explanation follows naturally from the very nature of the constraints
themselves. Reranking the general constraint (Parse-pl) higher than the specific constraint referring to the marked feature [low] results in the difference between Arabic and Icelandic, for instance.

4. Now, we will look at another intriguing case in which the derivational theory relying on rule ordering clearly makes predictions distinct from Optimality Theory. We have seen in our discussion of Korean neutralization and assimilation that a derivational account is capable of generating two grammars with two different orderings, while an optimality account in no way allows two possibilities. An analogous case is found in Klamath where two processes of deglottalization have been analyzed as exhibiting a bleeding relationship in the literature. Just as natural an account can be put forward where the two rules are in a feeding order, but an optimality account excludes such a possibility on principled grounds.

Klamath, an Amerindian language spoken in southwestern Oregon has both glottalized obstruents and glottalized sonorants, in addition to voiceless (aspirated) sonorants. Deglottalization of stops occurs in preconsonantal position, except before the voiced nonglottalized sonorants (m, n, w, y, l), as exemplified by (13). [3] Sonorants, however, deglottalize before any obstruents, as well as glottalized and aspirated sonorants (Barker 1964, Kisseberth 1972, Kean 1973, Lightner 1976). [4]

(13) Klamath Deglottalization

- **a.** p'et'-a ‘a hole becomes larger’
  p’e-pt’-a ‘(distributive) holes tear out’
  wLet’-wal ‘lies spread eagled on top of’
  wLet-pg-a ‘is lying flat on the back’
- **b.** nčoq’-a ‘is deaf’
  nčoq-n’apg-a ‘is almost deaf’
- **c.** nt’op’-a ‘rots’
  nt’op-Wi:y-a ‘almost rotted’
- **d.** nt’op’-ye:g-a ‘starts to rot’
- **e.** n’o-k’a ‘little head’
  n’o-n-k’a ‘(distributive) are breathless’
  -iwy’aq ‘put in plural objects’
(14) Klamath Deglottalization (bleeding order)

/n̓q̓oq̓'-l'-g-a/  ‘ears are stopped up’
[n̓q̓oq̓'-l-g-a]

/q̓l'/  /l' g/  /-q̓'-l'-g-/-  UR
   ___   lg   q' l g   Sonorant Deglottalization
   q̓l'        ___  ___  Obstruent Deglottalization
   [q̓l']   [lg]   [q̓l'lg]

(15) Klamath Deglottalization (feeding order)

/q̓l'/  /l' g/  /-q̓'-l'-g-/-  UR
   q̓l'   ___   q l' g   Obstruent Deglottalization
   ___   lg   q'l'g    Sonorant Deglottalization
   [q̓l']   [lg]   [q'l'g]

In an OT account, the different behavior of glottalized obstruents and glottalized sonorants is reflected in the
family of Deglottalization constraints. Note that the most general deglottalization (*C’C) plays an important role.

(16) Relevant Constraints in Klamath

*Obs’ Son’, *Obs’ Obs, *Son’ Cons >> Parse >> *Cons’Cons (Cons’ denotes a laryngeal (glottalized) consonant.)

a. Glottalized obstruents followed by an obstruent

<table>
<thead>
<tr>
<th>UR</th>
<th>*O’S’</th>
<th>*O’O</th>
<th>*S’C</th>
<th>Parse</th>
<th>*C’C</th>
</tr>
</thead>
<tbody>
<tr>
<td>/q’g/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>q’g</td>
<td></td>
<td>* !</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>→ qg</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| /q’g’/ |       |      |      |       |      |
| q’g’  |       | * !  |      |       |      |
| q’g   |       | * !  |      |       | *    |
| → qg’  |       |      |      |       |      |
| qg    |       |      |      |       | * * !|

b. Glottalized obstruents followed by a sonorant

<table>
<thead>
<tr>
<th>UR</th>
<th>*O’S’</th>
<th>*O’O</th>
<th>*S’C</th>
<th>Parse</th>
<th>*C’C</th>
</tr>
</thead>
<tbody>
<tr>
<td>/q’l/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>q’l</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>→ ql</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| /q’l’/ |       |      |      |       |      |
| q’l’  |       | * !  |      |       |      |
| q’l   |       | *    |      |       | * !  |
| → ql’  |       |      |      |       |      |
| ql    |       |      |      |       | * * !|
c. Glottalized Sonorants followed by an obstruent

<table>
<thead>
<tr>
<th>UR</th>
<th>*OS'</th>
<th>*OO</th>
<th>*SC'</th>
<th>Parse</th>
<th>*CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>/l'g/</td>
<td>l'g</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→</td>
<td>lg</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/l'g'</td>
<td>l'g'</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>l'g</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→</td>
<td>lg'</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>lg</td>
<td></td>
<td></td>
<td></td>
<td>**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>UR</th>
<th>*OS'</th>
<th>*OO</th>
<th>*SC'</th>
<th>Parse</th>
<th>*CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>/l'p/</td>
<td>l'p</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→</td>
<td>l'p</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/l'p'</td>
<td>l'p'</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>l'p</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→</td>
<td>l'p'</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>l'p</td>
<td></td>
<td></td>
<td></td>
<td>**</td>
</tr>
</tbody>
</table>

There is the general constraint, *Cons’ Cons, which dictates that preconsonantal glottalized consonants are ill-formed. There is a family of more specific deglotttalization constraints, *Obs’ Son’, *Obs’ Obs, and *Son’ Cons. Parse has to be ranked higher than *Cons’Cons because glottalized obstruents are not deglottalized before a plain sonorant, but the family of specific constraints needs dominate Parse to have any deglottalization effect. As can be seen in the tableaux (16), Obs deglottalization and Son deglottalization are not crucially ranked. Finally, as illustrated in (17), the choice of [q’lg] over [qlg] is determined by the crucial ranking between the Glottal Constraints and Parse. The second candidate is chosen due to the faithfulness condition. The bleeding order between ObsD and SonD is not reflected in any way in the OT account.
(17) Obs’-Son’-Cons

<table>
<thead>
<tr>
<th></th>
<th>*OS’</th>
<th>*OO</th>
<th>*SC</th>
<th>Parse</th>
<th>*CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>/q’l’g/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>q’l’g</td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>q’l g</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>q l’g</td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>qlg</td>
<td></td>
<td></td>
<td></td>
<td>**!</td>
<td></td>
</tr>
</tbody>
</table>

It appears at first glance that the same effect of the feeding relation between two rules in (15) could be obtained by reranking *C’C and Parse, as in (18).

(18) *C’C >> Parse yields [qlg]

<table>
<thead>
<tr>
<th></th>
<th>*OS’</th>
<th>*OO</th>
<th>*SC</th>
<th>*CC</th>
<th>Parse</th>
</tr>
</thead>
<tbody>
<tr>
<td>/q’l’g/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>q’l’g</td>
<td>*!</td>
<td>*</td>
<td></td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>q’l g</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>q l’g</td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>qlg</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

However, this ranking is not consistent with the other facts. When glottal obstruents are followed by a sonorant (both glottalized and plain), the ranking of *CC >> Parse yields non-optimal forms as in (19).

(19) *C’C >> Parse cannot be true in Klamath

<table>
<thead>
<tr>
<th></th>
<th>*OS’</th>
<th>*OO</th>
<th>*SC</th>
<th>*C’C</th>
<th>Parse</th>
</tr>
</thead>
<tbody>
<tr>
<td>/q’l/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>q’l</td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ql</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As shown in (19), /q’l/ surfaces as [q’l] (no changes) but the ranking of *Cons’ Cons >> Parse produces [ql]. This ranking has the same effect as subordinating the three obstruent/sonorant specific constraints by *C’C. However, the fact is that the obstruent target is different from the sonorant target. In other words, the specific constraints are
active and must precede the general *C’C constraint, mediated by Parse.

A simple ordering statement between the two rules would easily generate two grammars while an optimality account excludes on principled grounds the effect of the ‘feeding’ relation between Obs deglottalization and Son deglottalization.

6. One of the most powerful arguments for the sequential application of rules has been the fact that differences in closely related dialects are most efficiently explained in terms of rule ordering. However, meta relations such as rule ordering statements are not part of OT, and differences in ranking of constraints is the only source of variation across languages within this theory. It has argued that OT has not only the same descriptive coverage in dealing with dialectal variation as derivational theories, but it also offers more restricted accounts for Korean and Klamath.

* I am very grateful to Sharon Inkelas, Paul Kiparsky and Will Leben for their helpful comments.

NOTES

[1] According to Koutsoudas, Sanders and Noll (1974), all ordering is predictable, given appropriate universal principles. Their position of nonordering has been believed to be falsified by cases where closely related dialects differ only in ordering.
[2] One cannot simply appeal to the naturalness of ‘feeding rules’ since there are many cases of bleeding rules, and language change and dialectal variation have been assumed to involve changes in rule ordering.
[3] According to Kisseberth(1972), Stop Deglottalization and Sonorant Deglottalization were not differentiated, but Kean (1973) and Lightner (1976) argue that stops and sonorants deglottalize in distinct environments and the two rules cannot be collapsed into one.
[4] The /s/ does not seem to trigger deglottalization in some cases (e.g. [mol’s] ‘pus’).
CONJUNCTION AS A CASE FEATURE-CHECKER
Ed Zoerner
University of California, Irvine

1.0 The Puzzles
The facts of NP-coordinations pose at least four puzzles needing explanation. This paper identifies these puzzles, and demonstrates that all of them have a related solution.

First, an asymmetry exists between NP-coordination and non-NP-coordination regarding the presence of an overt conjunction. This asymmetry manifests itself in two ways. For one thing, a number of languages require an overt coordinator in an NP-coordination although they allow or even require parataxis in coordinations of other phrases. Chinese gives an example of this:

1. a. [Robin he Kim] mai-le yi-ben shu
and buy-ASP one-CL book
'Robin and Kim bought a book'
b. *Robin, Kim mai-le yi-ben shu

This phenomenon appears frequently, and in different language families; some languages which pattern with Chinese in this regard include:

2. Barasano: kede for NPs, parataxis elsewhere
Chemehuevi: wai for NPs, parataxis elsewhere
Tera: nde for NPs, parataxis elsewhere

No language, however, requires an overt conjunction in non-NP-coordination yet allows paratactic NP-coordination. An interesting implicational universal holds: if a language has an overt coordinator for non-NP-conjunction, it will have one for NP-conjunction. Also, even in some languages that always require syndetic coordinations, a distinction exists; one lexical item conjoins NPs while another conjoins all other phrases. The language of Nguna offers one such example; go conjoins NPs and poo conjoins other phrases such as VP:

3. Nguna (Shütz (1969: 49))
a. e pei na-vinaga go/*poo suu-goro
it be food and clothing
'It was food and clothing'
b. a go vano poo/*go tape na-peta seara
   I INT go and get yams some
   'I will go and get some yams'

Other languages with conjunctions patterning in such a fashion include:

4. Japanese: to, mo, ya for NPs; -te for APs and VPs
   Somali: iyo for NPs, o for all other phrases
   West Futuna-Aniwa: ma for NPs, u for all others

   However, no language requires the use of a special coordinator to conjoin
   any phrase other than NP without requiring a distinct NP-conjunction as well; no
   language, for example, has one conjunction to conjoin APs, PPs and NPs but
   another for VP. NP stands apart from all other phrases; call this the asymmetry
   puzzle.

   Second, as Emonds (1986) among others points out, English permits
   nonstandard Case to surface in coordinations but not elsewhere:

   5. a. *Me left
      b. Robin and me left
      c. Me and Robin left

   Native English speakers produce forms along the lines of (5b) and (5c)
   freely despite prescriptive injunctions to the contrary; this too requires explana-
   tion. Most languages do not permit such flexibility of Case-realization:

   6. Dutch
      a. Ik heb een klok
         1S-NOM have a clock
         'I have a clock'
      b. *Me heb een klok
         1S-ACC have a clock
      c. *Robin en me hebben een klok
      d. *Me en Robin hebben een klok

   Why English permits nonstandard Case while most languages do not needs
   an account; call this the nonstandard puzzle.

   The third puzzle consists of the fact that all nonfinal conjunctions must bear
   identical Case. In English, this holds of both standard and nonstandard Case; the
   various Cases may not 'mix and match'.
7. a. He, she, they and Robin (all) left
b. Him, her, them and Robin left
c. *He, she, them and Robin left
d. *He, her, they and Robin left
e. *Him, she, they and Robin left
f. *He, her, them and Robin left
g. *Him, her, they and Robin left
h. *Him, she, them and Robin left

Call this the identity puzzle; to date no theoretical account explains why multiple occurrences of nonstandard Case as in (7b) should improve on single occurrences of same as in (7c).

The fourth puzzle stems from the third. In English, the final conjunct stands exempt from the aforementioned requirement on Case identity. The final conjunct may in fact surface with differing Case, as in:

8. a. He, she, they and me all left
b. ?Him, her, them and I all left
c. Robin saw he, she, they and me yesterday
d. ?Robin saw him, her, them and I yesterday

(8a) sounds perfectly natural, although the final conjunct bears Accusative Case (ACC) rather than Nominative Case (NOM) as do the other conjuncts. (6b), which contrasts a final NOM with non-final ACC, sounds less natural (probably due to the status of ACC in English as 'default' Case) but still improves on the bad forms in (7). (8c, d) show that such a Case contrast may occur in object positions as well. Call this the final-conjunct puzzle; the theory needs to account for the peculiar Case treatment English gives its final conjuncts.

This paper proposes that a single explanation answers all four of these puzzles. Specifically, given a representation of coordination as a set of syntactic shells in which a conjunction assigns Case to its complement, and checks Case features of nonfinal conjuncts at LF, all the data above fall out. To reach this argument, this work first motivates a new structural representation of polyterm coordination, and then shows how it follows that under this representation a conjunction must assign Case. It continues by developing the notion of conjunction as a Case feature-checking element, and then demonstrates that these ideas resolve each of the puzzles in turn.

2.0 Representing Coordinations

This work accepts the view as in Munn (1992) that a coordinating conjunction (call it &°; English 'and' and 'or' qualify) heads its own functional phrase (&P). It digresses from such previous work, however, in the way in shows polyterm coordinations; it forwards the claim that a single base-
generated &° can project any number of &P-shells to accommodate any number of conjuncts; a coordination with \( n \) terms will consist of \( n-1 \) &P nodes. For example, a three termed English coordination of NPs appears as:

![Diagram of coordination structure](image)

This structure draws its inspiration from the VP-shell analysis of Larson (1988, 1990). All &° positions save the lowest remain underlingly empty at PF. The base-generated (generally lexical) &° undergoes a Form-Chain operation at LF, and in so doing coordinates all conjuncts within its checking domain; note that each &° position stands in a head-spec relation with a nonfinal conjunct.

Adopting this structure confers several advantages. First, the structure of (9) directly produces the correct PF word order of conjuncts and conjunction; no small matter. A standard view of conjunction holds that a base-generated &° appears between each conjunct; this requires positing an unmotivated reduction rule, which affects the underlined terms below:

![Diagram of coordination structure](image)

On the other hand, the &P-shell structure, with its single base-generated lexical &°, arrives at the correct result without appeal to such reduction. Regardless of the number of conjuncts, the &° will always precede the final conjunct.¹

Furthermore, the &P-shells establish a structural hierarchy amongst conjuncts, which correctly depicts binding asymmetries, as noticed for example by Munn (1992: 20):

11. a. John’s dog and he/him, went for a walk  
   b. *He, and John’s dog went for a walk.

These binding facts hold regardless of the number of conjuncts; this falls
out directly given the hierarchy of the &P-shell analysis. Such a hierarchy also explains the fact noted in McCloskey (1986) that Irish allows pro to enter a coordination only when it stands as the first conjunct:

12. a. Bhíos [&P pro-féin agus Eoghan] i láthair
    be-PT EMPH and Owen present
    ‘Owen and I were present’

     b. *Bhíos [&P Eoghan agus pro-féin] i láthair

In (12a), the verb governs the first conjunct in [Spec, &P] and hence licenses the pro with its agreement features. Because a governing head &º intervenes between the verb and pro in (12b), such licensing cannot take place.

More central to solving the puzzles, however, the &P-shell analysis crucially represents a structural distinction between the final conjunct and all nonfinal conjuncts. The former occupies [Comp, &º], while the latter all occupy a [Spec, &P] position within the &P-shells. This immediately leads to the prediction that a final conjunct may behave differently from a nonfinal one by virtue of its unique structural position. Before investigating this claim regarding Case, note that [Comp, &º] does in fact hold some exclusive properties. For example, an echo wh-phrase may only appear as a final conjunct:

13. a. Robin sold [&P pens, pencils and what]?
     b. *Robin sold [&P what, pens and pencils]?
     c. *Robin sold [&P pens, what and pencils]?

Also, notice that an 'etcetera'-type phrase whose categorial status differs from other conjuncts may only surface as a final conjunct:

14. a. Robin [&P[\text{v- runs}, [\text{v- jumps}] and [\text{NP stuff like that}]]]
     b. *Robin runs, stuff like that, and jumps
     c. *Robin stuff like that, runs and jumps

The ensuing attempt to solve the Case puzzles will take advantage of this empirically justified asymmetry between [Comp, &º] and [Spec, &P] that the &P-shell analysis provides.

3.0 **Coordinations and Case**

Once one grants that &º heads its own category, it follows that &º can assign Case. Consider for example the following partial diagram of a sentence with a coordinated NP in a direct object:
15.

\[ V' \]

\[ V^\circ_{+\text{CASE}} \rightarrow \parallel \&P \]

saw

\[ \text{NP} \]

\[ \&' \]

Robin

\[ \&^\circ \]

NP

and

Kim

Assume that all Case-marking takes place within VP, and that subsequent raising to AgrP satisfies checking requirements only. In (15), the verb cannot directly assign Case to an NP because of the intervening functional &P node. Nor can the verb merely assign Case to the &P node itself. For one thing, a functional &P simply does not equal an NP for purposes of bearing Case; and in any event every NP in the coordination would still lack Case-marking, in direct violation of the Case Filter.

Nor does the &^\circ itself bear an independent lexical specification to assign Case. The following demonstrates this:

16. *I tried [\&P Robin and Kim] to leave early

Here, the subject of the embedded clause stands in a Caseless position; the presence of the &^\circ does not salvage the construction. The &^\circ therefore does not assign Case on its own.

Rather, to satisfy the Case Filter, the V^\circ must percolate the ability to assign Case through the &P and &' nodes to the lexical &^\circ. The following diagram depicts this:

17.

\[ V' \]

\[ V^\circ_{+\text{CASE}} \rightarrow \&P_{(+\text{CASE})} \]

saw

\[ \text{NP} \]

\[ \&'_{(+\text{CASE})} \]

Robin

\[ \&^\circ_{+\text{CASE}} \]

NP

and

Kim

Note in passing another advantage of the &P-shell analysis over a flat representation of coordination as in (10); the former but not the latter can show such percolation under a standard head-to-head relation. Crucially, in the English form in (17) it does not necessarily transmit any particular Case to the &^\circ, but empowers it to assign Case. More on this in Section 5.2. The &^\circ so appointed by a verb stands in an appropriate structural relation for NP-Case-marking to
satisfy Case Filter requirements.

Here the importance of the distinction between [Comp, &°] and [Spec, &P] becomes apparent; this work claims that a Case-empowered &° directly assigns Case only to its complement, but not to any NP in a [Spec, &P] position. Just as &° licenses a wh-phrase or 'etcetera' phrase only under a head-complement relation, &°-Case-assignment takes place under this structural relationship. Nonfinal conjuncts, which stand in [SPEC, &P], receive no direct Case-marking from the &°. Nonfinal conjuncts may surface with any Case on their own, without direct Case-marking; the following section discusses the LF conditions that delimit the grammaticality of such Case constructions.

4.0 &° and Feature-checking

Recall the earlier claim that an &° term raises at LF to the highest &° position to conjoin all nonfinal conjuncts through a head-spec relation; such movement of course leaves behind traces. The following depiction of a four-termed NP coordination at LF in English illustrates this:

18. \[&P \text{ Robin} [\& \text{ and}, [\&P \text{ Kim} [\& \text{ t}, [\&P \text{ Terry} [\& \text{ t}, \text{ Pat}]]]]]]

In the terminology of Chomsky (1993: 12), the [Spec, &P] positions together form the checking domain of the lexical &° head (while the lowest [Comp, &°] constitutes the complement domain). As noted, the Form-chain movement of an &° as shown above brings each member within the checking domain into a head-spec relation with the &° itself.

Chomsky (1993) discusses instances in which a verb raises to the head AgrP for purposes of checking off agreement features of elements within the checking domain formed by such movement. The spirit of this idea applies straightforwardly to coordinations as represented here; conjuncts, which lie in the checking domain of &°, enter into a coordination licitly by virtue of an &° head checking their features successfully at LF. Because each member of an &°-chain in fact manifests the same single base-generated head, it follows naturally that each &° position will check off identical features. This means that in grammatical forms, all elements in an &°'s checking domain will share the same features.

Within the checking domain of an &°, any conjunct that bears an inappropriate feature causes the form to crash. As a simple example, consider the feature [±N] in an NP-coordination:

19. *[Robin, in, Kim and Terry left]

The ungrammaticality of the above now has a primarily syntactic rather than purely semantic explanation; the P° 'in' lacks the [+N] feature that the &° must check off at LF.
This notion of feature-checking applies directly to Case-realization as well. Consider a Case-marked NP as having a [+Case] feature, and one lacking Case as [-Case] (or, alternatively, as lacking such a feature altogether). An &° that undergoes Form-chain will therefore check this [Case] feature of all nonfinal conjuncts just as it checks features such as [±N].

5.0 Solving the Puzzles
To sum up the key points thus far: an &° head directly assigns Case to a (final) conjunct NP within its complement domain, and at LF checks off features of all nonfinal conjuncts, which lie within its checking domain. This established, all four puzzles have related explanations.

5.1 The Asymmetry Puzzle
The asymmetry puzzle calls attention to the fact that NPs have a special status regarding coordination; NP-conjunction generally requires syndetic rather than paratactic coordination, and very often a language will have a special lexical item expressly for such a purpose. This now has a natural explanation given the fact that an &° (rather than just a V°) bears the responsibility of assigning Case and checking Case. Under the natural assumption that overt lexical items inherently serve as Case-assigners better than phonetically null terms do, it follows that many languages will eschew parataxis in NP-coordinations.³ It also follows under similar reasoning that a language may have one particular &° for Case-assignment but another for all other coordinations, which do not involve Case.

5.2 The Nonstandard Puzzle
In a sense, the nonstandard puzzle has two parts, since nonstandard Case can surface in either a final or nonfinal conjunct; recall:

20. a. [Robin and me] left
    b. [Me and Robin] left

A verb percolates Case-assigning ability to an &° standing in an appropriate head-to-head relation. Whether or not the verb in so doing also transmits a particular Case becomes an important point. Should the verb not dictate a particular Case to the &°, the possibility for nonstandard Case realizations arises. The following necessary parameter speaks to part of the nonstandard puzzle:

21. A language’s &° does/does not filter out the Case of a V°.

This parameter conforms to Minimalist principles in that it places the onus of describing differences amongst languages on the functional head &°. Dutch
and English choose opposite selections of the parameter; the following diagram may help clarify:

22.  

<table>
<thead>
<tr>
<th>a. Dutch</th>
<th>b. English</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Diagram" /></td>
<td><img src="image" alt="Diagram" /></td>
</tr>
</tbody>
</table>

Most languages follow Dutch in selecting the 'does' version of (21). This means that the $\&^\circ$ will receive whatever standard Case the verb normally assigns; $\&^\circ$s in subjects will receive and therefore assign NOM, $\&^\circ$s in objects will receive ACC from the verb, and so on.

English, on the other hand, selects the (apparently marked) 'does not' selection; its $\&^\circ$ terms merely inherit the ability to assign Case, but filters out any specific Case to assign. This leads to the possibility that an English $\&^\circ$ will enjoy a wider range of Case-assigning possibilities than a verb does; in (20a), the $\&^\circ$ 'chooses' to assign nonstandard Case to its complement. The particular selection of the parameter in (21) may depend in part upon the complexity of Case-realization that a language shows generally; languages such as English which lack a complex set of overt Case-markers (English showing Case on pronouns and 'whom' only) will stand a greater chance of selecting the 'does not' option.

Part of the nonstandard puzzle remains, however, since an English conjunct in [Spec, $\&P$] may bear nonstandard Case, as in (20b); this Case may even differ from the Case of the [Comp, $\&^\circ$] term:

23.  

| a. Both [him and I] left early |
| b. Robin gave both [he and us] a nickel |

However, this also results from the parameter mentioned above. Since English $\&^\circ$s do not receive any particular Case specification from a verb, it follows that they will not face the same restriction on their Case-checking of nonfinal conjuncts at LF that other languages' $\&^\circ$s do. This means that languages which permit nonstandard Case in final conjuncts will also allow it in nonfinal ones; languages which prohibit nonstandard Case in coordination-final position will prohibit them elsewhere as well; this prediction appears to hold crosslinguistically.
5.3 The Identity Puzzle
The identity puzzle pointed out the fact that all nonfinal conjuncts must bear identical Case; recall for example the superiority of 'Him, her and Robin left' over '*Him, she and Robin left'. This fact also falls out immediately under the present analysis. As noted in section 4, all [Spec &P] terms lie in the checking domain of &° and simultaneously undergo feature-checking at LF. An &°-chain can only check off a single set of features; any conjunct bearing an aberrant feature causes the form to crash. This directly explains the data in (4) and (5). All nonfinal conjuncts will bear identical Case--standard or nonstandard --because the &° will seek to check off, for instance, only [+NOM] or only [+ACC].

5.4 The Final-Conjunct Puzzle
The distinction between a checking domain and complement domain that the &P-shell structure creates also enables a simple explanation of the final conjunct puzzle, which actually boils down to nothing more than the nonstandard puzzle and the identity puzzle data combined. The final conjunct in English stands exempt from having to bear the same Case as all nonfinal ones; this because the [Comp, &°] does not lie within the checking domain of the &°. The final conjunct receives Case from the &°, but the grammar needs nothing beyond the aforementioned parameter in (20) to account for all the final-conjunct puzzle data.

6.0 Conclusion
NP-coordinations have unique properties, as the four puzzles given at the beginning of this work indicate. Previously, no theory had managed to capture the facts in much more than a descriptive way. The theoretical move to the &P-shell structure, however, relates all the puzzles to two factors: Case-assignment to [Comp, &°] and feature-checking of [Spec, &P]. It does so within the theoretical requirements of Minimality. The success that the &P-shell analysis enjoys in solving these four puzzles suggests that continued investigation will likely reveal further empirical benefits of the analysis.

NOTES

My thanks go to Robert May for helpful commentary on an earlier version of this work.

1. The structure also can account for the only other possible surfacing of conjunctions; that of an &° appearing between each conjunct. A form such as 'Robin and Kim and Terry and Pat', which carries a degree of emphasis, results when the &° raises at PF rather than LF, and all the traces immediately copy the phonetic content of this antecedent. The underlined terms below show 'traces
come to life' under this idea:

i. \[ \text{&} \text{Robin [\& and [\&F Kim [\& and [\&F Terry [\& and Pat]]]]] } \]

The above actually just provides one instance of the general phenomenon that traces may assume phonetic content for purposes of emphasis; see Zoerner (1994) for details.

2. This forces a slight revision of the Case Filter, which becomes a requirement on syntactic positions rather than on NPs per se. Under the new idea, positions such as \[ \text{[Spec, IP] (in finite clauses) and [Comp, V°] } \]

must contain a Case-bearing NP. Not every NP need receive direct assignment of Case, however.

3. The claim that overt &°s assign Case better than null &°s do has a correlate in Larson's (1988, 1990) VP-shell analysis. Larson depicts the underlying structure of a double-object construction such as 'Robin sent a letter to Kim' as:

i. \[ \text{[VP Robin [\text{\textunderscore V} a letter [\text{\textunderscore V} sent to Kim]]] } \]

The verb 'sent' must raise to fill the empty position, says Larson, because the empty V° slot cannot assign Case on its own; it requires lexical content.

REFERENCES


PARASESSION:
HISTORICAL ISSUES IN
SOCIOLINGUISTICS/
SOCIAL ISSUES IN HISTORICAL
LINGUISTICS
Diachronic Aspects of Russianisms in Siberian Turkic

Gregory D. S. Anderson

University of Chicago

0.1 Languages belonging to the Turkic family represent one quarter of the approximately three dozen indigenous languages of Siberia; these include Yakut, Tuvan, Xakas, Shor, Dolgan and Tofalæ. The Russian language has had a profound influence on the Turkic languages of Siberia during the past centuries. Of course, during this time numerous Russian lexical items have entered the Siberian Turkic languages. At the earliest period, loans were frequently limited to terms in the military/government, trade, and religious spheres and were predominantly borrowed from a variety of spoken Russian vernaculars, undergoing assimilation to the phonology of the Siberian Turkic language in question. In the Soviet period, as bilingualism spread and literacy in Russian became mandatory, loans became extremely numerous, including such syntactic items as coordinating conjunctions, formerly lacking as a class in the Siberian Turkic languages. As the generations of Turkic-Russian native bilinguals matured, a partly conscious process somewhat akin to 'deculturalization' was witnessed in the pronunciation of many loans especially among young urban-dwelling speakers: an earlier assimilated form came to be replaced gradually by the actual pronunciation of the Russian source, often violating the phonotactics of a "native" utterance; this cline of pronunciations from completely assimilated to completely Russian varied individually according to such sociolinguistic and demographic factors as level of education, daily contact with Russians or urban vs. rural residence, etc. Today, codeswitching phenomena are frequently attested in the speech of the Siberian Turks. While some of the Siberian Turkic languages have enjoyed a boom in terms of overall number of speakers in the last few decades, e.g. Yakut or Tuvan, the future of others, e.g. Shor or Tofalæ, is bleak. In addition, Chulym Turkic and Baraba Tatar have already been largely assimilated to either Russian, Tatar, or both, and are omitted from most of the discussion below.

Phonologically speaking, the Turkic languages of Siberia are characterized by a more developed use of consonantal assimilation than in other Turkic areal groups, and a series of diachronically and morphophonemically active intervocalic weakenings of velars; also, consonant clusters are limited, and in some languages geminates are frequent, e.g. Yakut uos 'lip' < *ayiz, appit 'our.horse' < *at(i)miz, akkit 'your.horse' < *at(i)niz; Xakas sanaa 'to.the.ski' < *sanaya, müüs 'horn' < *amüntiz; Upper Chulym soooq Middle Chulym suaq 'cold' < *soyúq; Baraba xannan 'from.the.khan' < *qandán, paštì 'head[Accusative]' < *bašni; Tuvan saqtìp 'having thought' < *sayinip; Shor tutpa 'don't hold!' < *tutma; Altai attìn 'of.the.horse' < atini; Tofalæ qannonbas 'not.shamanizing' < *qamlamaz.

1.1 Development of the Siberian Turkic literary languages

Four Siberian Turkic languages enjoy status as literary languages in their respective regions, viz. Yakut, Tuvan, Xakas, and Altai; Shor had a literary language until 1940 when the issuing of publications in the Shor language was ceased; recent efforts have been made to revive this language.2 In the years following the Revolution, Latin-based alphabets were developed for Xakas, Altai, Tuvan3, Yakut4, and Shor; as official policy, this was abandoned in favor of
Cyrillic-based alphabets for all the Siberian Turkic languages by 1941. During this period, the influx of Russianisms became quite pronounced in these languages, and contrary to the early contact period, the loans came predominantly from the written medium, aided greatly by the mass importation of Russian loans into the fledgling Turkic literary languages. Also, although instruction was officially to be offered in a variety of indigenous languages during the 1930's, in reality Russian was used almost exclusively due to the near complete lack of suitable materials and non-Russian speaking teachers (Slezkine 1994: 243). In the post-World War II era, as the titular republics and regions of the Siberian Turkic peoples became increasingly inundated with Russian immigrants (and those of other Soviet nationalities), so too did the pressure from the Russian language become particularly intense. Naturally, as the contact with Russians increased, so did the incidence of bilingualism, and ultimately, of full linguistic assimilation.

1.2 Recent Soviet census findings

In the Soviet era, one must rely on the erratic and politically motivated Soviet census findings. According to Silver (1986: 76), the number of native language speakers of various Soviet minority languages may have been artificially inflated by the census' explanatory comments: if unable to decide on a native language, the respondent was instructed to choose the dominant language used in the family during early childhood. Establishing command of a second language, especially of Russian, is basically subjective, and frequently is underestimated in official counts (Silver 1986: 92); here, too, the wording on the census itself causes some confusion and inaccuracy, as the question asks for a second language of the USSR that the respondent 'freely commands'. Obviously a certain amount of politics and ethnic self-identity play a role in determining percentages of bilingualism in Russian. For example, the Uzbekhs went from 15% bilingualism in the 1970 census to 49% in 1979 but back down to 23.8% in 1989 (Silver 1986: 90; Ryan 1990: 112), numbers which could in no way reflect actual reality. Percentages of bilingualism probably run higher for most of the Siberian Turkic peoples than are listed officially in census statistics.

With the statistical caveats aside, several basic developments in Russian-Turkic language contacts in Siberia during the Soviet period can be determined. The larger Siberian Turkic groups, e.g. Tuvan or Yakut, have preserved their languages to a greater degree than the smaller ones, e.g. Tofalar or Shor, which are rapidly losing out to Russian. Among the Tuvan, Yakut, Xakas, and Altai--i.e. the Siberian Turkic peoples possessing literary languages--one generally sees a slight, though perceptible, trend to greater assimilation to Russian and a massive increase in bilingualism when comparing census data from 1926 and 1959, to the later Soviet censuses of 1970, 1979, and 1989. Tuvans have had the fewest percentage of their population that have given up their mother tongue in favor of Russian; indeed, nearly a full 99% of Tuvans listed Tuvan as their native tongue even in the 1989 census figures. As mentioned above, the Baraba Tatars (8,000 in the 1959 census) have been largely assimilated to the Russians and Tatars, whose language represents their medium for schooling. The Chulym Turks are likewise not distinguished from the Tatars nor the Russians in census statistics, but numbered 4,500 in 1959; most Chulym Turks have been assimilated to Russian linguistically, having already been assimilated culturally to the surrounding populations. Both of these groups are omitted from the data below.
For most of the Siberian Turkic languages there has been a gradual increase in percentage of people who consider Russian to be their native language rather than the Turkic ancestral tongue. Compare the data on people who consider their native language to be Turkic from the 1926, 1959, 1970, 1979, and 1989 census readings:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>YAKUT</td>
<td>99.7%</td>
<td>97.6%</td>
<td>96.3%</td>
<td>95.3%</td>
<td>94.0%</td>
</tr>
<tr>
<td>DOLGAN</td>
<td>99.5%</td>
<td>93.9%</td>
<td>89.8%</td>
<td>88.2%</td>
<td>84.0%</td>
</tr>
<tr>
<td>TU VAN5</td>
<td>------</td>
<td>99.1%</td>
<td>98.7%</td>
<td>98.8%</td>
<td>98.6%</td>
</tr>
<tr>
<td>XAKAS</td>
<td>89.9%</td>
<td>86.4%</td>
<td>83.7%</td>
<td>80.9%</td>
<td>76.6%</td>
</tr>
<tr>
<td>SHOR</td>
<td>93.7%</td>
<td>84.1%</td>
<td>73.5%</td>
<td>61.2%</td>
<td>57.5%</td>
</tr>
<tr>
<td>ALTAI</td>
<td>75.7%</td>
<td>89.1%</td>
<td>87.2%</td>
<td>86.4%</td>
<td>85.1%</td>
</tr>
<tr>
<td>TOFALAR</td>
<td>29.1%</td>
<td>89.1%</td>
<td>56.3%</td>
<td>62.1%</td>
<td>42.8%</td>
</tr>
</tbody>
</table>

In addition to the ever-increasing percentages of people abandoning the ancestral Turkic language in favor of Russian, a large percentage of the Siberian Turkic speakers consider themselves fluent in Russian; generally an increase is found when comparing data from 1970, 1979, and 1989:

<table>
<thead>
<tr>
<th></th>
<th>1970 %</th>
<th></th>
<th>1979 %</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFALAR</td>
<td>90.1%</td>
<td>TOFALAR</td>
<td>96.1%</td>
<td></td>
</tr>
<tr>
<td>SHOR</td>
<td>81.1%</td>
<td>SHOR</td>
<td>85.8%</td>
<td></td>
</tr>
<tr>
<td>XAKAS</td>
<td>78.2%</td>
<td>XAKAS</td>
<td>84.2%</td>
<td></td>
</tr>
<tr>
<td>DOLGAN</td>
<td>69.5%</td>
<td>DOLGAN</td>
<td>81.1%</td>
<td></td>
</tr>
<tr>
<td>ALTAI</td>
<td>62.7%</td>
<td>ALTAI</td>
<td>79.4%</td>
<td></td>
</tr>
<tr>
<td>YAKUT</td>
<td>43.1%</td>
<td>YAKUT</td>
<td>58.3%</td>
<td></td>
</tr>
<tr>
<td>TU VAN</td>
<td>39.3%</td>
<td>TU VAN</td>
<td>60.0%</td>
<td></td>
</tr>
</tbody>
</table>

i. 1989 % bilingualism in Russian

ii. urban vs. rural bilingualism

<table>
<thead>
<tr>
<th></th>
<th>urban</th>
<th>rural</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHOR</td>
<td>93.2%</td>
<td>84.5%</td>
</tr>
<tr>
<td>TOFALAR</td>
<td>91.6%</td>
<td>81.3%</td>
</tr>
<tr>
<td>XAKAS</td>
<td>87.7%</td>
<td>74.6%</td>
</tr>
<tr>
<td>DOLGAN</td>
<td>81.3%</td>
<td>63.2%</td>
</tr>
<tr>
<td>ALTAI</td>
<td>77.1%</td>
<td>52.3%</td>
</tr>
<tr>
<td>YAKUT</td>
<td>69.1%</td>
<td></td>
</tr>
<tr>
<td>TU VAN</td>
<td>60.0%</td>
<td></td>
</tr>
</tbody>
</table>

During this same period (1970-89), the Turkic-speaking peoples of Siberia frequently lost ground percentagewise to the (predominantly immigrant) Russian population in their titular republics and oblast's; of these, only Tuva still represented a majority in the Tuvin Republic, while the Yakut, Khakass, Dolgan, and Altai all constituted minorities in their respective administrative regions, although ethnic Altai indeed constituted a greater percent of the population of Gorno-Altai in 1989 than in 1970.7

<table>
<thead>
<tr>
<th></th>
<th>1970</th>
<th>1979</th>
<th>1989</th>
</tr>
</thead>
<tbody>
<tr>
<td>TUVA ASSR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TU VAN</td>
<td>58.6%</td>
<td>60.5%</td>
<td>64.3%</td>
</tr>
<tr>
<td>RUSSIAN</td>
<td>38.3%</td>
<td>36.2%</td>
<td>32.0%</td>
</tr>
</tbody>
</table>
2.0 Diachronic aspects of Russianisms in Siberian Turkic languages

At the earliest period of Russian-Turkic contacts, loans came from a variety of Russian dialects, and thus an apparent lack of consistency in the assimilation of the Russian loans to Turkic phonology is seen.\textsuperscript{8} Over time, although some variation was still to be seen, due partly to transcription of the Turkic words by untrained ears, local variation on the Turkic dialectal level, as well as a continuation of various Russian dialect loan sources, there were nevertheless general tendencies in loan assimilation that could be observed within a given Siberian Turkic language. In the Soviet era, a variety of calques, loan translations, neologisms, and various blends thereof were initially tried as means to supplement the lexicons of these languages, e.g. Yakut olo\textit{x} tur\textit{u}ga 'economics' (literally ['life'] ['standing-POSS']) or Tuvan demobilizasta\textit{a}š\textit{k}in 'demobilization' or militaris\textit{t}idilge 'militarization'; however Russian forms soon won out. As mentioned above, the growth of bilingualism in Russian that triggered a tendency to 're-Russianize' the (previously) assimilated loans was observed among some, especially urban-dwelling, members of the Turkic populations. While for some speakers loanwords in Russian orthography\textsuperscript{9} in no way reflect actual pronunciation, among young, urban bilinguals, these represent the actual pronunciations observed. In fact, the speech of these Siberian Turks, like that of many natively bilingual peoples in a similar socio-cultural milieu, is today sprinkled generously with Russian words and phrases.

2.1 Yakut: Assimilation to Codeswitching

The case of Yakut (Sakha) typifies Russian's impact on the Siberian Turkic languages. First contact between speakers of these two languages happened sometime in the early seventeenth century. By 1920, approximately 2800 Russian loans are found in Yakut (Sleptsov 1964: 12). The influence of Russian grew ever stronger in the Soviet period. As of 1989 69% of the Sakha people considered themselves bilingual in Russian; in many cities this approaches 100%.

During the past centuries of Russian-Yakut contacts, loans were assimilated to the rather un-Russian phonology of Yakut, e.g. kūbūörune < губерния [guberniya] 'province', bürištüüpüünlük < преступник [prestupnik] 'criminal' or nehiliyek < наслег [naslek] 'administrative unit of Tsarist Yakutia similar to the
Soviet era sel'sovet'. With respect to vocalism, the following tendencies are to be observed. Stressed vowels are generally replaced by long vowels or diphthongs, with the word-stress shifting to the last syllable, as word-final syllable stress is regular in Yakut. Unstressed vowels are frequently replaced by harmonically conditioned variants. Initial r- in Russian gets a prothetic A- (a harmonizing [−high] vowel), while initial clusters get either an epenthetic vowel after the initial consonant (frequently a copy of the following stem vowel) or prothetic I-. The 'yotacized' vowels of Russian (especially я [ya] and ю [yul]) are generally replaced by front vowels in Yakut. In early loans, Back Harmony restrictions are rarely violated, and never so if the stressed vowel in the Russian source is in the initial syllable of the word. In most Yakut dialects, the harmonic pattern is determined by the placement of original stress in the Russian source, with bidirectional spread, if applicable; in Kolyma-area loans, the vowel harmony pattern is always determined by the initial syllable, à la 'native' Yakut words, regardless of where the stress fell in the loan source.

Among the tendencies in the assimilation of Russian consonantism to Yakut phonology, the following can be enumerated. Initial *p- and *f- were replaced by [b-]10, while medially, the replacement sound in Yakut was [−p(p)-]. Russian *v always was realized as Yakut [b]. All labial sounds sometimes appeared as [m] as the result of sporadic, distant assimilation to a nasal sound in the word, a process that has been active in the Turkic languages for centuries (cf. the benimen isoglosses of the Old Turkic period). Russian *g- was generally replaced by [k-], which along with Russian *k- further changed to [x-] before [a] and [o], in accordance with Yakut phonotactics.

Russian *n is realized as [n] in Yakut when preceding *i (or less frequently *e); in addition, there are a variety of sporadic [n] and [l] correspondences between Russian sources and Yakut loans, as well as in the pronunciation of individual speakers of Yakut. A further characteristic of the assimilation of Russian loans to Yakut phonology is the multiplicity of sources for [s]. In final position, Yakut [−s] may reflect Russian *-č, *-c, *-s, *-š, *-z, *-ž, or *-šč, and all of these except *č-
in initial position. In medial position, the various developments are complicated by the frequent change of s > -h- in intervocalic position.\textsuperscript{11} Russian medial clusters of three or more consonants and final clusters of any size were reduced in a variety of lexically specific ways.

(6) ńukulay < Николай [Nikolay] 'Nikolai'
nireexi < неряха [nerya] 'sloven(ly)'
sonouk < залог [zalok] 'guarantee'
siertibe < жертьа [žertvo] 'victim'
čeppier < четверг [četverk] 'Thursday'
čierbe < червь [červ'] 'worm'
xaaltis < галстук [galstuk] 'tie'
biraaskay < братский [bratskiy] 'brotherly'

During the course of the twentieth century, the Sakha people have become increasingly familiar with Russian. Words of Russian origin in many people's speech began to gradually sound more like their Russian sources. For example, initial [p-] became common, as did the sequences [ka-] and [ko-], e.g. pampiliat < памфлет [pamflet] 'pamphlet', kaaska < каска [kaska] 'helmet', or kollokubuyum < КОЛЛОКВИУМ [kollokviium] 'colloquium', where one originally would have found [b-], [xa-], and [xo-], respectively; also some medial and final clusters began to be pronounced: taranspartir < транспортиров [transportir] 'protactor' or kuurs < курс [kurs] 'course' (Sleptsov 1975: 109ff.). In the present, one frequently finds forms that are pronounced by many urban dwellers in exactly their Russian pronunciation, often violating both Yakut syllable cannons and vowel harmony restrictions alike, for example gruzovik, gips, rasa, rtut', vrač, grafstvo, xronika, etc. In addition to the speech of young urbanites, examples of codeswitching can be found in the dialogue of characters in the works of various Sakha authors.

(from SLEPTSOV 1975: 7)

(7) i. ȉi ڇín kíñler pravil'naia suox povedenìlyarinan aayabin: üle pokazatele ürdüüriin ihin oxshuu ljuboj miysteje, ljuboj vremeye, besprestanno, ol üle mexanizacija laayittan, mexanizacijata suyúttan nezavisimo ñiittiix tusax.

'This I consider their incorrect conduct: the struggle against the rise of the index at work must be observed in any place (or) any time, continuously, regardless of whether this work is mechanized or non-mechanized.'

ii. uot xaya ere Kiŋ Maarga barbitin xantan bilen biha sonno, imenno onno, tiydiŋ? interesno, a? ne tak li?

'From where did you learn that fire appeared in some Kiŋ Maar and came directly just here? Interesting, huh? Isn't that so?'

2.2 Tuvan: Assimilation to Codeswitching
The history of Tuvan's contacts with Russian differs in several ways from that of Yakut's. First, the initial contact between Tuvans and Russians was not until the 1830's, some 200 years following the initiation of the yasak fur-tribute among the Sakha. However, Russian influence began to be felt fairly rapidly in the early twentieth century; according to Tatarintsev (1974: 15), in 1888 there were approximately 150 Russians in Tuva but by 1918, this number had risen to nearly 12,000, or 20% of the population. At the early period (pre-20th century), virtually no Tuvans spoke Russian, but many Russians in Tuva spoke the local language. By the time of the Revolution, some bilingualism had begun, which increased rapidly following the assimilation of Tuva into the USSR in 1944. Although Tuvans have preserved their language remarkably during the Soviet period (still 99% of the population is native speaking in 1989), 60% of Tuvans consider themselves fluent in Russian; in present-day Kyzyl, where bilingualism approaches 100%, sentences with codeswitching are not infrequently encountered.

As in Yakut, early loans were assimilated to Tuvan phonology; more recent loans are likewise pronounced by elderly persons and those less fluent in Russian in an assimilated form, although a certain amount of lexically-determined 're-Russinization' can be observed, with considerable individual variation. In order to assimilate to Tuvan phonology, Russian words generally did not have to undergo as severe modifications as was observed in Yakut. For example, Tuvan possesses [8], [2], and [v], all lacking in Yakut but found in Russian; in addition, Tuvan has a less developed system of round harmony than does Yakut. Of course, as in Yakut, unfavored syllable structures were altered, stressed vowels were treated as long and as the trigger for the vowel harmony pattern, to which unstressed vowels were assimilated. Among consonants, both voiceless and voiced anterior stops were considered weak in the weak/strong opposition in Tuvan (cf. Sat 1973: 95-6), i.e. they are realized as [d] and [b].

(8)  

<table>
<thead>
<tr>
<th>Tuvan</th>
<th>Russian</th>
<th>Pronunciation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>biristaa</td>
<td>верста</td>
<td>[versta]</td>
<td>'Verst'</td>
</tr>
<tr>
<td>oodurba</td>
<td>отрубы</td>
<td>[otrubi]</td>
<td>'Bran'</td>
</tr>
<tr>
<td>mombuš</td>
<td>помочь</td>
<td>[pomoč']</td>
<td>'Help'</td>
</tr>
<tr>
<td>moža</td>
<td>вожжи</td>
<td>[vožći]</td>
<td>'Reins'</td>
</tr>
<tr>
<td>sappik</td>
<td>сапог</td>
<td>[sapok]</td>
<td>'Boot'</td>
</tr>
<tr>
<td>ögü(:)rze</td>
<td>огурец</td>
<td>[ogurec]</td>
<td>'Cucumber'</td>
</tr>
<tr>
<td>çaadir</td>
<td>театр</td>
<td>[teatr]</td>
<td>'Theatre'</td>
</tr>
</tbody>
</table>

As the level of familiarity and comfort with Russian grew among Tuvans, so too did the number of words that were affected by the re-Russinization process. Sat (1973: 94ff.) lists several steps in this process that one is likely to encounter in these 'decreolized' variants in individual Tuvan's speech, e.g. introduction of palatalized sounds, restoration of clusters through loss of epenthetic vowels (kilass > klass), 'de-harmonizing' Russian words (sekredeer > sekretaar'), and introduction of non-Tuvan sounds (ögürze > ogurees > ogurec). Mongush (1983: 57) considers the various degrees of re-Russinization of lexical items in the speech of Tuvans to reflect various stylistic registers or 'gradations', the use of which are determined by the usual sociolinguistic and demographic factors.
<table>
<thead>
<tr>
<th>Language</th>
<th># w/Tk. lg as native lg.</th>
<th># w/Tk. lg as 2nd lg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>YAKUT</td>
<td>1205</td>
<td>5654</td>
</tr>
<tr>
<td>TUUVAN</td>
<td>11</td>
<td>482</td>
</tr>
<tr>
<td>XAKAS</td>
<td>21</td>
<td>426</td>
</tr>
<tr>
<td>ALTAI</td>
<td>23</td>
<td>618</td>
</tr>
</tbody>
</table>

As is obvious from (13), a significantly larger number of Russians command Yakut both natively and as a second language than other Siberian Turkic languages. In Yakutia a kind of mutual bilingualism in Russian and Yakut arose in various settlements in the Kolyma region in the northeast; for 300 years these peoples intermarried, and, according to P. Ryabakov (cf. Korkina 1989: 64), the Russian population of Srednekolymsk knew Yakut better than Russian at the end of the nineteenth century; the relative prestige status of Yakut in northeastern Yakutia is further demonstrated by the widespread adoption of Yakut as the mother tongue of the Evenki (85%), Even (70%) and Yukaghir (23%) of this area (Slezkine 1994: 101, 1989 census data).

---

1 According to Rassadin (1972: 333), by the early 1970's Tofalar children rarely spoke the language, frequently answering their Tofalar speaking parents in Russian.

2 For example, according to Fænsen (1983), in 1980 there were the following publications in the Siberian Turkic languages: Yakut 75 monographs, 2 periodicals, 28 newspapers, Tuuvan 55/2/4, Khakass 10/1/1, Altai 23/1/1. The issuing of literature in Shor was halted in 1940 (though in 1939 there were 30 monographs and 1 newspaper printed in Shor). However, according to Kempe (1992: 78), a new program in Shor studies was established at Novokuznetsk University, and a few elementary schoolbooks and a small dictionaries were issued in Shor in 1991.

3 During the first years of the Tuvan People’s Republic Mongolian was the literary language of the Tuvans, so the dates on the establishment of a Latin-based alphabet was a little later for Tuvan than for the other Siberian Turkic peoples.

4 There were early, mainly missionary works issued in both Yakut and Altai utilizing the Cyrillic alphabet; also, early studies on the Siberian Turkic languages, e.g. Boehlingk (1851, reprinted 1964) or Katanov (1903) likewise used a modified Cyrillic orthography.

5 Tuva was the independent Tuvan People’s Republic (Tannu-Tuva) during the 1926 Soviet census.

6 This increase in percentage probably reflects the fact that many of the fully linguistically assimilated Tofalar (Karagass) of the 1926 census had begun listing their nationality as Russian by 1959; the overall number of people identifying themselves as Tofalar decreased dramatically during this period (from 2829 to 586).

7 How the post-Soviet exodus from Siberia of ethnic Russian (et al.) immigrants, especially in the Far North, will affect the demographics of the titular regions of the Siberian Turks, particularly Yakutia, remains to be seen.

8 For example in Pekarsky’s masterful dictionary of 1907-30, 64% of the loans had more than one variant listed (Sleptsov 1964: 99), with some due simply to (ortho)graphic variation; even so, this percentage of variants is not inconsiderable when coupled with the generally monolithic nature of Yakut, which exhibits only slight dialectal variation. Variations can be relatively insignificant phonetically, e.g. aptaniamaiya ~ aptanumiyia ~ aptanuomuya ~ abtanumuya ‘autonomy’, or the variants can be considerably divergent from one another, e.g. kiniral ~ žanaraal ~ ŋadaraal ‘general’ (Sleptsov 1975: 53-9, 61).

9 As of 1977, by official decree, most words of Russian origin were to be written in Russian orthography in all languages of the former USSR using the Cyrillic alphabet. Only those early loans which were fully assimilated and sufficiently dispersed among the dialects and speakers of a given language preserved their altered orthographic form, e.g. Yakut ostuol ‘table’ < сто́л [stol].
The Xakas have a high rate of bilingualism in Russian in the present day; in some raions of Xakasia, e.g. Bogradskii, reported bilingualism approaches 95% (Krivonogov 1984: 166). The only Xakas who do not speak Russian are over 60 and live in extremely rural areas.

(from TIUKPIEKOV 1993: 4)
(11) i. ponimaes', potrava üčün nime polar saγaa?! Xoy, xadarbasta, začem čaban polča?

'Do you understand, what the crop-damage will mean?! The sheep won't graze, what is the shepherd for?'

ii. mexanizatorlar toγizin uvažat' polbinča...pu vreditel'stvo!
Vreditel'ler ibire toldıra! Ničevo, tik xalyışpaspın! Tölederbln!
Poniames', krugliy god töliržîn!

'the machine specialists show no respect... this is sabotage!. The saboteurs are all around! It's nothing, I won't leave free of charge! I'll make (you) pay! Do you understand, you will pay year round! '

Similar phenomena are attested in other Turkic languages of Russia as well. For example, one frequently encounters Russian and Bashkir words within a single sentence (from Garipov 1969: 154).

(12) mineŋ bratimdä frontta ranil ittelär
'my brother was wounded at the front'

irtägä televizorža kakuy peredaça bula
'what show is on the television tonight?'

Min bögün vsiu noč ne spal
'I didn't sleep all night today'

4.4 Russian Knowledge of Siberian Turkic languages

While at present the bilingualism of the Siberian Turks in Russian is generally unilateral in nature, i.e. the Turkic population speaks Russian and their native language, and Russians rarely know the local language, this has not always been the case. In fact, at the earliest contact period, Cossacks and merchants frequently spoke the local Turkic vernacular, e.g. Tuvan, Xakas, or Yakut, while the Turkic people rarely knew Russian, i.e. the exact reverse-type of unilateral bilingualism was found. However, at present knowledge of a local Turkic language is not unknown among ethnic Russians in Siberia. In fact, as of 1979 (Chisl 1984: 80-4), there were people who ethnically identified themselves as Russians but who listed their native language as the local Siberian Turkic language in four different areas in Siberia: Yakutia, Tuva, Xakasia, and Gorno-Altai. In addition, a small percentage of Russians consider themselves bilingual in the local Siberian Turkic language.
(from TATARINTSEV 1974: 53-4)

(9)

i. Ėto značit (*ol deerge/*inčangaš) amgi üenig
gradostritel'stvozunun (*xooray tuduškununun) deynelin kõör,
ĉuruktu peredelat' (*ede kîlîr).

'that means (it is necessary to) proceed from the standard of contemporary
town-building (and) alter the plans.' {from a meeting}

ii. bo aytirigni putat'tap (*buduldurup) algandîr men. zanimat'sjalap-la
(*öorenip-lel*kičeelep-le) turgan kîži men.

'I bombed on the question, (but) I did study.'

iii. diñnadiñ be, xudožestvenniy svist azi, tivalaarga, čūū deer čüvel..iye,
uran sigît deer čüve dam čitti?

'did you hear that they reproduced the artistic whistling, or how is it called
in Tuvan..yes uran sigît?'

2.3 Xakas: Assimilation to Codeswitching

The history of contact and interaction between Xakas and Russian follows
the same basic pattern as for the other Siberian Turkic languages. Xakasia became
part of Russia in 1707, and the number of Russians steadily increased over the next
two centuries. As bilingualism grew during the Soviet period, the loans gradually
took on the shape of their Russian sources in the pronunciations of various Xakas
individuals. As Xakas exhibits a wide range of dialectal variation, the sounds
within an older loan from Russian could have either been preserved or altered
depending on the dialect in question. Long vowels in Russian loans appear less
frequently in Xakas than in Tuvan and especially Yakut; also, clusters with /l/ and
sometimes /h/ are preserved word-initially in Xakas.

(10) putulka < бутылка [butilka] 'bottle'
kirles < крьльцо [kri'lco] 'porch'
pilke < вилка [vilka] 'fork'
istan < штани [stani] 'trousers'
is kemeyke < скамейка [skameyka] 'bench'
klüs < ключ [klyut'č] 'key'
pool < пол [pol] 'floor'
krede < гряды [gra'da] 'bed (of garden)'

Xakas differs from both Tuvan and especially Yakut in possessing a weaker vowel
harmony system, with systematic morphemic violations of Round Harmony, and
vowels neutral to Back Harmony. Some dialects, e.g. Bell'tir, show even less
stringent harmony requirements, and the system has begun to disintegrate
altogether.
According to Sleptsov (1964: 115), in certain areas where contact with Russians was intense, [p-] was preserved as a conscious effort on the part of Sakha to Russianize their pronunciation, even when the rest of the word was highly altered; there are approximately 200 loan words beginning with [p-] in pre-Revolutionary lexical sources (e.g. Pekarskij and Kulakowskij), a feature especially common in the speech of the Anabar Sakha.

Note that among many Yakut speakers even today, words that are pronounced in the nominative case with these final sounds as in Russian, nevertheless show an alternation with [-h-] in inflected forms, e.g. garağ > garahü, plașč > plahu (Sleptsov 1975: 134). Also, in general, nominative forms of Russian loans may be atypical in terms of stress placement in a given Turkic language, but inflected forms are often treated regularly (cf. Baskakov 1972: 78).

**References:**


**Chislennost’ i sostav naseleniiia SSSR.** 1984. Moscow: Finansy i Statistika.


**Itogi svesoiuznoi perepisi naseleniiia (SSSR) 1959 goda.** 1962. Moscow: Gosstatizdat.

Katanov, N. F. 1903. *Opyt issledovania uriankhaiskogo iazyka*. Kazan'.


Naselenie SSSR. 1990. Moscow: "Finansy i Statistika".


Vsesoiuznaia perepis' naseleniiia 1926 goda SSSR t. IX: RSFSR. 1929. Moscow: TsSU SSSR.
1.0 INTRODUCTION
Oral discourse in many languages is characterized by the occurrence of discourse markers, or PRAGMATIC MARKERS, which in Modern English include forms such as so, well, like, I mean, or you know. While numerous sociolinguistic studies have been concerned with the identification of these forms in Modern English (ModE) and with a specification of their functions, which include both the structuring of discourse and the maintaining and furthering of social interaction,1 this paper explores the existence of such forms from a diachronic perspective in the texts of Old English (OE) and Middle English (ME).

While it may initially seem implausible to speak of pragmatic markers in texts which exist only in written form, it is generally agreed that medieval discourse exhibits many structural and linguistic elements characteristic of oral discourse, what Ong (1984:3) has called an 'oral residue'. For example, Green (1990:270-72), while questioning many of the tenets of the so-called 'oral-formulaic theory', has acknowledged a fundamental 'symbiosis of orality and writing' in the middle ages, and Fleischman (1990:23) points out that 'many of the disconcerting properties of medieval vernacular texts ... can find more satisfying explanations if we first of all acknowledge the extent to which our texts structure information the way a spoken language does'. When reexamined within the framework of contemporary discourse analysis, these 'disconcerting' features of medieval discourse—traditionally dismissed as fillers, metrical expedients, or other defects of style—may provide the potential candidates for pragmatic markers. Some of these MYSTERY FEATURES (to adapt Longacre's 1976 term) are listed in the Appendix along with studies which implicitly or explicitly treat them from a discourse analytic perspective.

2.0 THE 'MYSTERY FEATURES' OF OLD AND MIDDLE ENGLISH
Undertaking a kind of 'historical discourse analysis/historical text linguistics' (Enkvist and Wärvik 1987:222), or what Fleischman (1990:37) terms 'New Philology',2 this paper investigates whether the mystery features of Old and Middle English:
(a) have the syntactic, semantic, and distributional characteristics of Modern English pragmatic markers,
(b) share the textual and interpersonal functions of their Modern English counterparts, and
(c) develop diachronically in a way analogous to the postulated synchronic development of the modern forms.
The discussion is based on a book-length study of pragmatic markers in English (Brinton in press), with additional support from the studies cited in the Appendix. The main forms considered are exemplified below.3
(1) OLD ENGLISH

a. Hit gelamp þa æfter þyrste þæt þa unge-leafluffan hæðenan ge-bundon þone bisceop. and to ðære byrig rautenna geleddon on bendum (Ælfric, Lives of saints [St. Apollinaris] 154-56).
   ‘It happened after a time that the unbelieving heathens bound the bishop and led him in bonds into the city of Ravenna’.

b. Hwæt, we for dryhtene iu dreamas hefdon, / song on swegle selrum tidum (Christ and Satan 44-45).
   ‘What, we before had joys in front of the lord, song in the heavens in better times’.

   ‘then Fabianus ordered that he should offer incense to the foul Jove, or himself walk over burning coals with bare fee. What then Tibertius went boldly over the burning coals with unburned feet’.

(2) MIDDLE ENGLISH

a. Criseyda gan al his chere aspien. / And leet it so softe in hire herte synke. / That to hireself she seyde, ‘Who yaf me drynke?’ (Chaucer, Troilus and Criseyde II, 649-51).

b. And so they lay todydir untill underne of the morne; and all the wyndowys and holys of that chambir were stopped, that no maner of day myght be seyne. And anoné sir Launcelot remembryd hym and arose up and wente to the wyndow … (Malory, Le Morte d’Arthur 11, 2, 795, 17-20).

c. And so bifel it on a Saterday. / This carpenter was goon til Osenay (Chaucer, The Canterbury Tales [‘Miller’s Tale’] A.3399-400).

d. Whan sir Bors was departed frome Camelot he mette with a religious man … (Malory, Le Morte d’Arthur 16, 6, 955, 1).

e. And cryde ‘Awake!’ ful wonderlich and sharpe; / ‘What! Slombrestow as in a litargie?’ (Chaucer, Troilus and Criseyde I, 729-30).


g. ‘I am yong and unkonnyng, as thow woost, / And as I trowe, with love offended moost …’ (Chaucer, The Canterbury Tales [‘Knight’s Tale’] A.2393-94).

h. That in his herte he demed, as I gesse. / that ther nys lover in this world at ese/ So wel as he (Chaucer, Troilus and Criseyde III, 1727-29).

i. ‘His answere shal she have, I undertake’ (Chaucer, The Canterbury Tales [‘Merchant’s Tale’] E.2317).

j. I have, God woost, a large feeld to ere, / And wayke been the oxen in my plough./ The remenant of the tale is long ynoough (Chaucer, The Canterbury Tales [‘Knight’s Tale’] A.886-88).

k. For truste wel, hir flesshly love was deedly hate (Chaucer, The Canterbury Tales [‘Parson’s Tale’] I.204).
2.1 Characteristics. The syntactic, semantic, and distributional properties of OE and ME mystery features can be compared with those identified in studies of pragmatic markers in Modern English:

(a) **Marginality in word class**: Lexicographers are at a loss, for example, to account for OE *hwæt*, calling it an ‘interjection’ or a ‘particle’.

(b) **Heterogeneity of forms**: OE and ME mystery features include individual lexical items such as ME *gan* (a verb) or *anon* (an adverb), phrases such as ME *for the nones*, reduced clauses such as ME *I gesse* or *ye knowen*, and full clauses such as OE *pa gelamp þæt* or ME *then* (it) *happened that*.

(c) **High frequency (in oral discourse)**: ME *anon*, for example, occurs more frequently than expected for a word with the meaning ‘at once’; *gan*, too, would appear to be more common than an ingressive aspect marker would normally be. Moreover, the frequency of *pa*’s in OE narrative discourse is very high.

(d) **Phonetic ‘shortness’ or reduction**: OE *pa* exists alongside a longer form *panne* or *ponne*, which occurs in ‘non-narrative’ contexts (Wårvik in press), while ME *gan* is an aphone form of OE *on-*/aginnan and occurs concurrently with the prefixed form *beginnen*, which is usually a pure ingressive marker.

(e) **Asyntactical, or occurrence outside the core syntactic structure**: OE *hwæt* is always appended loosely to a clause; the ME *I gesse* forms, in their pragmatic uses, occur parenthetically, either clause finally, initially, or medially; and ME *bifel*-constructions are also occasionally parenthetical.

(f) **Restriction to, or possibility in, sentence-initial position**: OE *hwæt* and *hwæt pa* almost always appear sentence initially, as do OE *gelamp-* and ME *bifel*-constructions. ME *I gesse* parentheticals may, and frequently do, occur sentence initially, and ME *anon*, while positionally quite free in Chaucer, seems to be restricted to initial and *anon* collocations in Malory. By definition, preposed *whan*-clauses occur sentence initially.

(g) **Lack of semantic or propositional content**: Many of the mystery features are traditionally described as ‘meaningless’, ‘empty’, or ‘colorless’. ME *anon* often seems bleached of any real sense of immediacy. ME *gan* cannot be translated ‘began’ because of the presence of incompatible aspeccular elements. OE *hwæt* has no interrogative sense, but rather is said to have ‘vague meaning’. ME *for the nones* is described by dictionaries as having ‘no special meaning’, and the *gelamp-* and *bifel*-constructions both contain extremely general verbs of happening.

(h) **Difficulty in translating**: OE *hwæt* has been variously—and unsatisfactorily—translated as ‘what ho’, ‘list’, ‘alas’, ‘lo’, ‘indeed’, and ‘hear (me)’.

(i) **Optionality**: The absence of any one of the forms studied renders the discourse neither ungrammatical nor unintelligible. Even the *gelamp-* and *bifel*-constructions, which would seem to be grammatically more central, can be omitted, with the original complement clause assuming main clause status.

(j) **Stylistic stigmatization (due to frequency, orality, and apparent lack of meaning)**: When seen as a metrical expedient, inserted by (incompetent) poets to add a syllable to a line of verse or to move the infinitive into rhyme position, *gan* is considered a defect of ME style. Similarly, *for the nones* is described as a ‘stop-gap’, ‘mere tag’, or ‘weak explicative’. Moreover, the *gelamp-* and *bifel*-constructions might be considered clumsy, unsophisticated, or naive.
Of the forms examined, OE *hwæt* would seem to be most similar to a pragmatic marker: it is a phonetically short item of indeterminate word class, with no obvious lexical or grammatical meaning, occurring optionally outside of the clausal structure in sentence-initial position. A number of OE and ME forms fail to meet some of these criteria: for example, ME *gan* never appears sentence initially, *anon* belongs to a clearly recognizable lexical category, and OE *gelamp-* and ME *bifel-* constructions, as well as not being phonetically 'short', constitute part of the core syntactic structure. But the same types of exceptions can be found in many of the Modern English pragmatic markers.

2.2 PRAGMATIC FUNCTIONS. Stronger evidence for the status of mystery particles as pragmatic markers is provided by their functions. Pragmatic markers in Modern English serve two broad purposes, which following Halliday (1970) can be termed 'textual' and 'interpersonal'. Together these can be classified as 'pragmatic', as opposed to propositional or referential (see Leech 1983). Generally speaking, in the textual mode, the speaker structures utterances as text, while in the interpersonal mode, the speaker expresses subjective attitudes and evaluations as well as acknowledge and maintains a social exchange with the hearer.

2.2.1 Textual functions. The mystery features of Old and Middle English serve two types of textual roles:

(a) Salient event marking: In Old English, *pa*, accompanied by verb second order, is commonly recognized as a foregrounder or peak marker, and in Middle English, the historical present is also seen as a foregrounder. It can be argued that both *gan* and *anon* have foregrounding functions as well (see 2a and 2b), but denote different levels of the foreground. *Gan* denotes 'pivotal' events which would constitute part of a high level summary of the text, while *anon* marks 'backbone' events which would be part of a more detailed summary. Furthermore, *gan* focuses on the action itself and slows the narrative down, while *anon* emphasizes the sequence of events and serves to speed up the pace of the narrative. In a somewhat similar way, *hwæt pa* in Old English can be seen as moving the narrative forward by indicating that the event which follows is to be inferred from the one that preceded (see 1c).

(b) Narrative segmentation: Episode or topic shift is a function attributed to *pa*, to aspectual forms and to word order in Old English and to *pan* 'then', to the perfect, and to the historical present in Middle English. More importantly, both the OE *gelamp-*construction and the ME *bifel-*construction have the primary purpose of demarcating the beginning, and sometimes the end, of a narrative unit, or episode (see 1a and 2c). The *that-*clause denotes the inciting or instigating event of the episode and the main clause verb of happening serves as a metacommentary on the narrative structure, orienting the listener to the structural shape of the narrative, while the accompanying adverbial is a temporal or causal frame which grounds the episode within the narrative. Both the framing and the orienting functions are assumed by preposed *whan-*clauses in later Middle English (see 2d).

Additionally, one may note that the multifunctionality that characterizes pragmatic markers in Modern English is also true of mystery features. For instance, OE *pa* has been identified as a foreground 'dramatizer', a sequencer of events, a marker of colloquial speech, and an indicator of narrative segmentation (Enkvist 1986).
2.2.2 Interpersonal functions. In the interpersonal component, the mystery features of Old and Middle English exhibit both more speaker-oriented (i.e. subjective) and more hearer-oriented (i.e. interactive) functions:

(a) Subjective functions: Many of the mystery particles serve as devices for internal evaluation, a function traditionally designated as 'emphatic' or 'intensive', whereby the speaker highlights the importance of various narrative situations without breaking the narrative frame. Since evaluation is logically aligned with foregrounding, it seems clear that both ME anon and gan can have this function. In Old English huwēt may serve the purposes of both internal and external evaluation. An evaluative dimension has also been attributed to the historical present and to the perfect in Middle English. The ME I gesse parentheticals have the subjective function of indicating the speaker's epistemic uncertainty (hedging) or less often certainty, while the speaker's sense of surprise or astonishment may be expressed by (eala) huwēt in Old English and by what in Middle English (as well as in Modern English) (see 2e).

(b) Interactive functions: The interactive functions of OE and ME mystery features are quite varied. OE huwēt can focus the attention of the hearer or reader (cf. ModE y'know what?); similarly, ME what (ho) claims the attention of an interlocutor (see 2f). The main function of OE huwēt, however, is to preface information which is, or is presumed to be, shared (see 1b). By confirming shared knowledge, OE huwēt may establish intimacy or solidarity between the speaker and hearer ('positive politeness') and is often intended to elicit the hearer's favorable reception of the information. The parenthetical ye knowen has an analogous function in Middle English (see 2g). ME God woot and trusteth me wel are attempts by speakers to persuade hearers of the truth of their utterances by invoking authority (see 2j and 2k). Finally, I gesse parentheticals are primarily a means of speaker self-effacement or deference serving the purpose of 'negative politeness' (see 2g, 2h, and 2i).

In the absence of an oral context, the interactive functions of the mystery features are particularly significant in suggesting their classification as pragmatic markers.

2.3 Grammaticalization. While studies of the development of pragmatic markers in Modern English have generally been concerned only with the relation of propositional to non-propositional meaning, a number of studies have considered their development as a process of grammaticalization, seen as a synchronic phenomenon. Most important here are Romaine and Lange (1991) on (be) like and Thompson and Mulac (1991) on I think. While recognizing a lack of historical depth in their study, Romaine and Lange suggest that for both the discourse marker use of like and the 'quotative' use of be like, the original semantic and syntactic properties lead to the pragmatic functions that develop: 'Semantically, it is because like has the referential meanings of 'comparison', 'for example', 'as if', and so on, that it is suitable for use in a construction reporting hypothetical discourse or thought. Syntactically, it is because it can occupy a slot immediately preceding the comparison ... that it can function as an anaphor whose scope is forward or backward' (1991:246). Using quantitative data, Thompson and Mulac (1991) argue that the development of 'epistemic parentheticals' I think and I guess, albeit not a 'textbook case', shows many of the features of grammaticalization, including, most
importantly, the fact that the forms are decategorialized from subject + verb to a ‘unitary epistemic morpheme’, that they can continue to function (in ungrammaticalized form) as regular complement-taking verbs, and that in their grammaticalized form they retain vestiges of their earlier meanings (cf. Hopper 1991).

The mystery features of Old and Middle English do seem to undergo many of the morphosyntactic and semantic changes identified with the process of grammaticalization, though never, of course, being fully ‘grammaticalized’ in the sense of being incorporated into a recognized grammatical paradigm (see Lehmann 1985:307, 309 on ‘paradigmatication’).

2.3.1 Morphosyntactic changes. In respect to morphosyntactic changes, it should be acknowledged from the outset that mystery features do not undergo either phonological reduction or morphological bonding, which are thought to be typical of grammaticalization. However, function words (auxiliaries, case markers, etc.) are generally considered clear instances of grammaticalization, even though they retain their lexical independence and often their full phonetic substance. It should also be acknowledged that the mystery features often begin the grammaticalization process as phrases or clauses rather than as individual words, but this too has not been seen as a barrier to grammaticalization (Heine et al. 1991:24-25; Thompson and Mulac 1991:318, 324).

Mystery features are subject to one of the central processes of grammaticalization, namely, decategorialization, or loss of the morphological and syntactic characteristics of their original category, with movement down the scale from more major to more minor word class membership (Heine et al. 1991:213; Hopper 1991:30-31; Hopper and Traugott 1993:103-13). OE ħwæt loses the characteristics of an interrogative pronoun, evolving into a particle of indeterminate status. ME gan undergoes a shift from a full complement-taking verb into a quasi-auxiliary found only in the third person preterite. As Thompson and Mulac (1991) argue for I think, the ME I gesse and ye knowen parentheticals develop into more or less unitary epistemic particles similar to adverbs, restricted to first-person subjects and present tense verb forms. ME bifie-l-constructions become more unified and particle-like in nature, as they are increasingly followed by that-less complements or are found parenthetically. Even gelamp-constructions are characterized by Gorrell (1895) as ‘introductory particles’.

To some extent, mystery features also undergo syntactic fixation, in which they lose syntactic variability and come to occupy a fixed slot (Lehmann 1985:308, 309). In its pragmatic functions, ħwæt (pa) is restricted to initial position, while ħwæt always precedes clauses with first- or second-person subjects. At least in later Middle English (Malory), anon becomes confined to initial position as well. ME gan, which may be followed by to or to-less infinitives in earlier Middle English, later becomes restricted to bare infinitives and may not be separated from its infinitive.

Comparing grammaticalized and non-grammaticalized forms, there is strong evidence that the mystery features experience ‘divergence’, ‘split’, or ‘form-meaning asymmetry’, that is, the retention of full lexical characteristics in some contexts alongside grammaticalization in other contexts (Heine et al. 1991:212-13; Hopper 1991:24-25; Hopper and Traugott 1993:116-20; Lehmann 1985:311). OE ħwæt continues to function as an interrogative pronoun/adjective/adverb (in the
propositional mode) while also serving as a complementizer (in the textual mode) and an interactive marker (in the interpersonal mode). ME *gan* remains as an ingressive aspectual marker and ME *anon* as a temporal adverbial while at the same time taking on pragmatic functions. ME suppositional verbs such as *gessen, witen, trown* may still be used as regular verbs with nominal and sentential complements and a full range of personal subjects, maintaining their 'full' meaning and denoting actual acts of cognition. Finally, the impersonal verbs of happening in both Old and Middle English continue to denote single, isolated events.

The phenomenon of 'layering', the continuation of older, more highly grammaticalized forms next to newer, less grammaticalized forms (Hopper 1991:22-24; Hopper and Traugott 1993:123-26), can also be observed among the OE and ME mystery features, in, for example, the overlapping use of *bifel*-constructions and proposed *whan*-clauses for denoting episode boundaries in Middle English and in the concurrent loss of 'you know' senses of *what* and rise of *ye known* parentheticals during the same period.

2.3.2 Semantic changes. The mystery features of Old and Middle English all exhibit a 'semantic aptitude' (Lehmann 1985:315), or appropriateness for the type of pragmatic marker that they become. It is the general interrogative sense of *hwet* that allows its development as a marker which questions common knowledge, expresses surprise, and focuses attention. It is the subjective epistemic/evidential sense of the *I gesse* parentheticals that permit their development as markers of negative politeness. It is the inceptive semantics of *gan* that motivates its development as a textual marker which focuses on the ensuing action and slows the pace of the narrative, while it is the perfective sense of *anon*—its portrayal of an event as occurring in an instant or as a whole—that motivates its development as a textual marker which emphasizes the sequence of events and speeds up the pace of the narrative. Finally, it is their very general verbal meaning of happening that makes the *gelamp* and *bifel* verbs suitable for episode boundary marking. Thus, like other items undergoing grammaticalization, it is the original lexical meaning of the mystery features which in large part determines the extended meanings or functions that develop.

In all cases, the semantic development of the mystery features provides evidence for referential (propositional) meaning being the source for pragmatic (textual and interpersonal) meanings. And such unidirectionality from propositional to pragmatic meaning has been seen as characteristic of grammaticalization, where there is a 'tendency to recruit lexical (propositional) material for purposes of creating text and indicating attitudes in discourse situations' (Traugott in press). 11

The semantic development of the mystery features of Old and Middle English can frequently be explained by the conventionalization of conversational implicature, what Traugott (in press) calls the 'semanticization' of implicature and Heine et al. (1991) call 'context induced reinterpretation', which, along with metaphor, has been seen as a central process in grammaticalization. For example, the textual meaning of saliency or sequentiality of ME *anon* is an implicature of the word's sense of suddenness or urgency: if something happens in great haste, this implicates that it follows quickly in succession and that it carries some importance. The interpersonal 'you know' sense of OE *hwet* is an implicature of the word's interrogative sense: that is, from a questioning of what the hearer knows is inferred an expression
of the speaker’s confirmation of what the hearer knows. The surprise sense of ME what is a result of the speaker’s request for reasons, such a request suggesting that she or he is puzzled, surprised, or impatient and hence needs reasons. The textual meaning of gan is also contextually implied, an implication of the lexical meaning of the verb that arises in the context of an event sequence: rather than denoting the beginning of an action occurring in isolation, it comes to denote the beginning of an action seen as part of a series of actions; the textual meanings of the gelamp- and bifel-constructions are similarly implied by the narrative context.

3.0 CONCLUSION
In the history of English, there appears to be a continuity in pragmatic functions, at the same time that the forms—which are highly ephemeral—are in a continual process of renewal; such replacement is also characteristic of grammaticalization (Hopper 1991). Thus, OE hwæt is comparable to ModE y’know or y’know what; OE hwæt þa is comparable to ModE so in its implicational sense; ME anon is comparable to ModE now; and ME gan is perhaps comparable to ModE colloquial forms such as up and, take and, and go and. Gelamp þæt in Old English is continued as it bifel that in Middle English, as it came to pass that in Early Modern English, and as it happened that in Modern English oral narrative. However, it is replaced, in written narrative, by preposed when-clauses beginning in Middle English. Only the I gesse parentheticals seem to have persisted unchanged throughout the subsequent history of English, with only minor changes in the inventory of suppositional verbs used.

In conclusion, it seems clear that, despite the lack of oral discourse, we can speak confidently about pragmatic markers in earlier periods of the language. Taxonomically, both the pragmatic markers of Modern English and the mystery features of Old and Middle English have mixed characteristics, though they are all generally marginal in word class membership. The various textual and interpersonal functions of the mystery features point even more decisively to their equivalency with modern pragmatic markers. Moreover, the diachronic development of mystery features seems analogous to the synchronic development of pragmatic markers, with a semantic appropriateness between original and derived form and with pragmatic meanings derivable from propositional meanings, often by means of conversational implicature. Finally, both the synchronic and diachronic developments are describable as processes of grammaticalization exhibiting, above all, syntactic fixation and decategorialization.

Appendix: Studies of the ‘mystery features’ of Old and Middle English

OLD ENGLISH
(and) þa ‘(and) then’ (Enkvist 1972, 1986, 1994; Enkvist and Wårvik 1987; Foster 1975; Taejin 1992; Turville-Petre 1974; Wårvik 1990a, 1990b, 1994)

hwæt (þa) ‘what (then)’ (Brinton 1990a, in press)

þa {gelamp, geweard, getimode, getidde, gesælde, wæs} þæt ‘then it happened that’ (Brinton 1993, in press)

her ‘here’ (Clemoes 1985; Fries 1993)
nu ‘now’ (Fries 1993)

sona, þærhite ‘immediately, at once’ (Wårvik 1994)
word order (Hopper 1979, 1992)
imperfective and ingressive aspectual forms (Richardson 1991a, 1994)

MIDDLE ENGLISH

pan ‘then’ (Wårvik in press)
this (Novelli 1957)
but (Donaldson 1981)
anon ‘at once’ (Brinton in press)
do (Wright 1989, 1991)
for the nones ‘for the occasion’ (Lumiansky 1951)
then (it) {bifel, fel, happed} that ‘then it happened that’ (Brinton in press)
I {gesse, traw, wene, woot, undertake, know} ‘I think’, it {seemeth, thinketh} me ‘it seems to me’ (Brinton in press; Robertson 1933)
God (it) woot ‘God knows’ (Brinton in press)
trusteth me wel ‘trust me well’ (Brinton in press)
gan/con ‘began’ (Aertsen 1991; Brinton 1990b, in press; Richardson 1991a)
as it is told, that y of told, etc. (Wittig 1978)
ye known, thou woost, etc. ‘you know’ (Brinton in press)
wel (Finell 1989)

preposed whan–clauses (Brinton in press)
the historical present (Benson 1961; Richardson 1991a, 1991b; Zimmerman 1973)

perfect aspect (Richardson 1991a; Zimmerman 1973)

Notes

1. There is a large and growing literature on pragmatic markers in Modern English (for a summary, see Brinton in press, Chapter 2). Book-length studies include Erman (1987), Goldberg (1980), Schourup (1985), Schiffrin (1987), and Stenström (1994).
2. ‘For over twenty years the study of discourse has been almost exclusively concerned with synchronic analysis and … since we can no longer resort to the excuse that discourse studies are young and immature, we might find it necessary very soon to turn our minds to diachronic studies of discourse as well’ (Calvo 1992:26).
3. For convenience, constructions such as (1a) will be referred to as ‘gelamp–constructions’, those such as (2c) as ‘bifel–constructions’, and those such as (2h and 2i) as ‘I gese parentheticals’.
4. The distinction between pivotal and backbone is made by Jones and Jones (1979).
5. The structure and function of gelamp–/bifel–constructions offer further evidence that foregrounding and backing are not a binary distinction. For instance, while the that–clause of these constructions depicts temporal events in iconic order, and hence presumably foregrounded, the clause itself is syntactically backgrounded and often atypical of foregrounding in respect to both verbal aspect and subject type (see further Brinton in press).
6. One diachronic study of pragmatic markers is Finell (1989) on well in English, and Abraham (1991) on doch, eben, halt, and ja in German. See also the comments of Traugott on well and right (1982:251, 252, 255) and on let’s and let alone (in press).
7. See Brinton (in press, Chapter 2) for a review of the literature on grammaticalization; see also Bybee et al. (1994), Heine et al. (1991), Hopper and Traugott (1993), and Traugott and Heine (1991).
8. The one exception is ME God woot, which is found in the reduced form Goddot(h).
9. The mystery features also do not seem to undergo condensation, or loss of scope, since in their pragmatic functions they relate not to individual words or clauses but to larger stretches of discourse.
10. An exception to the loss of syntactic variability is the increased positional freedom of *I gesse* parentheticals, which come to occur in a number of different sentential slots.
11. This development need not, as the mystery features amply illustrate, be a unilinear development from propositional through textual to interpersonal meaning, but the different pragmatic meaning may develop quite independently from different propositional meanings of the same word (see Brinton in press, on the development of *huwe* and *anon* especially). This view accords with Traugott's (in press) sense of their being 'correlated diachronic continua' of development.

References


Donaldson, E. Talbot. 1981. Adventures with the adversative conjunction in the General prologue to the *Canterbury tales*; or, What's before the *but*? So many people, longages and tongues: Philological essays in Scots and mediaeval English presented to Angus McIntosh, ed. by Michael Benskin and M.L. Samuels, 355-66. Edinburgh: Middle English Dialect Project.


Enkvist, Nils Erik and Brita Wärvik. 1987. Old English *pa*, temporal chains, and narrative structure: Papers from the 7th International Conference on Historical Linguistics, ed. by Anna Giacalone


Richardson, Peter. 1991b. Tense, discourse, style: The historical present in 'Sir Gawain and the Green Knight'. Neuphilologische Mitteilungen 92.343-49.


in press. The ambiguous adverbial/conjunctions *pa* and *ponne* in Middle English: A discourse-pragmatic study of *then* and *when* in early English saints' lives. Historical pragmatics, ed. by Andreas Jucker. Amsterdam and Philadelphia: John Benjamins.


How Recent Contact Erased Ancient Traces in the Gender Systems of the Oromo Dialects
Robbin Clamons
University of Minnesota

In one respect, the gender systems of western and central dialects of Oromo, a Lowland East Cushitic (LEC) language, are less like the gender systems of other Oromo dialects and also other LEC languages, and more like those of non-LEC Cushitic and Semitic languages. In yet another way, the gender systems of the westernmost dialects are less like those of other Oromo dialects and other Cushitic languages altogether, and more like those of Semitic and Omotic languages. The position of Oromo within the Afroasiatic family is identified in Figure 1.1

![Diagram of Afroasiatic language family]

Figure 1

I argue that the variations in the different dialects resulted from contact of the western Oromo communities with distant relatives, and that reconvergence was supported both by attitudes of cultural identity and agencies that fostered social integration. Within the Le Page and Tabouret-Keller (1985) model, spoken language projects a speaker's world view onto a social screen of two or more individuals. Social factors serve to focus a set of language norms within certain overlapping social groups, or blur them across other social groups. Both institutions fostering social integration and cultural attitudes of identity have been identified as focusing factors in other sociolinguistic contexts (Tabouret-Keller 1992).

In the eastern and the southern dialects of Oromo, the assignment of gender is based on phonological features as well as semantic and pragmatic factors. This is also true in other LEC languages. In the western dialects, on the other hand, only semantic and pragmatic information is relevant in gender assignment, as in non-LEC Cushitic, Semitic, and Omotic languages. Expressions with initial i feminine and initial k masculine alternants are characteristic of Cushitic
languages. In the eastern and southern Oromo dialects there is a robust $t / k$ paradigm, but the westernmost dialects have only $k$ initial forms.

I argue in this paper that both of these differences reflect simplifications in the western varieties, and further, that these changes occurred under pressure of language contact: the simplification in the assignment system because of contact of Tulama and Mecha speakers with speakers of languages having semantically transparent gender systems, and the generalization of initial $k$ in the $t / k$ paradigm because of further contact of Mecha speakers with speakers of Semitic and Omotic languages, which do not have $t / k$ paradigms. Because Oromo became an important language for large groups of native speakers of non-LEC Cushitic and Semitic languages who were assimilated into the center of the Oromo community, the opaque phonological basis for gender assignment was replaced with a transparent semantic basis. Further assimilation of Omotic speakers finally led to the neutralization of the $t / k$ paradigm in the westernmost dialects.

1. Northwestern Migration of the Oromo People. In the 16th century, ancestors of the modern Tulama and Mecha, nomadic Oromo, embarked on a migration that took them to the north and west of their homeland in the southeastern highlands of Ethiopia (Lewis 1966, Hassen 1990). This excursion carried them through areas that had remnant populations of Highland and Central Cushitic, Semitic and Omotic peoples. They moved northward into the area of modern Shoa, which was at the time dominated by Amhara. The Mecha separated from the Tulama and migrated to the west and southwest. Figure 2 shows the relative geographic location of the modern Oromo dialects, and neighboring speech communities.

![Figure 2](image-url)
Hassen (1990:18ff) argues that two social factors facilitated this migration. The first is the political situation in the area at that time, and the second is the store of cultural attitudes and customs of the Oromo that fostered assimilation of outsiders into their group. The jihad of Imam Ahad (1529-1543) and the subsequent retaliation of the Christian emperor Galadewos (1540-1559) disrupted the political and economic situation to the north and west of those lands occupied by these Oromo nomads. These areas were ravaged by war, famine and heavy slave raiding that financed both factions in the conflict. Much of the population was destroyed or dispersed, both among the more sedentary agriculturally based Oromo and also among the other ethnic communities of the area. Before order could be restored by either Christian or Moslem factions, the nomadic Oromo community had pressed northward and established dominance in the area. The Tulama settled permanently in the region centered around modern Finfine (Addis Ababa) while the Mecha moved further to the west.

Oromo attitudes of cultural identity and institutions of group and individual adoption supported the absorption of peoples from the diverse language groups encountered by the Tulama and Mecha. Among the Oromo there were two forms of adoption, guddifacha, which was the adoption of an individual by a foster parent, and moggaasa, which was the adoption of either individuals or groups into a gaassa "clan". Both of these forms of adoption were formally recognized and ritualized, and also binding and unbreakable. Although it would be exceptional, a man might adopt an eldest son if his own children weren't powerful enough to satisfy him even among modern western Oromos (Bartels 1983:136). Adoption of outsiders by Oromos was economically and politically beneficial to both parties, offering protection and material support to the adopted individuals and groups while at the same time enlarging and strengthening the Oromo clans. Hassen (1990:21ff) argues that this practice promoted the assimilation of other groups by the Oromo, rather than the reverse.

Modern Mecha Oromo distinguish Boraana, who are supposed to be directly descended from the original migrants, and gabaro, who have been assimilated from other ethnic groups. The Boraana are considered to be the channel through which god blesses the people, and they therefore enjoy certain social privileges and fulfill certain social obligations for the community. In fact though, the distinction between gabaro and Boraana is to some extent a fiction. A man from Wollega explains that "... [in] our country the purity of descent in even the highest Boraana lineages is questionable. My own grandfather, who was elected my people's ritual leader of their gada, was a son of a Mao [Omotic] mother ..." (Bartels 1983:160).²

Oromo social practices are frequently organized on principles of community, rather than through descent and kinship groups (Bartels 1976, Lewis 1974, 1984). Neighbors and fellow members of formal and informal voluntary associations are often more important than kin in organizing social processes in Oromo culture. Work, conflict resolution, and preparations for funerals and births are typically organized by cooperating neighbors. Members of other ethnic groups may even be included in voluntary associations that are formed to take care of these kinds of social processes. In fact, this reliance on voluntary associations among neighbors is widespread among Cushitic people of this area, and it would have been a familiar institution for those who were assimilated into the Oromo clans.
The adoption customs and the community based organization processes of the Oromo brought conquered groups into the center of Oromo communal life, and individuals from other speech communities had to have shifted to the language of the Oromo people, in addition to sharing other cultural customs with them. Language communities that were in contact with the Tulama and Mecha communities over a considerable span of time include the Amhara, whose language, Amharic, is an Ethiopian Semitic language, and the Hadisyya, who speak a Highland East Cushitic language. In their migrations to the west, the Mecha also came into contact with speakers of Omotic languages. There is widespread multilingualism in western areas among speakers of Omotic languages to this day. Cooper, Singh and Ghermazion (1976) report that over half of the native speakers of Omotic languages in their surveys in western regions also speak Oromo.

2. Gender Assignment in Oromo. The gender assignment systems differ across the range of Oromo dialects. In dialects spoken to the east and south of Shoa, assignment has a phonological basis. Thus in these dialects, sareé 'dog' is feminine in unmarked contexts, as in 1a, because of the final non-low high tone vowel. The gender assignment system in Tulama and other western dialects is semantically transparent. Expressions denoting females and diminutives or pejoratives are feminine, and expressions denoting males and all other non-humans are masculine. In these dialects, sareé is masculine, as in 1b, unless the dog referred to is is to be specifically identified as female, or is held in particular contempt or regard by the speaker.

(1) a. sareé bareed - duu
   dog F beautiful F
   'beautiful dog'
   'beautiful female (/ dear or nasty) dog'
   (eastern / southern dialects)

b. sareé bareed - aa
   dog M beautiful M
   'beautiful dog'
   'beautiful male (/ dear or nasty) dog'
   (western dialects)

In all of the dialects of Oromo, an affective meaning is associated with a shift from the grammatically determined gender. Speaker B from the Harar dialect scorns the dog under discussion in the brief discourse in 2 (Clamons 1993:277).

(2) A: sareé takka ganda xeesa arkinne.
   dog F one F village in we saw
   'We saw a dog in the neighborhood.'

B: sareé-n xun bashoo tiyya jala fige.
   dog M SUB TOP that M cat F my F after ran M
   'That (grrrr) dog chased my cat.'

She expresses this by shifting the referring expression from the grammatically prescribed feminine gender into the masculine gender. Because of the semantic assignment basis in western dialects, only masculine expressions are eligible for gender shifting.
### 2.1 Gender Assignment in Other Afroasiatic Languages

Throughout the LEC branch, assignment of gender is based on phonological and semantic factors. In Somali, Rendille and Qafar, masculine and feminine nouns generally have distinctive accentual patterns. In all three languages there are pairs of polysyllabic feminine nouns having a final high tone accent, and polysyllabic masculine nouns having a final low tone, as illustrated in the pairs in 3 from Qafar (Hayward and Corbett 1988: 274.277), in 4 from Rendille (Oomen 1981:43), and in 5 from Somali (Kraska 1992:16).

<table>
<thead>
<tr>
<th></th>
<th>a. awká</th>
<th>b. áwka</th>
<th>c. gabá</th>
<th>d. i̠ba</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>'girl'</td>
<td>'boy'</td>
<td>'hand F'</td>
<td>'foot M'</td>
</tr>
<tr>
<td>4</td>
<td>a. in̠m</td>
<td>b. in̠m</td>
<td>c. ar̠m</td>
<td>d. ar̠m</td>
</tr>
<tr>
<td></td>
<td>'girl'</td>
<td>'boy'</td>
<td>'wife'</td>
<td>'husband'</td>
</tr>
<tr>
<td>5</td>
<td>a. in̠n</td>
<td>b. in̠n</td>
<td>c. org̠</td>
<td>d. org̠</td>
</tr>
<tr>
<td></td>
<td>'girl'</td>
<td>'boy'</td>
<td>'she-goat'</td>
<td>'he-goat'</td>
</tr>
</tbody>
</table>

In Dasenech, although there may be some relationship between the gender of an expression and a referent's size, importance or strength, gender and phonological form also correlate with gender to a certain extent. Consonant final nouns are always masculine, and nouns ending in final ti are always feminine. Oomen (1981) and Kraska (1992) have both argued that feminine proto forms *Vt / *tV have eroded from the ends of feminine nouns that have the characteristic final high tone pattern in modern Rendille and Somali. This analysis could be extended to account for the patterns both in Dasenech and in eastern and southern Oromo dialects.

Gender assignment is described quite differently for non-LEC languages within Afroasiatic. In Hadiyya, for example, the semantic basis is entirely transparent: only expressions referring to females are feminine. Sasse (1984:117) claims that "...in addition to the semantic category of natural sex, which is of minor importance in the Cushitic, gender categories primarily denote the semantic notions of social significance (masculine) vs. social insignificance (feminine)". Agaw languages have no lexical or phonological bases; gender can only be assigned "...within a structured piece of speech" (Appleyard 1984: 582).

Semitic languages also have two gender systems, with gender distinctions found in pronouns, nouns and verbs. Amharic, for example, has a two gender classification system. As in Agaw, the assignment of gender is based on semantic or pragmatic criteria. Cowley et al (1976: 84) write that the "... so-called masculine forms are really unspecified as to gender, while the feminine forms specifically refer to something female, relatively small, or toward which the speaker feels affection." This affective use of feminine gender is illustrated in the Amharic expressions in 6.

<table>
<thead>
<tr>
<th></th>
<th>a. yih mes'haf</th>
<th>b. yicc mes'haf</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>this M book</td>
<td>this F book</td>
</tr>
</tbody>
</table>
|   | 'this book'    | 'this (favorite/ little) book(let)'

In Hamer, an Omotic language spoken in southwestern Ethiopia, feminine gender is associated with the values that are usually associated with masculine gender in other Afroasiatic languages. Nouns denoting commonplace, large dependent or grouped objects are marked with the feminine -tono and -no
suffixes. Thus in addition to use of feminine expressions for reference to females, such as *ratcono* 'Rac woman', a noun used to refer to large referents such as *ammo* 'large field', is feminine, while a noun used to refer to small referents such as *amá* 'small field' is masculine, as in (7) (Lydall 1976:407).

\[
\begin{array}{llll}
\text{(7) a.} & \text{rac} & '\text{Rac (clan)}' & \text{b.} & \text{ami} & '\text{field}' \\
& \text{ratca} & '\text{Rac man}' & & \text{amá} & '\text{small field}' \\
& \text{ratcono} & '\text{Rac woman}' & & \text{ammo} & '\text{large field}' \\
& \text{ranco} & '\text{all Rac}' & & \text{amino} & '\text{all fields}' \\
& \text{ranca} & '\text{several Rac}' & & \text{amna} & '\text{several fields}' \\
\end{array}
\]

2.2. Reconvergence Hypothesis. Gender assignment in all of the non-LEC Afroasiatic languages is quite different from that of the eastern and southern dialects of Oromo and other LEC languages, in which a certain gender is associated with a noun. In these other languages nouns have no formal, underlying gender, but are assigned gender in each context on the basis of pragmatic context and cultural values. Although masculine is described as the explicit or implicit default gender, these languages do not have gender fixed in the lexicon, and it is the referent of an expression that is associated with certain values if the expression is masculine, others if it is feminine. Notice that this is equivalent to the gender assignment function in modern western Oromo varieties.

The difference in gender assignment in the western and eastern Oromo communities could have developed in two ways. A system originally based on semantic factors, like the western systems, could have developed into a more opaque system because of regular historical changes, analogy, etc., or a system based on a complex of semantic, lexical and phonological factors could have been simplified to a semantically transparent system, which I contend is what happened in the history of Oromo. Comparative linguistic evidence in support of this analysis is found in the data from other LEC languages (Qabar, Rendille, Somali, and Dasenech) which also assign gender on the basis of phonological form. If it were the case that an originally semantically transparent gender assignment basis had been preserved in the western Oromo varieties, then formal factors relevant to gender assignment would represent innovations, not only in the eastern and southern varieties of Oromo, but also in these other LEC languages. Since nouns with non-low final vowels are all feminine and also all have a final high tone accent in a number of the modern languages, it is likely that the assignment of gender was based on phonological form in Proto-LEC, and that the western dialects of Oromo have a simplified gender assignment system.

In non-LEC branches of Cushitic, and in Semitic and Omotic languages, gender is assigned on the basis of semantic value or pragmatic context, without reference to phonological form. Cultural and social institutions that operated to foster absorption of other speakers from these other groups into the migrating western Oromo communities introduced significant numbers of speakers of these more distantly related languages into the center of Oromo life. Since the Oromo community wanted to promote the inclusion of neighboring individuals and groups, they would have been motivated to accommodate innovations in the language. It seems likely that speakers of these other languages who shifted to Oromo would have carried over their semantically transparent system of gender assignment. Recognition of this new assignment function as the norm within these western groups would create a focus within the regional community,
balancing the blurring of gender assignment functions across the larger Oromo group. Examples of the generalization of gender assignment to a semantically transparent system are found in similar contact situations.

2.3. Convergence in Other Contact Situations. Gumperz and Wilson (1971) consider the simplification of the gender assignment function in Kupwar varieties of Urdu and Marathi. Kupwar village is a stable multilingual society in which Kannada and Marathi have been in intimate contact for six centuries, with Urdu introduced three or four centuries ago. Because of the stable social separation of speakers of the three languages, all three have been maintained, but many of the syntactic structures have converged in the Kupwar varieties. One convergence is found in the gender assignment systems.

In Urdu there are two genders, with most referring expressions for human females and males assigned feminine and masculine gender, respectively, and with nouns denoting inanimates assigned predictably to either gender on a phonological basis. Masculine is the default gender, used when gender is unspecified and when agreement is blocked by syntactic rules. In Marathi, there is a three gender system with a similar assignment function, except that neuter is the default gender. Kannada also has a three gender system, but assignment is entirely semantically based, expressions denoting female humans are feminine, those denoting male humans are masculine, and all others are neuter.

In the Kupwar variety of Kannada, the assignment function is the same as in other varieties of Kannada. But the Kupwar varieties of Marathi and Urdu now have semantically based assignment systems similar to the Kannada system. Kupwar speakers now classify referring expressions as feminine only if the referent is human and female. In Kupwar Urdu all other forms are masculine. In Kupwar Marathi, expressions referring to human males are masculine, all others neuter. The assignment function in these varieties has collapsed to a semantically based system just as in the western dialects of Oromo. This is illustrated in the data in 5 that show the difference in gender choice in Urdu and Kupwar Urdu (Gumperz and Wilson 1971: 156).

(8) a. wəhā  nədii  a-iː.  
   there  river  came F
   'There was a flood.'  
   
   b. hwe  nədi  ay-ːa.  
   there  river  came M
   'There was a flood.'  
   
   c. yəllii  hwəLii  bət-tu.  
   there  river  came N
   'There was a flood.'

In the Kupwar Urdu, the noun nədī 'river' is assigned masculine gender, although it would be assigned feminine gender in High Urdu.

Altogether, Gumperz and Wilson consider twenty-three modifications in the three languages. The modification in the gender assignment systems of Urdu and Marathi have the greatest number of changes, and represent the only instance where two languages have changed to become more like one. All of the changes that they observe are reductions or generalizations. The change in the gender systems of Urdu and Marathi to semantically predictable systems also corresponds
to the change in the Oromo assignment in the western varieties. Weinreich (1953: 42-43) discusses a number of contact situations, similar to the contact situation that western Oromos have been in recent centuries, in which grammatical accommodation results in the simplification of grammatical categories. He observes that the assignment of gender tends to be simplified to a semantically transparent basis, similar to that described in Gumperz and Wilson, and that I propose here for the western Oromo dialects. The acceptance of new speakers into the heart of the community provides a new focus that promotes the establishment focused language conventions within the new group.

3. The t / k Paradigm. In eastern, southern, and central dialects of Oromo there is a paradigm that has alternating t feminine and k masculine forms of demonstratives, possessive pronouns, relatives, and interrogatives. In the question in 9a from the eastern Arssi dialect, for instance, 'which' and 'your' are tamtu and tee, in agreement with haadha 'mother', but in the question in 2b, they are kamtu, and kee, in agreement with abbaa 'father'.

(9) a. tamtu haadha tee?
   which F mother your F
   'Which is your mother?'

b. kamtu abbaa kee?
   which M father your M
   'Which is your father?'

This t / k alternation has been neutralized in western dialects throughout the paradigm. This is illustrated in the example in 10, from the western Wollega dialect.

(10) a. kamtu haadha kee?
    which mother your
    'Which is your mother?'

b. kamtu abbaa kee?
    which father your
    'Which is your father?'

The masculine k initial form has been generalized, and 'which' is appropriately kamtu in both masculine and feminine expressions.

3.1. Simplification of the t / k Paradigm. Bryan (1959) is a survey of languages which have a paradigm with a t / k initial alternation. Some are Cushitic and others not in the Cushitic subfamily. She hypothesizes that these forms derive from a substratum. Later, Bryan retracted this hypothesis (Bynon and Bynon 1975: 358), but Hetzron (1980: 20) has suggested that perhaps Cushitic formed the substratum of other t / k languages. He points out that an alternation in Afroasiatic determiners may be posited between feminine *t and masculine *kW. Whatever the distant history, the paradigm was clearly not an innovation in Oromo.

While Cushitic languages have characteristically robust t / k paradigms, other Afroasiatic languages do not. Omotic languages do have two gender systems, with marking typically found on nouns and adjectives, personal pronouns and third person singular verbs. Assignment is semantically based, as it is in western varieties of Oromo. But although there are some reflexes of an earlier t / k system, notably an -o / -e alternation in nouns of some Omotic languages, there are no overt alternants in the modern languages.

Because there is an opposition of two markers, rather than an opposition involving one marker and an unmarked form, this area of the gender marking
system is more susceptible to collapse. The absence of an analogous opposition in Omotic languages, however, provides external pressure, because of the incidence of bilingualism that can be assumed, given the history of migration and assimilation of Omotic people by the Oromo in the western highlands, as well as the extensive present day bilingualism within Omotic communities in western areas.

3.2. Direction of Generalization. It is sometimes assumed that masculine gender will always be generalized. Hetzron (1980: 13), writes for example: "...that when only one gender survives in a language, it is the masculine that is likely to be generalized, not the feminine." Lehiste (1988) points out that although interference in the form of underdifferentiation in the gender assignment system of a second language is expected in second language speakers whose first language has no equivalent assignment function, the direction of generalization may well be unpredictable. It is the case that the generic gender in some languages is the feminine. Alpher (1987) identifies some languages in which feminine gender is the generic form. Bani (1987) and Vormann and Sharfenberger (1914) both describe languages with default feminine gender. In American Norwegian, new nouns are always classified as masculine, in Australian German, however, feminine is the default gender.

Alpher (1987) does make some predictions about the direction of generalization of gender which are consistent with the situation in the Oromo data. He claims that in a patriarchal society, a tendency toward generalization of the masculine form is expected, that in matriarchies, or societies where women are integrated together in a group with young and old males in the social pattern, feminine may well be the default gender. Baxter (1978), Haberland (1963), Knutsson (1976), Legese (1963) and Lewis (1963) all include descriptions of Oromo cultural organization, characterizing it as patriarchal. The position of women is described as marginal, although it is clearly little understood. In fact it is the masculine form that is generalized in all of the varieties of Oromo, except where feminine singular morphology is equivalent to collective marking, which is found throughout Afroasiatic. The masculine form is general in the plural relative and in the plural anaphoric in eastern varieties. In the western varieties the masculine form is always the default form, and it has been generalized in all the $t/k$ forms, and in the assignment function as well.

4. Conclusion. There are two aspects of the gender systems that differ across the dialects of Oromo, the assignment system and the marking system. The assignment of feminine or masculine gender is based on phonological, semantic, and pragmatic information in the eastern and southern varieties, but on only semantic and pragmatic information in the varieties spoken in and to the west of Shoa. Further, the eastern and southern dialects of Oromo have a distinctive $t/k$ paradigm that is also found in the other Cushitic languages, but neutralized to an all $k$ system in western dialects. I have argued that the western varieties have been simplified.

Western Oromos have been in intimate contact with speakers of languages that assign gender on semantic grounds, with feminine the gender for diminutives and masculine the default gender, and western communities have assimilated large numbers of these speakers into the Oromo culture. This change is similar to the change to a semantically transparent system that is attested in unrelated languages
in similar contact situations. As second language speakers of Oromo redefined the assignment function, native Oromo speakers would have adopted the new conventions in order to strengthen group identity. Western Oromos would be less likely to focus on a common pattern with eastern and southern speakers as the geographic and social distances grew greater between the groups. As speakers focused within the migrant community, by establishing a new norm, the shared language features of the larger Oromo speech community blurred.

In Tulama and all dialects spoken to the east of Shoa, the \( t / k \) paradigm has been maintained. In the dialects west of Shoa, where Oromos have both assimilated speakers and maintained contact with unassimilated speakers from Omotic and Semitic speech communities, the variants are generalized to \( k \), the Cushitic masculine form. I have suggested that the \( t / k \) marking collapsed in the westernmost varieties of Oromo under the pressure of contact with languages that have no analogous pattern. Again, focus of language norms within the Mecha community promoted diffusion of language conventions across the Oromo dialects.

Although there is considerable contact between Oromos in eastern and southern communities and other ethnic groups, the contact is typically with speakers of other LEC languages, or involves contact with Arabic. This important contact with Arabic is diglossic, with native Oromo speakers typically using Arabic as the language of religion and instruction, or in trade, and there are clear influences on eastern dialects, notably in the addition of voiceless velar fricatives initially, and in the introduction of large numbers of Arabic loans. The western groups, however, have absorbed large numbers of speakers of other Cushitic, Semitic and Omotic languages, thus the substrate pressure on the gender systems in these western dialects has been sufficient to facilitate modifications in the grammar.

Notes

1 My thanks to Amal Osman for continuing help with Oromo, and to Fahmi Katabay, Stephanos Madda, Abraham Oluma and other members of the Oromo community of the Twin Cities for information about Oromo and the Oromo people. Thanks to Larry Hutchinson for much useful advice since the earliest research study leading to this paper, and thanks for helpful suggestions and discussion to Bruce Downing, Tim Dunigan, Ann Mulkern, Bernd Heine, Nancy Stenson, and Jerry Sanders. I can in good conscience lay claim to all errors in the data and the analysis.

2 The \( gada \) system is a generation grade set that formed the foundation of traditional Oromo government. See Legesse Asmarom (1973) for a description of a modern \( gada \) institution among the Boraana of southeastern Ethiopia and northern Kenya.

3 Although as few as eleven percent of speakers in some rural Oromo communities speak Amharic in modern Ethiopia, a very high percentage of native Oromo speakers in towns are bilingual in Amharic (Cooper, Singh and Ghjermazion 1976).

4 There is no affective meaning associated with feminine gender in Hadiyya as in other Afroasiantic languages (Bender et al 1976).

5 Appleyard (1984), Castellino (1976) and Hetzron (1976) do indicate some correlation between \# a and feminine in some of the Agaw languages but this appears to be overridden by pragmatic evaluation.

6 There is also an interaction with plural and for inanimates, these feminine endings also denote a 'global plural' or 'group singular', in contrast to a discrete 'particular plural', that specifies
several particular single cases as opposed to all cases, and is formed with the suffix -na. This is illustrated in the examples given in 7.

7 Heine (1980) indicates that the Burji (Highland East Cushitic) of Kenya are assimilating to Oromo, but the long term bilingualism of Burji speakers across generations has not had a direct impact on the southern Oromo in the gender system.

8 Gragg (1980) examines some of the complexities encountered in attempting to sort out the origins of lexical items that are shared across languages in this area, but it is clear that eastern Oromos have adopted relatively larger numbers of Arabic words.

REFERENCES


Grammaticalization in AAVE

Patricia Cukor-Avila
University of North Texas

Guy Bailey
University of Memphis

1.0 Introduction

One of the primary goals of quantitative sociolinguistics has been to use the present to explain the past. In order to do this, sociolinguists have had to develop a set of analytical procedures that allow them to observe linguistic change as it is actually taking place. This ‘synchronic approach’ has been used quite frequently in the study of sound change (cf. Labov 1994), but it has been used much less often in the study of grammatical developments. Nevertheless, we suggest that this synchronic approach to grammatical change seems to be particularly appropriate for the study of grammaticalization (cf. Schwenter and Traugott 1994).

Recent developments in African American Vernacular English (AAVE) provide an ideal research site for exploring grammatical change in progress. As research on AAVE has broadened to include the speech of older rural African Americans as well as that of younger urban ones, sociolinguists have come to realize that several unique grammatical features of AAVE are not relics of an earlier creole but have emerged over the last fifty years or so. As a way of studying grammatical change in progress, this paper examines two of those features: the use of be+V+ing for habitual meaning, what we call be₂, and the use of had + simple past, what we call innovative had + past, which is used where other varieties of English use the simple past. The data for this examination come from two sources: an ethnolinguistic study of four generations of African Americans in the rural east-central Texas community of Springville (see Cukor-Avila 1995) and mechanically recorded interviews with former slaves made primarily in the 1930s and 1940s (see Bailey, Maynor, and Cukor-Avila 1991).¹ Taken together, these sources give us linguistic evidence from real and apparent time that spans 130 years.²

2.0 The Grammaticalization of be₂

The use of an uninflected be (as in They be wavin’, you know, doin’ their hands and stuff like that) is among the most widely recognized characteristics of AAVE. Early work on AAVE noted that this uninflected be (often called invariant be or be₂) was used to mark habitual aspect (see Fasold 1972). Because be₂ was so different from anything in white vernaculars and because Caribbean creoles often mark habitual aspect, some linguists argued that be₂ was most likely a relexification of an earlier creole form (see Dillard 1972). Much of the early work on AAVE, however, was based largely on the speech of children and teenagers, and with one exception (Wolfram 1974), all of it was done in Northern urban areas -- areas where large African American populations had developed only since World War II. In other words, in the early stages of work on AAVE, inferences about the historical development of that variety were made on the basis of the speech of children in areas that historically had not had large African American populations.

As researchers began to explore the AAVE used by older adults and rural residents, data which challenged the hypothesis that be₂ was a relexification of a creole feature began to emerge. Bailey and Maynor (1987, 1989), for example, found that be₂ occurred less frequently among older adults than among children,
adolescents, and younger adults and that it had a strikingly different syntactic and semantic distribution. Among elderly adults in urban areas and rural residents in general, \textit{be}$_2$ appeared before all types of predicates (that is, before V+\textit{ing}, locatives, adjectives, and NPs) with roughly equal frequency and without semantic restriction. Examples (1) through (8), taken from interviews with a male born in 1917, illustrate the wide syntactic and semantic distribution of \textit{be}$_2$:

(1) That \textit{be} a row here and that’s a row there.
(2) Tha’s a piece of land over there where \textit{be} a turn row betwixt it.
(3) They [chicken snakes] just \textit{be} knotted up when they suck eggs.
(4) And May used to be the wet part of the year; it don’t \textit{be} now.
(5) Well, it don’t \textit{be} too many [thunderstorms] right around here.
(6) We hear tell in different places \textit{be}, \textit{be} storms.
(7) [If you] \textit{be} sick and they wash your clothes they still want to pay for it.
(8) He \textit{be} full [right now].

In examples (1) and (2), for instance, \textit{be} occurs before an NP; in (7) and (8) it occurs before adjectives. In (3) and (7) it is used for habitual actions; in (1) and (8) it is used for actions at one point in time. These uses of \textit{be} are quite different from the uses reported for children, adolescents, and young adults. Among these groups, most of the instances of \textit{be} have habitual meaning, and most of them occur before V+\textit{ing}, as examples (9) and (10) illustrate:

(9) Sometimes them big boys \textit{be throwin’} the ball.
(10) They \textit{be doin’} the breakin’ durin’ PE and durin’ class time.

What these differences between earlier and more recent varieties of AAVE suggest is that \textit{be}$_2$ has become increasingly restricted syntactically to a position before V+\textit{ing} and semantically to a function as an habitual marker. Like contemporary AAVE, earlier AAVE had an invariant \textit{be}, but in earlier AAVE it was simply an alternate form used in place of \textit{am}, \textit{is}, and \textit{Ø} without any syntactic or semantic restrictions (\textit{are} was extremely rare in earlier AAVE and is still uncommon). Over the last half century, this invariant \textit{be} has become restricted to ‘progressives’ that have durative or habitual meaning. Figure 1 illustrates this development. Among the oldest people in the corpus (those born before 1945), \textit{be}$_2$ accounts for only 6% of the progressive tokens with durative/habitual meaning; among adults born between 1950 and 1965, it accounts for half of the tokens. Among children and adolescents, \textit{be}$_2$ comprises more than three quarters of the progressive tokens with durative habitual meaning; further, more than three quarters of the tokens of \textit{be}$_2$ occur before V+\textit{ing}.

The motive for the grammaticalization of \textit{be}$_2$ seems to lie in the messiness of the English progressive system and the lack of a clear relationship between form and function for present tense forms of \textit{be} in early AAVE. Comrie (1976) notes that English progressives can be used for a number of meanings other than actions in progress. For instance, they can be used with future meaning (I’m \textit{leaving Monday}), with habitual meaning (I’m \textit{jogging every day this spring}), and for continuous actions (I’m \textit{getting slower every day}). Early AAVE shows no tendency to use one form of \textit{to be} as opposed to another for these different meanings. In fact, in early AAVE there is no one-to-one relationship between form and function for any of the forms of \textit{to be} except for \textit{am} (which is almost always used for first person singular). \textit{Is} and \textit{Ø} alternate in both the third singular and
plurals (with their use constrained in part by whether or not the subject is an NP or a pronoun [see Bailey, Maynor, and Cukor-Avila 1989] and in part by the following predicate [see Rickford 1992]), be alternates with all three of the other forms. In early AAVE, then, the progressive signaled a wide range of meanings, as it does in other varieties of English, and except for am, the forms of to be used in the progressive had no unique function associated with them.

Figure 1. be + V+ing as a Percentage of all Progressives with Durative/Habitual Meaning (after Bailey 1993)

What has happened over the last half century is that this mismatch between form and function has in part been resolved as be has been grammaticalized to signal progressives (i.e., to positions before V+ing) that have habitual and durative meanings. Among young adults and children, then be_2 is used for habitual/durative actions, is, Ø, and am for true progressives. Although is and Ø are used in both the singular and plural of the third person (unlike am, which is restricted to first person), these forms are constrained by both the preceding subject and the following predicate. Along with the development of gonna as a future, the grammaticalization of be_2 creates a more optimal relationship between form and function for both the progressive subsystem and the present tense of be in AAVE.

One question that arises here, however, is why be_2 was the form reanalyzed to represent durative/habitual actions. The trigger for this reanalysis probably lies with a second kind of invariant be that occurred in earlier AAVE. Earlier AAVE included widespread deletion of would, and when would deletion occurred before be, it left an invariant be as in When I was a girl, we be picking cotton every fall. Because would is one of two ways of marking past habitual action in English (used to, of course, is the other), the deletion of would before be left an invariant be that marked past habitual action, as in the example above. This invariant be that derives from would deletion is thus quite similar to be_2 in contemporary urban AAVE.
except that the former refers to past actions. It is an easy step, of course, for the earlier invariant be that results from would deletion to trigger the use be₂ for present habitual actions. Bailey’s 1993 examination of the development of be₂ over time suggests three stages in the development of this habitual be. In the first stage, the stage represented by most of the older adults in the corpus, be₂ is simply an alternative form of the amis/∅ with no specialized function. During this stage it coexists with a second kind of invariant be derived from would deletion. In the second stage, be₂ takes on habitual meaning perhaps by analogy with invariant be from would deletion. Examples (11) through (14), which come from an interview with a female resident of Bryan, Texas, born in 1937, illustrate this stage. In stage two, however, be₂ still occurs before a variety of predicates. In the third stage, be₂ becomes restricted syntactically to positions before V+ing, as in examples (15) through (21) below which come from a woman born in 1945.³

(11) Really, you be more partial to them [grandchildren] than to your own.
(12) I found in fast food restaurants people be dirty sometimes.
(13) Some of the girls wear them [boots] and they be turned down . . . high-heeled boots, they be turned flat.
(14) FW: What causes those allergies?
 INF: Well, all the growth and everything you be around.
(15) FW: I always get up at 6:30.
 INF: So Randy be getting in the bed [when] you be get up.
(16) [When] we was working at night, we be watching a cute little guy come in.
(17) . . . ‘cause we be going to bingo [every week].
(18) She be sitting up there [at work] and she be kerplunk.
(19) . . . some went to $23 a month that she be getting here.
(20) [They] be fighting like R. and P.
(21) I be doing those doctors [cleaning their offices].

Thus invariant be, which was once a lexical variant used anywhere am, is, and ∅ could be used, has become grammaticalized to serve as a marker of durative/habitual meaning in the progressive system. As it has become grammaticalized, its syntactic distribution has become increasingly restricted. The context for this grammaticalization was the messiness of the English progressive system and the lack of a clear relationship between form and function in the present tense of be in early AAVE. The immediate trigger for the grammaticalization was the presence of a second kind of invariant be, one that results from would deletion and that signals past habituality. The end product is a kind of sorting out of the relationship between form and function.

3.0 The Grammaticalization of had + past

The use of innovative had + past in AAVE represents a similar sorting out of the relationships between form and function.⁴ Traditionally in English the past perfect has been used to express a situation that occurred prior to another past action, a type of ‘past in the past’ (for example, Betty had already served dinner when we arrived at the party). The relatively infrequent use of this form in speech is due to the fact that in the majority of instances events expressed in the past perfect can also be expressed using the simple past, except when a speaker wishes to convey a specific
chronological ordering of events. Our data show that over the last half century there has been a reanalysis of both the form and the function of the past perfect in the speech of our informants. The reanalysis of the past perfect is coupled with the increase in the use of this form, as is shown generationally in Figures 2 and 3, this being a direct result of its expansion from backgrounded to foregrounded discourse contexts (Hopper and Thompson 1980).

Earlier descriptions of AAVE by Labov et al. (1968) and Labov (1972) and Wolfram and Fasold (1974) note that the past perfect is used frequently in this dialect and that its construction, more often than not, resembles the form had + simple past rather than had + past participle. While both of these research teams have accurately described the form of the past perfect in AAVE, they have neglected to take into account its function at the level of discourse. Labov and Wolfram and Fasold noticed that the past perfect was primarily used in narrative style yet neither study investigated the semantic functions of its use in this context. Likewise, the majority of innovative had + past constructions in the Springville data are found in narrative contexts. Since by definition, narrative clauses already refer to events in the past in reference to speech time (Schiffrin 1981), the fact that urban AAVE speakers and rural AAVE speakers with urban ties are using the past perfect frequently in this context suggests some type of reanalysis of the function of this form; therefore, in order to distinguish this form from the traditional past perfect, we refer to the reanalyzed form as innovative had + past. Table 1 below illustrates the distribution in discourse of all had + past constructions for a representative group of Springville informants. The data show that the discourse distribution of innovative had + past favors its use in narrative constructions, the highest proportion of which are found in orientation clauses.

![Frequency of Past Perfect Over Time](image)

**Figure 2. Past Perfect as the Percentage of Total Past Over Time**
Figure 3. Innovative had + past as a Percentage of Total Past Perfect Over Time

Table 1

Discourse Distribution of Innovative had + past Constructions in the Speech of Springville Residents

<table>
<thead>
<tr>
<th>Speaker</th>
<th>orientation clause</th>
<th>complicating action</th>
<th>single event*</th>
<th>listing*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brandy (b. 1982)</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Sheila (b. 1979)</td>
<td>11</td>
<td>8</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>Lamar (b. 1976)</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Travis (b. 1965)</td>
<td></td>
<td>6</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Vanessa (b. 1961)</td>
<td>13</td>
<td>7</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Elsie (b. 1939)</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>40</strong></td>
<td><strong>28</strong></td>
<td><strong>19</strong></td>
<td><strong>12</strong></td>
</tr>
</tbody>
</table>

*Listing and single event occurrences are outside of narrative clauses.

Orientation clauses usually precede complicating action clauses, or the actual telling of the narrative, and serve to 'identify in some way the time, place, persons, and their activity or the situation' (Labov 1972: 364). These clauses differ from other narrative clauses in that they typically express resultative notions by reporting on existing states which are not temporally ordered, or extended processes which may begin before the narrative action itself and continue during that action...
(Schiffrin 1981: 48). Hopper and Thompson’s (1980) distinction between foregrounded and backgrounded events is relevant here since orientation clauses provide background information, which is not temporally ordered within the narrative, and occurs prior to the actual sequence of events described in complicating action clauses, which relate more foregrounded information.

Examples (22) and (23) illustrate the use of innovative had + past in orientation clauses. In example (22), Vanessa’s use of had + past suggests that she views the events of ‘becoming friends’ and the manager ‘calling her her sister’ not as sequenced events, but as a result of some other action, perhaps the fact that they worked together a lot and got along so well. In this brief orientation section to a much longer narrative, Vanessa is offering background information, in essence ‘setting the scene’ for the rest of the narrative. Additionally, there is no explicit reference point stated in conjunction with the use of these forms. Instead there is an implied reference time which is the beginning of the actions that comprise the narrative. As speakers expand the use of had + past into contexts where there is no past reference point, either explicit or implied, we can assume that had + past is no longer functioning in its traditional role and that it is has taken on a new function in the dialect. In both examples (22) and (23), then, innovative had + past functions as a type of remote past signaling that the events described occurred prior to the telling of the narrative. Dahl (1985: 147) reports that the past perfect used to signal past temporal frames is the first step towards a situation where they are used as a general remote past.

(22) Vanessa (b. 1961)
   a. When I was workin’ at Billups
   b. me an’ the manager we had became real good friends
   c. an’ so she had started callin’ me her sister.
   d. So I liked workin’ there
   e. because uh, we did the work together.
   f. We made it easy for each other.
   g. An’ jus’ because she was the manager
   h. she didn’ put everything off on us.
   i. It was mostly like me an’ her work together an’ do everything.

(23) Vanessa (b. 1961)
   a. Nuh uh, lemme tell you
   b. what I did though Lucy.
   c. You know on FM 2542
   d. one day you had took me that road to go back to uh, Shiloh, remember?
   e. I gon do the same thing one night you know.

The complicating action section of a narrative ‘tells the story by relaying a series of temporally-ordered narrative events’ (Schiffrin 1981: 48). These events are typically expressed in the simple past unless speakers wish to convey a change in the order of events, in which case the past perfect is used. Examples (24) and (25) illustrate the use of innovative had + past to describe narrative sequenced past events. It is clear that for both Travis and Sheila innovative had + past can be used to relate foregrounded events, in other words, where event time equals reference time. This is shown in example (24) where Travis relates a series of sequential actions (buying groceries, getting biscuits, and buying biscuits) that are not in
reference to any explicit past event, and in example (25) in Sheila’s short narration of the events in the movie ‘Child’s Play.’

(24) Travis (b. 1963)
 a. Well we used to go at Piggly Wiggly.
 b. They said their stuff, you know, we
 c. my mama had bought some, some uh, groceries from them one time
 d. an’ she had got some biscuits
 e. she had bought from them,
 f. an’ it was, it was sour.
 g. So we jus’ stopped, you know, tradin’ over there.

(25) Sheila (b. 1979)
 a. This doll try to get into this boy body,
 b. An’ then they had killed him at the en’ of it.

As the use of innovative had + past spreads from backgrounded to foregrounded speech contexts, and the past reference point associated with traditional past perfect usage is lost, speakers begin to use this form to express past events in non-narrative contexts as well. This is shown in examples (26) and (27) where innovative had + past is used to relate single events in the past. In example (26), Vanessa’s use of had tol’ indicates that sometime in the past Kelsey told her that she (Kelsey) had applied for a job at Daylight Donuts. What is interesting about Vanessa’s response in this short dialog is that she uses the simple past in the subordinate clause, rather than the past perfect which would have been the more likely choice in light of traditional uses of this construction. (That is, She tol’ me she had put an application in would imply that ‘she told me she had applied for the job before she got it’). We can assume then that temporal ordering is not the reason for the use of innovative had + past; rather these examples suggest a function in which innovative had + past refers to an unspecified time that perhaps more remote than the immediate past. This is exemplified in Travis’ comment in example (27). We had recorded Travis earlier that same day, and in our conversation he said that he liked to paint and had even sold some of his drawings. During our conversation later that afternoon with his younger brother Lamar, the subject of Travis’ art work came up again, which prompted him to make the comment I had told you all about my pictures. The implied past reference point can only be the time of our present conversation, which would then put the time when Travis told us about his pictures somewhere in the remote past, in this case, during our previous conversation.

(26) Vanessa (b. 1961)
 V: So what you been doin’? I’ve seen Kelsey. She started back to school.
 L: Yeah. She’s workin’ at Daylight Donuts.
 V: Yeah she had tol’ me she put an application in.
 L: Yeah I put in a application over there on Diamond Shamrock.
 V: You did?

(27) Travis (b. 1963)
 T: Uh, that big picture in the fron’ room that goes all [L interrupts]
 L: See the school in back of there?
FW: Yeah, uh huh. It says Springville School. Yeah I'll go in in jus' a second an' check it out.

T: Yeah, I, I, I had told you all about my pictures.

FW: Yeah I know. I wanted to see your stuff. That's great.

Sheila is the only person listed in Table 1 who uses innovative had + past as a listing device for unsequenced events, making her the most advanced speaker in the Springville corpus in the use of this form. While the events she lists in example (28) appear to have some kind of order, those listed in example (29) are in no particular order; in fact, temporal order is irrelevant in this passage. She is simply listing a series of related, but unsequenced, past actions: walking around Springville, getting paper to use for grave rubbings, going around town picking up rocks, and going to Bonnie's house to take pictures. For the most advanced speakers in Springville, innovative had + past functions more like an alternative form to express events in the past rather than as a form to distinguish remote from more recent past.

(28) Sheila (b. 1979)

FW: You need to give Fred, you don't, you don't', you said you had a picture of him but it's not here?

L: Nuh uh. It's at my mom house. 'Cause we had went to -- one day my mom had took us out to eat. We had went to, to go eat. Then we had went to the mall, then we had went to Quick As A Flash. An' he paid for the pictures so we could take 'em. We took pictures.

(29) Sheila (b. 1979)

S: She gave it all to us. An' then we had, yeah then we had walked aroun' Springville explorin' everything. We had got paper an' put it on the graves an' scratched it. An' we had went all over Springville pickin' up rocks an' uh, grass. We went all to the graveyard an' everything.

FW: Huh. Where's the graveyard?

S: Right down there.

FW: Oh that's right. Down there. That's right.

S: We had went to Bonnie house. Took pictures of that an' Miss Loretta house.

The data from Springville suggest that as speakers reanalyze the function of had + past, the frequency of its use as a past tense form increases. Figure 4 illustrates this phenomenon for the most advanced Springville speakers. The data also show that the diffusion of innovative had + past and the semantic reanalysis of this form have been a gradual process, with the 'traditional' past perfect (realized as had + past in the youngest generations) coexisting with the reanalyzed form in all speakers (cf. Lichtenberk 1991 on the coexistence of forms during grammaticalization). Further, the data from the youngest speakers also suggest that as innovative had + past continues to spread throughout discourse contexts, it does so at the expense of the older, traditional past perfect, so that innovative had + past occurs in environments where previously only the simple past would be used (i.e., listing). This situation is exactly parallel to the situation described earlier with deleted would, where as be₂ expands, the invariant be that results from would deletion disappears.
Figure 4. Frequency of Innovative \textit{had} + past for the Most Advanced Springville Speakers

The discourse distribution of innovative \textit{had} + past illustrated in Table 1 suggests that as the usage of this form moves through discourse contexts, it becomes more foregrounded, especially in the speech of the Post-1970 generation. Figure 5 illustrates a continuum for the use of innovative \textit{had} + past as its discourse function shifts from expressing traditional backgrounded events to backgrounded and foregrounded events in narratives, as this form begins to grammaticalize over time in the speech of Springville residents:

\begin{figure}
\centering
\includegraphics[width=\textwidth]{continuum}
\caption{Grounding Continuum for Innovative \textit{had} + past in Springville Speech}
\end{figure}

The expansion of innovative \textit{had} + past into non-narrative contexts represents the further grammaticalization of this form. This is shown in Figure 6 which illustrates the complete path of grammaticalization as \textit{had} + past is semantically and grammatically reanalyzed for Springville speakers:
For the oldest generation of Springville speakers and the former slaves, *had* + past is only used to express backgrounded events, in other words past before past. Among the next generation of *had* + past expands to a new environment within the same backgrounded speech context. In other words, *had* + past becomes an option to express not only past before past events, but also to express backgrounded events in a narrative, hence its use in orientation clauses. As an explicit past reference point is lost in conjunction with the use of innovative *had* + past in orientation clauses of narratives, the probability that speakers will begin to use this form to express more foregrounded events within the same speech event increases. Thus *had* + past expands into the complicating action clauses of narratives, the next stage along the grammaticalization path of this form. Because speakers freely use *had* + past with time adverbials to express single foregrounded events within narratives, we can posit an easy transition for its use outside of narratives to express single past events. This would be the next stage as *had* + past grammaticalizes. The final stage in the grammaticalization path of *had* + past represented in the Springville data is its use for the listing of unsequenced past events. The expansion of *had* + past to this context is best understood in light of the fact that unsequenced listings are no more than a series of single past events.

4.0 Conclusion

Although most work on grammaticalization relies on historical documents, a growing body of work, such as Schwenter’s 1994 analysis of the grammaticalization of the perfect in Alicante Spanish, is beginning to look at grammaticalization in progress. As this analysis of *be*₂ and innovative *had* + past suggests, quantitative sociolinguistics provides an extremely useful methodology for studying grammaticalization as it is taking place. The linking of sociolinguistic methodology with the substantive and theoretical insights of work on grammaticalization creates an exciting opportunity for research that not only tracks grammatical change but also explores the motivations for it.

---

1 The research for this project was supported by a series of grants from the National Science Foundation (BNS-8812552, BNS-9009232, and BNS-9109695) and by grants from the University of Michigan and Texas A&M University.

2 For a more detailed explanation of our field methods and interview contexts see Cukor-Avila and Bailey (under review) and Cukor-Avila (1995).
It is interesting to note here that we have three tokens of \( be_2 \) (\( be+V+ing \)) with habitual meaning among our elderly adults. In spite of the fact that none of the tokens can plausibly be derived from \( would \) deletion, all of them have past reference. They seem to reflect a kind of transitional stage from the invariant \( be \) that derives from \( would \) deletion to the \( be_2 \) of contemporary urban AAVE.

The authors wish to thank Kathy Carey whose insights about semantic change and grammaticalization processes and intuitive suggestions about the \( had + \) past data are greatly appreciated.

---


Cukor-Avila, Patricia and Guy Bailey. (under review) An approach to sociolinguistic fieldwork.


The death of ‘prefixing’:
contact induced typological change in northern Australia
Ian Green
Australian National University

1. THE MINDI LANGUAGES. The Mindi languages of northern Australia present the socio-historical linguist with an intriguing picture of typological diversity. Consisting of five modern languages, the Mindi varieties are a genetic unit, forming a sub-group of the Australian language family. (The sub-group takes its name from the form of its innovative first inclusive dual pronoun, mindi, reflected in all daughter languages). The Mindi sub-group is geographically discontinuous. To the west is the ‘Yirram’ or ‘Western Mindi’ branch, made up of two languages: Badj (with dialects Djamindjung and Ngaliwurru) and Nungali. More than 200 kilometres to the east, and separated from the Yirram bloc by the Wardaman and Mudburra languages, is the ‘Barkly’ branch. This consists of Djingili, Nganga and a chain of three dialects, Binbin.ga, Gudandji and Wambaya, making up the so-called ‘McArthur’ language¹. Genetically the Barkly branch divides in two; Djingili makes up one sub-branch, while Nganga and McArthur make up the other, referred to here as ‘Eastern Mindi’. (1) below gives the structure of the Mindi family tree², while the Appendix presents a map showing the geographic position of the Mindi languages within central northern Australia.

(1) THE MINDI SUB-GROUP (after Chadwick 1978, 1984)

Mindi

Western Mindi (Yirram) | Barkly
---|---
Badj | Nungali | Djingili | Eastern Mindi

Ngaliwurru | Djamindjung

McArthur | Nganga

Binbin.ga | Gudandji | Wambaya

Evidence for the close genetic linkage of the Western Mindi and Barkly blocs comes not primarily from lexical comparison, lexical diffusion in the region having obscured somewhat the genetic picture. Wambaya, for example, has only 17% cognate vocabulary with its close relative Nungali, but scores 35% with its eastern neighbour Garrawa. Similarly, Djingili has a mere 9% of its vocabulary in common with Ngaliwurru, compared to 28% with its southern neighbour Mudburra. The major evidence for the genetic unity of the Mindi languages is, rather, morphological. Recognition of this is due to Chadwick (1984), who set out the correspondences in pronominal paradigms and nominal class markers between Western Mindi and Djingili, and identified a number of cognate innovations from proto-Australian that established the sub-grouping. While Chadwick’s analysis focussed on Djingili, largely to the exclusion of the other Barkly languages, his arguments have subsequently proved extendable to Eastern Mindi, and the overall Mindi sub-grouping is now uncontroversially accepted by Australianist scholars (cf. Dixon et al to appear (b)). Surprisingly, however, no detailed reconstruction
that would reveal the character of the proto-language has been attempted. But this remains an urgent task, since the Western Mindi and Barkly blocs exhibit a striking typological difference. For while the Western Mindi languages are ‘prefixing’ in nature, the Barkly varieties are, on the other hand, to be regarded as ‘suffixed’.

2. MINDI TYPOLOGICAL DIVERSITY: PREFIXING VS. SUFFIXING. There are two major respects in which the Western Mindi’s are prefixing languages. Firstly, their verbs are formed with pronominal prefixes. Verbs in the Western Mindi languages are typically composed of a non-inflecting main verb stem followed by an ‘auxiliary’, which constitutes a separate phonological word. The auxiliary consists of a sequence of pronominal and mood prefixes to an auxiliary verb stem (AVS) which is itself inflected for various tense categories. The prefixed pronominals index core participants: intransitive subject (S), transitive subject (A) and object (O). The mood prefixes mark the imperative, irrealis positive and irrealis negative categories. The auxiliary verb has a semantic classificatory function, indicating the action-type class to which the verb belongs. Each of the Western Mindi varieties has of the order of 20 of these verbal classifiers, ranging over intransitive (e.g. ‘sitting’, ‘standing’, ‘coming’, ‘going’, ‘be in the process of burning’) and transitive (e.g. ‘handling’, ‘seeing’, ‘hitting’, ‘telling’) action types.

(2) WESTERN MINDI VERB STRUCTURE:

VERB --> Main Verb Stem # Auxiliary
AUX --> ( { Irrealis Neg Imperative } ) - S/A -(O) - (Irrealis Pos) -AVS.Tense

e.g. Nungali (Bolt, Hoddinott and Kofod 1971b):
ngilidiga wa -wu -rrum -am
cry 3sgS Irr Pos come Future
He will be coming crying.

Secondly, the Western Mindi bloc exhibits prefixing in its nominal noun-class marking. Synchronically this is found only in Nungali. Nungali has four noun classes, essentially masculine, feminine, vegetable, and residue. Noun class membership is primarily marked by concord, with obligatory prefixes on NP modifiers (e.g. adjectives, possessive pronouns, numerals) as well as demonstratives. There is also lexically specified class marking of noun stems themselves, prefixes being obligatory for some noun stems, optional for some and proscribed for others. The noun class prefixes vary for a three way macro-case opposition of Ergative/Locative/Instrumental vs. Absolutive vs. Dative/Oblique. This macro-case opposition is supplemented by a system of suffixal case-marking which is invariant for noun-class:

(3) NUNGALI CLASS-CASE PREFIXING (Bolt, Hoddinott and Kofod 1971b)
Masculine Absolutive di-nad du-ngunin big man
Masculine Ergative nji-nad nji-buri big dog
Masculine Oblique giya-malin giya-ngara for the good kangaroo
Feminine Absolutive nga-nad nga-ngaran big woman
Feminine Ergative nganji-nad nganji-ngarung big woman
Feminine Oblique ganji-malin ganji-yilgin to the good mother

Body-part terms are distributed among the four classes. For those body part terms that belong to the vegetable and neuter classes there is a second order class marking, whereby a morpheme denoting the gender of the possessor follows the initial class-case prefix. These morphemes do not vary for case, and have the form ya masculine and na feminine. Thus compare mi-ya-luwal vegetable:Absolutive-
masculine-knee 'his knee' with mi-na-luwal vegetable: Absolutive-feminine-knee 'her knee', and ni-ya-man.ga neuter: Absolutive-masculine-ear 'his ear' with ni-

na-man.ga neuter: Absolutive-feminine-ear 'her ear'

In place of the Western Mindi auxiliary, consisting of pronominal and tense-mood prefixes to an auxiliary verb stem, the Barkly languages have an auxiliary which is described (e.g. Chadwick 1975, 1978; Nordlinger 1993) as consisting of just a sequence of bound pronominals and tense-aspect-mood (TAM) morphemes. A generalised structure for this Barkly auxiliary is given in (6) below. (Note that not all languages have all the TAM marking options shown in the formula.) In Djingili this auxiliary is a suffix to the main verb stem. In Eastern Mindi it is a second position enclitic, following either the first word or the first constituent of the clause. In Wambaya, for example, this enclitic behaves phonologically as a suffix when mono-syllabic, and as a separate word (i.e. as a separate stress domain) otherwise. And in contrast to the twenty or so action-type categorising auxiliary verbs of the western bloc, the Barkly unit varies in form according to just a three-way directional opposition, determined by whether the speaker construes, and wishes to mark the action as, accomplished with motion towards the speaker, or with motion away from the speaker, or as motion-neutral. But there is no systematically segmentable auxiliary verb root associated with this categorisation; in fact there is no morphological head in this unit at all (Nordlinger 1993:162-164), the motional categorisation being effected not via a discrete morpheme but rather through variation in the form of the final TAM affixes.

(4) BARKLY GENERAL VERB STRUCTURE

AUX --> \{ Neg Interrogative Future

\} - S/A - (Future) -(O) - TAM. Direction

Past

(5) e.g. Djingili (Chadwick 1975)

(a) ngadja -ni -mindi -dju

see 3sgA 1Inc diO Pres neutral

He's watching you and me.

(b) ngadja -ni -mindi -djiimi

see 3sgA 1Inc diO Past towards

He came and saw you and me.

(c) ngadja -ni -wurr -adu

see 3sgA 3plO nonPast away

He's going to see them.

(6) e.g. Wambaya (Nordlinger 1993)

iligirri-nmandji ngurr -uba yarru

river ALLATIVE 1Inc pl S nonPast away go/come

We're all going down to the river.

Class marking strategies in the Barkly group provide an even starker typological contrast with the Western Mindi branch than do verb structures. Semantically, the Barkly languages in this respect align closely with Nungali, all possessing a matching set of masculine, feminine, vegetable and neuter classes. And the Barkly's noun class markers also show macro-case oppositions, although here the opposition is primarily a two-way absolutive vs. non-absolutive contrast, and less frequently involves a division of the non-absolutive into ergative and dative type categories in the same way that Nungali regularly does. The outstanding difference, though, is that in the Barkly class-case marking is not generally effected through
prefixes, but rather by suffixes. Class-case suffixing is obligatory for head nouns and NP modifiers, as illustrated by the Wambaya data below4.

(7) WAMBAYA CLASS-CASE SUFFIXING (Noldinger 1993)

<table>
<thead>
<tr>
<th>Gender</th>
<th>Absolutive</th>
<th>non-Absolutive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masculine</td>
<td>bungmadj-dji</td>
<td>bungmadj-ni</td>
</tr>
<tr>
<td></td>
<td>old man</td>
<td></td>
</tr>
<tr>
<td>Feminine</td>
<td>bungmadj-na</td>
<td>bungmadj-nga</td>
</tr>
<tr>
<td></td>
<td>old woman</td>
<td></td>
</tr>
<tr>
<td>Vegetable</td>
<td>manugudja-ma</td>
<td>manugudja-mi</td>
</tr>
<tr>
<td></td>
<td>conkerberry fruit</td>
<td></td>
</tr>
<tr>
<td>Residue</td>
<td>bagi-ga</td>
<td>bagi-gi</td>
</tr>
<tr>
<td></td>
<td>big</td>
<td></td>
</tr>
</tbody>
</table>

But there is one major exception to suffixal class-case marking in the Barkly group. Unlike other nominal constituents, the demonstratives of the Eastern Mindi branch regularly affect their noun class-case marking by applying prefixes to the stem. The class-case demonstrative prefixes are largely identical to the regular nominal suffixes, as can be seen by comparing the final syllables of the Wambaya data in (7) with the initials in (8):

(8) WAMBAYA REMOTE DEMONSTRATIVE: PREFIXES+STEM (Noldinger 1993)

<table>
<thead>
<tr>
<th>Gender</th>
<th>Absolutive</th>
<th>Ergative/Locative</th>
<th>Dative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masculine</td>
<td>i-ni-yaga</td>
<td>ni-nki-yaga</td>
<td>ni-nagi-yaga</td>
</tr>
<tr>
<td>Feminine</td>
<td>na-ni-yaga</td>
<td>nga-nki-yaga</td>
<td>nga-nagi-yaga</td>
</tr>
<tr>
<td>Vegetable</td>
<td>ma-mi-yaga</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>Residue</td>
<td>ya-ni-yaga</td>
<td>----</td>
<td>----</td>
</tr>
</tbody>
</table>

Some demonstrative forms and adjectives with demonstrative-like discourse functions (‘one’, ‘different’, ‘other’) in Djingili also have some class marking by prefix, although these show no variation for case (Chadwick 1984:25).

3. THE MINDI LANGUAGES IN AREAL PERSPECTIVE. The Western Mindi and Barkly groups (as diagrammed in the Appendix) sit on opposite sides of the ‘prefixing’ boundary. The Western Mindi group has non-prefixing languages to its south, and prefixing languages on all other sides. The Barkly group has prefixing languages on its northern borders, but is otherwise surrounded by suffixing-only languages. Now the distinguishing feature of all languages categorised as prefixing is that they possess structures in which bound pronouns attach to and precede a verbal element; this element may be either a main verb stem or some type of auxiliary verb. These bound pronouns have the phonological status of prefixes; the verbal element is typically segmentable as a separate morpheme, and is usually TAM inflected. But it is not the case that a post-prefixal verbal element can be universally extracted in the prefixing languages, there being some tendency for auxiliary verb stems to fuse with TAM affixes into a synchronically non-segmentable unit. The Barkly languages can thus be viewed, as per Blake’s (1990) reclassification, as aberrant prefixing languages; that is, they are ‘prefixing’ in that they attach bound pronouns to a following portmanteau morpheme which has both TAM and, insofar as it classifies the verb for a three-way motional contrast, auxiliary verb-like functions. But they are aberrant in respect of the status of their auxiliaries. Djingili is marginally so, in that its auxiliary (except in its simple verb function) is a series of suffixes to the main verb stem rather than a separate phonological word. The Eastern group is more strikingly so; while Djingili has its main verb stem and auxiliary juxtaposed, as have all the prefixing languages proper, the Eastern group has its auxiliary as a second position clitic, thus allowing it (given that word order is free) to be dissociated from the main verb stem. Note that the Barkly languages also diverge from the prefixing bloc in general in having
only three auxiliaries, prefixing languages typically having auxiliary inventories at least five to six times larger (Dixon et al to appear (a)).

It is the attachment of pronominals to a following auxiliary verb-like (i.e. verb classifying) element which can be seen as a critical feature that distinguishes the Barkly group from its non-prefixing neighbours, which otherwise have structures analogous to those of the Mindi group. For example, while Mudburra (McConvell 1980) also has bound pronominals, these are enclitics rather than proclitics or prefixes. Mudburra bound pronouns may follow verb stems, or they may suffix to a base morpheme that has some TAM function, thus replicating the initial [TAM - Bound Pronoun] sequence of the Mindi auxiliary. In addition Mudburra, in a manner reminiscent of the verbal classification of Western Mindi, constructs its verbs by combining non-inflecting verb stems with one of about thirty TAM marked auxiliary verbs. But the bound pronominals and auxiliary verbs remain discrete grammatical elements, and there is no systematic preposing of the former to the latter.

Equally, Warumungu (Simpson and Heath 1984) and Garrwa (Belfrage 1992) both show some TAM suffixation or encliticisation of pronouns. And both have a type of verbal classification, described as a verbal category of ‘associated motion’ (Koch 1984), which is semantically similar to that found in the Barkly group. Garrwa can mark actions performed ‘while in action’ with a serial construction that postposes the root djila ‘go’ to the main verb. Warumungu has four motional categorisations: motion away, motion towards, start-of-motion and an unmarked motion neutral. As in Garrwa, these are categories that are marked on the verb, in this case as suffixes. As in Mudburra, neither Garrwa nor Warumungu tie their bound pronouns to following main or classificatory verbal elements.

From an areal perspective it is also the Barkly rather than the Western Mindi group which in respect of its class marking is exceptional. Class marking is an areal feature of Australia’s central north, found over a region that covers a large portion of the prefixing bloc and extends into the Barkly group and south-east into Wagaya. In this region only the Barkly group and Wagaya, together with the Jarrakan languages to the immediate west of Western Mindi, encode class membership with suffixes rather than prefixes. And class suffixing in Wagaya is best explained as a recent acquisition from the neighbouring Barkly languages (Brammall 1991).

The recognition of the sub-grouping of Western Mindi and Barkly thus raises challenging historical questions as to the nature of proto-Mindi and the consequences of language contact. Was proto-Mindi prefixing, and has the Barkly group reorganized its structures in the suffixing mold under the influence of its more southerly neighbours? Or was proto-Mindi suffixing, the Western branch then developing its character - placing bound pronominals in front of classifying verbs and acquiring class prefixes - through diffusion from the north? And what sort of verbal classification did it have: a system like the twenty or so action-type categories of the contemporary Western languages, or something more like the three-way directional opposition found in the Barkly?

4. PROTOMINDI AS A PREFIXING LANGUAGE. In respect of verbs, all the comparative evidence points clearly to proto-Mindi having been a prefixing language, that is, one that attaches bound pronouns prior to verbal stems. While reconstruction of the proto-Mindi verb is naturally a lengthy task well beyond the scope of this paper, the following are for our purposes the critical points in the enterprise.
(a) Firstly, as implicit in Chadwick (1984), proto-Mindi had at least singular bound pronouns. That is, we can reconstruct fairly confidently the major portions of the free and bound pronominal paradigms of the proto-language; the reconstructable singular bound paradigms include forms distinct and not derivable from the equivalent free forms. For example, as illustrated below, *nganjdv is the proto-Mindi second singular bound A form, but *Nami (where N = n or nj) the second singular free stem. Similarly, *(ng)ana is the first singular bound O, but *ngarrgu the first singular free form that had O function in the proto-language.

(9) Djamindjung Jingili Binbin.ga proto-Mindi
  2 sg A bound nganhdh nganjja (nj)dv *nganjdv
  2 sg A free nami nama-njama njami *Nami
  1 sg O bound an ara ng *(ng)ana
  1 sg O free ngarrgu ngarru ngari *ngarrgu

(b) Secondly, as noted in Chadwick (1984), the auxiliary-initial imperative markers of Western Mindi and Jingili are cognate, and reconstructable for the proto-language as *ba. This imperative marker has been lost in the Eastern group, where imperative singular goes unmarked, and imperative non-singular employs the regular second person bound pronominal forms.

(10) AUXILIARY INITIAL IMPERATIVE AFFIXES

    Djamindjung Jingili Wambaya proto-Mindi
    2 sg    ba wa-Ø Ø *ba
    2 dual ba-wunji(i) wa-nju gu1 *ba-wunji
    2 plural ba-wurru (wa)-rru girr *ba-wurru

One further auxiliary-initial TAM marker appears to be shared between the Western and Barkly groups. For example, Binbin.ga has an auxiliary-initial yi which marks interrogative future negative (Chadwick 1978), and which would appear to be cognate with the nga~ya auxiliary-initial morpheme found in the Western group, and which has a general future negative function.

(c) Thirdly, as Nordlinger (1993:285) has demonstrated, a number of the TAM-direction affixes of the Barkly auxiliary show correspondences with the intransitive auxiliary verbs of the Western branch. For example, Nganga has an element agba in its habitual past motion-neutral affix. Jingili has a likely cognate form aga as its motion-neutral past. Within the array of Barkly TAM-direction affixes both Nganga agba and Jingili aga are irregular. A potentially cognate gagba is also found, again as an irregular form, within the Western branch, where it appears as the suppletive past tense of the auxiliary verb 'be' in Djamindjung. Nordlinger notes that similarly the Wambaya past motion-away affix (g)anj is likely to be historically related to the Djamindjung past tense form ganj of the verb 'go'. Equally, the Wambaya habitual motion-neutral past adji, presumably derives from yadji, the irrealis form of the Djamindjung auxiliary verb 'be'7. It is also likely that the Jingili motion-away affixes in rrv are cognate with the Djamindjung AVS rrum 'come', the opposing orientations in the two branches resulting from differing directional interpretations of an originally directionally unmarked motion verb8.

(d) Points (a) to (c) above together establish proto-Mindi as a language with a sequence of TAM and pronominal affixes that preceded two or three motional and copular verbs. There are no apparent correspondences for any of the Western transitive verbs. However, the Western transitive set includes a number of auxiliary verbs found widely in northern Australia, and which are most likely part of an ancestral set of classifying or compounding verbs; it is most unlikely that this set as a whole has diffused, both in form and function, over the northern Australian
region (Dixon et al to appear (a)). Ancestral transitive stems reflected in the Western Mindi classifier set include *wanda* 'take', *wa* 'bite', *injdi* 'eat', *yada* 'spear', *ma* 'have', *rra* 'handle' and *unggum* 'say'. Despite the lack of cognates in the Barkly branch, we can conclude, then, on the basis of a top-down reconstruction, that proto-Mindi had at least nine or ten auxiliary verbs. And although there is evidence in the Barkly group of only intransitives being pronominally prefixed, the simplest scenario is that all the auxiliaries of proto-Mindi were prefixed, and that in the Barkly the original transitive inventory has undergone wholesale deletion. This scenario is consistent with the typology of the modern prefixing languages, since no prefixing languages have prefixing for some of their classifier verbs, but not for others. A prefixing origin is the only plausible way of accounting for the consistent association of the bound pronouns with TAM-verbal classificatory morphemes in the Barkly auxiliary, since these morphemes are not systematically linked in any of the surrounding languages, such as Mudburra and Warumungu, to which we would look to provide precedents for any model of a non-prefixing proto-Mindi. On the other hand, a scenario that had proto-Mindi with the structure of Mudburra, for example, would require coincidental changes in both the Western and Barkly groups, each branch independently locking its inherited Mudburra-like [TAM - Bound Pronoun] sequence into a position prior to the auxiliary verb, and each preserving no trace of the original separability of the combined elements.

Turning to the matter of noun-class marking we find the character of proto-Mindi more difficult to access. Comparison of the data across the two Mindi branches permits reconstruction of only a partial set of proto class markers, as shown in the comparison of the Nungali (Bolt, Hoddinott and Kofod 1971b) and Wambaya (underlying forms from Nordlinger 1993) affixes in (11) below. Only a two-way, absolutive vs. non-absolutive, contrast is formally reconstructable, dative forms not showing clear cognates and being excluded from (11). Reconstruction of the proto absolutive set proceeds, except for the vegetable class, from recognition of the formal correspondence between the second syllable of the Nungali concordial form and the Barkly form. This procedure follows that of McConvell (1985). It assumes that there were innovations in the Nungali system which saw masculine, feminine and neuter absolutive markers (*di*, *nja*, and *nV* respectively) simply added to the original set. That the *ya* and *na* in the masculine and feminine agreement markers of Nungali had previous status as separate morphemes is supported by the existence of the *ya* masculine and *na* feminine second-order marking for possessor of vegetable and neuter class body parts (as discussed under example (3) above). And the absence of these second syllables from the prefixes for noun stems suggests that NP-head prefixing could itself have been a Nungali innovation, one which accompanied the innovation of the new absolutive prefixes, and which saw these newer prefixes applied to the head along with the non-absolutive prefixes from the concordial set.

<table>
<thead>
<tr>
<th>(11)</th>
<th>Nungali Abs (nouns)</th>
<th>Nungali Abs (adjectives)</th>
<th>Wambaya Abs</th>
<th>pMindi Abs</th>
<th>Nungali Erg</th>
<th>Wambaya non-Abs</th>
<th>pMindi non-Abs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Masculine</td>
<td>Feminine</td>
<td>Vegetable</td>
<td>Neuter</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>di</em></td>
<td><em>nja</em></td>
<td><em>ma</em></td>
<td><em>nu</em></td>
<td></td>
<td><em>nji</em></td>
<td><em>ni</em></td>
</tr>
<tr>
<td></td>
<td><em>di(ya)</em></td>
<td><em>nja(na)</em></td>
<td>*ma, <em>mi</em></td>
<td>*nu(wa), <em>ni(wa)</em></td>
<td></td>
<td><em>nganj</em></td>
<td><em>nja</em></td>
</tr>
<tr>
<td></td>
<td><em>dji, yi, i, Ø</em></td>
<td><em>na</em></td>
<td><em>ma</em></td>
<td><em>ga</em></td>
<td><em>wunji</em></td>
<td></td>
<td><em>mi</em></td>
</tr>
<tr>
<td></td>
<td><em>dJV</em></td>
<td><em>nJa</em></td>
<td></td>
<td></td>
<td></td>
<td><em>mi</em></td>
<td><em>gV</em></td>
</tr>
</tbody>
</table>


The reconstructable case oppositions of *na vs *nga in the feminine, and *ma vs *mi in the vegetable\(^9\), indicate that class-case marking was a feature of the proto-language, and did not originate independently in each branch. The forms of these markers, however, offer no clue as to whether they were originally prefixes or suffixes. Further, while the Barkly group has class-case prefixed demonstratives and some adjectives, these are not formally relatable in any clear way to those of Nungali. Nevertheless, the prefixed demonstratives of the Barkly perhaps are the key to the status of the class-case morphemes in proto-Mindi. Any account of proto-Mindi as having purely suffixal class-case marking would require an implausible wholesale reanalysis of noun suffixes as prefixes to following modifiers in Nungali, and a coincidental reanalysis confined just to demonstratives in the Barkly group. On the other hand, if proto-Mindi were thoroughly prefixing in its class-case marking, how is the Barkly shift to suffixing in all but the demonstratives to be accounted for?

Chadwick has suggested rather that nominal class-case suffixing in the Barkly group has developed from ‘postposed markers not previously attached to the noun stem’ (1978:336), but does not explain why demonstratives, in contrast to most other nominals, should have acquired prefixing. Reacting to Chadwick’s idea, Nordlinger (1993:281) proposes that postposed prefixed demonstratives are instead the source of the Barkly suffixes. This would explain the effective formal identity of demonstrative prefixes and nominal suffixes, and, since demonstratives would not be expected to be postposed to demonstratives, it would also explain why demonstratives themselves do not take class suffixes. Unfortunately, Nordlinger’s demonstrative hypothesis lacks the necessary formal support, since there is no trace in the Barkly suffixes of any specifically demonstrative morphology (e.g. such as a stem or an exclusively demonstrative case allomorph). It seems plausible then to combine these accounts, and suggest that in proto-Mindi the precursors of the modern noun class-case affixes were in fact postposed as separate words to the head noun, as suggested by Chadwick. At the same time proto-Mindi must have had class-case prefixed demonstratives, as suggested by Nordlinger, these demonstratives representing a first extension of prefixing strategies from the verbal to the nominal domain. The nominally postposed class-case markers functioned as part of the overall demonstrative paradigm; they had the same class-case stems as the other demonstratives, and they alternated, rather than co-occurred, with them. Precedents for such systems, with noun class markers or ‘articles’ belonging to the set of prefixed demonstratives, are found in prefixing languages such as Mara (Heath 1981), Maung (Capell and Hinch 1970) and Ngarinjin (Rumsey 1982). In the Barkly group the postposed class-case markers became absorbed into the noun stem as suffixes. However the other demonstratives were maintained as prefixal, the suffixisation process able to fuse independent postposed words to the noun, but not extending to the disruption of established words by stripping them of their prefixes. The Western Mindi branch saw noun class-case markers develop into prefixes to following modifiers in Nungali; prefixing was subsequently extended to head nouns. Class marking as a feature, however, was dropped altogether in Badj.

4. THE LOSS OF PREFIXING. The most plausible scenario motivated by the comparative evidence, then, is that proto-Mindi was a verbally prefixing language, and that significant changes were effected to its Barkly branch as a result of contact with suffixing-only languages.
The mechanisms of these changes are fairly straightforward. Speculating on a prefixing origin for the Barkly languages Blake (1990) has suggested that Djingili acquired its phonologically suffixing verbal structure under influence from Mudburra; though not elaborated in any detail, Blake’s proposal implies that the Mudburra feature of postposing tense-inflected (though not pronominally marked) classifier-like verbs to an uninflceted main verb served as the model for Djingili’s ‘putting the auxiliary permanently at the end of lexical verb stems’ (p55). But Blake’s suggestion is not properly informed by data on the Western Mindi verb, which shows clearly that Djingili did not have to have recourse to Mudburra influence to create a verb plus auxiliary construction. The verb plus auxiliary sequence in Djingili is not an innovation on, but rather inherited from, proto-Mindi. In proto-Mindi, as in the modern Western branch, the auxiliary most likely had the phonological status of a separate word. The according of suffixal status to the auxiliary then amounts to no more than a minor rearrangement of stress assignment within the overall verbal complex, the loss of the verb - auxiliary word boundary accommodating the Djingili verb to speakers entirely lacking prefixing strategies in their linguistic repertoire.

Across the Barkly group as a whole, in the Eastern sub-branch as well as Djingili, the same process has applied to the class-case markers. That is, a postposed element bearing grammatical information pertinent to the phrase as a whole has been reanalysed as suffixal to the phrasal head.

The Djingili verb, representing an initial response to contact with suffixing-only languages, would appear to preserve the earlier Barkly structure. The second position enclitic status of the Eastern Mindi auxiliary then constitutes a subsequent development. While the Djingili verb conforms to a suffixing-only model, it nonetheless has an aberrant structure relative to languages such as Mudburra and Warumungu, whose bound pronouns commonly occur as second position enclitics. The freeing up of the rigid verb stem plus auxiliary ordering of Djingili then brings this regionally aberrant structure into line with these surrounding languages.

This type of change could well have taken place through contact of proto-Mindi descendants with any of the non-prefixing languages surrounding the contemporary Barkly group. But there is another factor that points towards contact with the Garrwa/Warumungu part of the region, rather than the west, as the motivating factor in the changes. This concerns the reduction of the AVS inventory.

No doubt facilitating the shift to suffixing structures in the verb was the reduction of the original AVS inventory. In terms of identifying a likely influence on proto-Barkly this decidedly points to the Warumungu-Garrwa region. These languages lack the large scale type of action/event classing by auxiliary verb that would have been inherited into proto-Barkly from proto-Mindi. On the other hand, proto-Mindi verbal classification would not appear to be incompatible with the verbal systems of Mudburra and its close genetic relatives to the west, which all employ dual verb stem constructions that combine main verb stems with a wide range of classifier-like auxiliary verbs, but which lack a specific verbal category of ‘associated motion’. The reduction of the classifier inventory of proto-Mindi to a three-way motional opposition, then, appears to be best explained as a simplification of the system for speakers unfamiliar with verbal classification in general, and only able to interpret it in terms of their familiar system of associated motion.

There is some formal support for Garrwa in particular having influenced proto-Barkly. For example, the irregular dative/oblique marker *ngi reconstructable for just the third singular feminine in proto-Barkly appears to be a borrowing from
Garrwa, since *ngi* is the regular pronominal dative suffix in Garrwa. Further, reconstructable for proto-Barkly, and found in all languages of the Barkly group, is the procedure of forming reflexive verbs by placing a morpheme *ngg(̣)* in the bound pronominal object slot in the verb (that is, immediately preceding the TAM-direction marker, cf. (4) above). The Western Mindi strategy, however, is quite different. In Western Mindi reflexive verbs are derived by suffixing a morpheme *djV* to the AVS; the reflexive verb functions as an intransitive, and nothing may appear in the O prefixal slot. Thus compare (12) and (13).

(12) Djingili (Chadwick 1975)

*uluqadjja -ni -nggu -nu*
wash 3sgA Reflexive Past neutral
He washed himself.

(13) Nungali (Bolt, Hoddinott and Kofod 1971b)

*mung nga -ngayi -dji -na*
look 1sgA see (AVS) Reflexive Past
I looked at myself.

Wider comparative evidence (Dixon et al to appear (a)) establishes the Western Mindi strategy here as the ancestral one, and the Barkly branch as the innovator. Garrwa, out of all the languages of the region, alone provides us with a source for this Barkly development. Garrwa constructs all its free form reflexive pronouns (except for the irregular first singular) by suffixing *ngga* to a pronominal stem. This then appears to be a Garrwa intrusion into proto-Barkly, the bound pronominal subject plus reflexive morpheme sequence in proto-Barkly replicating the free form reflexive pronoun that we find in modern Garrwa.

Further investigative work, comparing the grammars and lexicons of the Barkly languages and their neighbours, and attempting both to identify and stratify diffusion in the region, is required to verify whether, as this data seems to suggest, Garrwa was indeed the major influence in the evolution of proto-Barkly from proto-Mindi. Hopefully such further investigations will also put us in a position to make inferences about the social conditions that attended these linguistic developments; we may then be able to determine whether these developments are the results of gradual convergence between discrete speech communities, or whether they came about more abruptly, through interference in language shift as speakers of a non-prefixing language adopted proto-Barkly, imprinting on it their suffixing-only native linguistic patterns.

1 The following orthographic conventions are used: retroflexes are represented with an underscore (*d*, *n*, *l*), lamino-palatals as *dj* and *nj*, dentals as *dh* and *nh* and velar nasals as *ng*; sequences of alveolar nasals followed by velar stops are given as *n.g.*

2 Chadwick (1984) is at points equivocal about the positioning of Djingili in the family tree; for further discussion of the branching presented here see Dixon et al to appear (b).

3 My use of the term ‘prefixing’ here follows Australian typological practice, where it is used as shorthand for ‘prefixing and suffixing’; that is, it denotes languages which are able to employ both prefixing and suffixing strategies in the formation of their agglutinative structures.

4 Underlying forms of roots and affixes are given in (5).
Jarrakan class suffixing appears to derive from the absorption into the stem of postposed generics; the forms are not cognate with those of the Mindi group, and the system has fewer categories (Dixon et al to appear (a)).

Note that there is as yet no higher level sub-grouping of proto-Mindi that would establish a prefixing or non-prefixing legacy for it. Both Mindi branches are classified as of the 'non Pama-Nyungan' type (Blake 1988) on the basis of their free pronominal forms, but no consequences for their prefixing/non-prefixing ancestry issue from this. Equally, there is no supporting archaeological data that would locate proto-Mindi clearly within an ancestral prefixing or non-prefixing bloc.

Note that the habitual and irrealis categories are regularly formally related in northern Australian languages (Green 1995).

For a wider view of the prehistory of the *rrV motional auxiliary verb, see Dixon et al (to appear (a)).

Note that the reconstruction assumes that the mi vegetable class allomorph in the Nungali absolutive agreement set is derivable from ma.


Bolt, J. E., W. G. Hoddinott and F. M. Kofod. 1971a. An elementary grammar of the Ngaliwuru language of the Northern Territory. Australian Institute of Aboriginal and Torres Strait Islander Studies, MS.

----. 1971b. An elementary grammar of the Nungali language of the Northern Territory. Australian Institute of Aboriginal and Torres Strait Islander Studies, MS.


----. 1984. The relationship of Jingulu and Jaminjungan. Australian Institute of Aboriginal and Torres Strait. Islander Studies, MS.


Simpson, Jane and Jeffrey Heath. 1984. Warumungu Sketch Grammar. Australian Institute of Aboriginal and Torres Strait Islander Studies, MS.

APPENDIX: THE MINDI GROUP IN AREAL PERSPECTIVE (adapted from Blake 1990)
A NEW HYPOTHESIS ON THE ORIGIN OF THE EASTERN ANDALUSIAN VOWEL SYSTEM

José Ignacio Hualde  Benjamin P. Sanders
University of Illinois  Northeastern Illinois University

Since the interest and originality of the Eastern Andalusian vowel system was pointed out in articles by Navarro Tomás (1939) and Alonso, Canellada & Zamora Vicente (1950), this phenomenon has attracted the interest of numerous researchers, who have studied it from very different theoretical perspectives (cf. Alarcos 1958, 1983, Salvador 1957, 1977, Hooper 1976, Gómez Asencio 1977, Zubizarreta 1979, Cerda Massó 1984, Manaster-Ramer 1989, Villena Ponsoda 1987, Listerri & Poch 1987, Lieber 1987, Mondéjar 1991, Sanders 1994, among others). In our view, nevertheless, the historical development of the system has not been adequately elucidated.

In the relevant Andalusian varieties, we find oppositions such as [lifro] ‘book’ vs. [lifro] ‘books’, [lifre] ‘free, sg.’ vs. [lifre] ‘free, pl.’ and [lifra] ‘pound’ vs. [lifrâ] ‘pounds’. In these pairs, the quality of the final vowel has a distinctive function. We, thus, find a greater number of vowel oppositions than in other Spanish dialects. Some authors have spoken of ‘vowel doubling’ with respect to this phenomenon. There are authors who have maintained that for each of the five vowels of standard Spanish there are two in Eastern Andalusian, one close or tense and one open or lax. This is the position that Gregorio Salvador adopted in his 1957 study where a ten vowel system i, ï, e, ẽ, a, ą, o, ô, u, u is proposed.

Nevertheless, it now seems certain that the high back vowel /u/ does not participate in this ‘vowel doubling’ (Salvador 1975, 1977) and the existence of an opposition between two variants of the high front vowel /i/ is also dubious (Moya Corral 1979:25, Sanders 1994). The doubling is thus very likely limited to the nonhigh vowels, with a surface opposition between tense and lax mid and low vowels resulting in an eight-vowel surface system:1 i, ẽ, ą, a, q, ô, u. The merger between the low-mid front vowel and the fronted low vowel has also been reported for certain areas (Alonso 1956, Alvar 1958).

The problem that we want to address in this paper is that of the origin of this ‘vowel doubling’ phenomenon. The standard, and seemingly unanimous, view on this matter is that the loss of final /s/ has somehow triggered the appearance of open or lax vowels. More specifically, the assumption is that the final aspiration found as an intermediate stage between the preservation of /s/ and its total loss has influenced the preceding vowel by opening it (Navarro Tomás 1939, Hooper 1976, Salvador 1977, Mondéjar 1991, Alarcos 1958, 1983, etc.). This opening or laxing would acquire a distinctive value once the aspiration was also lost. That is, we would have the following evolution:2 [lifros] > [lifroh] > [lifroh] > [lifro]

In the most advanced stage in this evolution, the open character of the final vowel in the example in can be used to obtain a phonological opposition with the final vowel of the corresponding singular form.3 This standard explanation has a certain plausibility, but leaves a number of facts unexplained. First of all, given this
hypothetical development it is somewhat surprising that similar vocalic effects are not found in other areas where final /s/ is also weakened. The aspiration and deletion of syllable-final /s/ are phenomena that are widespread in Spanish, being found with greater or lesser intensity in numerous areas of both Spain and Latin America; but only in Eastern Andalusian do we find the vocalic effects that we are concerned with.4

Secondly, even though the origin of the Eastern Andalusian ‘vowel doubling’ has always been sought in the loss of /s/, as distinctive as the opening of vowels in plural forms (and other originally /s/-final forms) is the closing that takes place in singular forms, at least in some varieties. A number of authors have noted that final /e/ and /o/ in Eastern Andalusian are considerably higher than in Castilian Spanish, showing a clear tendency towards /i/ and /u/, respectively (cf. Rodríguez Castellano & Palacio 1948, Alonso et al. 1950, Alvar 1990). In the study by Alonso et al. (1950), vowels of intermediate quality between /o/ and /u/ are often transcribed in word-final position in singular items. Alvar (1990:235) remarks that “en Almería, Jaén y Granada, en el singular niño o pobres se pronuncian con una vocal final muy cerrada, con lo que el singular queda así marcado, mientras que los equivalentes regionales de niños o pobres tienen una vocal final muy abierta (por supuesto falta la -s), con lo que la marca de abertura sirve para indicar el plural.” [in Almería, Jaén and Granada, in the singular niño or pobres are pronounced with a very close final vowel, with which the singular is thus marked, whereas the regional equivalents of niños or pobres have a very open final vowel (of course, the -s is lacking), with which the openness feature indicates the plural.] Both in the singular AND in the plural do we thus find final vowels which are different from those that are common in Castilian Spanish. In absolute word-final position we may even find vowels that are phonetically [u], [i], instead of the corresponding mid vowels: [pέśu] pecho ‘chest’, [trάxhi] traje ‘suit’ (Zamora Vicente 1970:293).

For reasons that are not entirely clear to us (but we will come back to this point), both synchronic and diachronic accounts of the Andalusian vowel system have always taken the opening of vowels in the plural as the fact that needed an explanation, glossing over the raising that takes place in the singular, which, when addressed at all, is considered a mere secondary phenomenon. The standard assumption is that the primary phenomenon is the opening or laxing of vowels in the plural and that the raising of vowels in the singular is only a secondary development with the goal of maximizing the distinction (cf. Alarcos 1985:198). It is difficult to understand, however, how the change schematized above may have had as a consequence the extraordinary raising that vowels in a different environment have experienced. In our view, the traditional account of the historical origin of the Eastern Andalusian vowel system has concentrated on the wrong set of data.

The alternative hypothesis that we want to defend here is that the aspiration and loss of final /s/ has resulted in a singular/plural distinction based on vowel quality in Eastern Andalucía (and not in other areas), because an independent vowel contrast already existed in this area at the phonetic level. This phonetic contrast was prior to and independent from the weakening of final /s/. To be more specific, in this area of Andalusia, as in areas of the Peninsula of Leonesian speech, unstressed
word-final /e/ and /o/ were and are articulated as very close vowels, approaching the quality of the high vowels. This final raising is limited to the absolute word-final position and does not take place when the word ends in a consonant. This difference in vowel quality, which would predate the weakening of final consonants in Andalusia, has become evident once these have been lost.

Our contention is, thus, that in a contrast like the one that we find between the final vowels of [lifbrɔ] ‘book’ and [lifbro] ‘books’, the origin of the differentiation may very well have been a process of raising affecting word-final vowels (against the traditional hypothesis). Final vowel raising could have created a subphonemic contrast that was made distinctive once final consonants were lost. In this way we can explain why the loss of /s/ and other final consonants has not created a vowel contrast in other dialects where final vowels were not raised. Furthermore, as already mentioned, the raising of mid vowels in absolute word final position, far from being a rare phenomenon, is well-attested in other areas of the Iberian Peninsula. Final raising is in fact one of the distinctive features of the Leonese macro-dialect and is found in varieties from Asturias and Cantabria to Extremadura. In certain Asturian varieties, for instance, we find pairs such as liibru ‘book’ vs. liibros ‘books’, where the raising of the final vowel in the singular has created a contrast between singular /-u/ and plural /-os/, without the aspiration of /-s/ playing any role in the creation of the vowel contrast, since this consonant is not weakened in this area.

Our hypothesis is that the vowel contrast that we find in Eastern Andalusia originated in the same fashion as in the Leonese varieties, although the raising was normally less pronounced, and that the weakening of final consonants only served to reinforce a contrast that already existed at the phonetic level. As Penny (1994) suggests, rather than there being a historical raising process, the case might be that final /-u/ (but not before /s/) from spoken Latin was preserved in some areas of the Peninsula.

A phenomenon that is parasitic on the contrast in final syllables is vowel harmony. Several authors who have studied Eastern Andalusian varieties have noted that the open or close character of the final vowel is communicated, at least in certain cases, to the stressed vowel and post-tonic vowels (and, with less certainty to pre-tonic vowels as well). Thus, in examples such as [mɔnɔ] ‘monkey’ vs. [mɔnɔ] ‘monkeys’ and [pɛɾɔ] ‘dog’ vs. [pɛɾɔ] ‘dogs’ the stressed vowel shares the openness or closeness of the final vowel.

Here again, those authors who have analyzed this phenomenon have assumed that a harmonization process takes place in the plural, the singular being the unmarked case. This assumption has been adopted both by authors who have approached the facts from a structuralist perspective, such as Alarcos (1958, 1983) and by authors who have adopted some version of generative phonology, such as Hooper (1976) and Zubizarreta (1979). These authors have proposed a laxing harmony rule which would take place in plurals such as [mɔnɔ] ‘monkeys’ or [pɛɾɔ] ‘dogs’. The fact is, however, that the tonic vowel is equally affected by the final in the corresponding singular forms [mɔnɔ] ‘monkey’, [pɛɾɔ] ‘dog’. If in a word such as singular perro the tonic vowel is close this must undoubtedly be a consequence of harmonization, since in the position that it occupies, adjacent to /r/,
we would otherwise expect an open vowel, given the allophony rules of other Spanish dialects (Navarro Tomás 1977) and general tendencies in Romance languages and elsewhere (cf. Prieto 1993). There is no reason to assume either that synchronically a rule of harmonization operates only in the plural or that historically the original phenomenon is laxing harmony instead of raising harmony. Alonso et al. (1950:212) do in fact describe the existence of harmony both in the plural and in the singular: “Cuando en la palabra van varias vocales idénticas, la cerrazón del singular o la abertura del plural se extienden a toda la palabra, con extraordinaria diferenciación.” [when the word contains several identical vowels, the closeness of the singular or the openness of the plural extend to the whole word, with extraordinary differentiation]. Similarly, when they analyze stressed /e/ they note that “la e tónica se cierra notoriamente en los singulares” (212) [tonic /e/ closes remarkably in the singular] and a similar description is also found with respect to stressed /o/: “la vocal se abre en el plural; se cierra en el singular” (214) [the vowel opens in the plural and closes in the singular]. Vowel harmony is thus found in the singular as much as it is in the plural, and could have had its historical origin as easily in one case as in the other. If, against the received opinion, we assume that the origin of vowel harmony in Eastern Andalusian may have been the raising of non-final vowels under the influence of close or tense final vowels, the Andalusian phenomenon loses its exceptional character within the Romance family. Raising harmony is again well attested in Asturias and Cantabria (cf. Rodríguez Castellano 1952, Catalán 1953, Neira 1955, Blaylock 1965, Galmés de Fuentes 1966, Díaz Castañón 1966, Penny 1969, 1969b, 1970, 1978, Hualde 1989, among many others) as well as in large areas of Italy.

A fact that we have not addressed so far is the laxing and palatalization of /a/. As Alvar (1991:228-31) explains, the palatalization of the low vowel in the context of certain word-final consonants is a phenomenon which includes cases with different geographical and social distributions. The palatalization of /a/ is the fronting of the tongue in the articulation of this vowel. This fronting may be accompanied by a greater opening in some areas; but in other areas this vowel is articulated as a front lax vowel that overlaps with realizations of the mid front vowel /e/, cf. Alonso (1956), Llorente (1962). Different consonants may also produce different effects. This is a phenomenon that may be taken to be unrelated from a genetic point of view. In any case, the palatalization of the low vowel could be independent from the weakening of final consonants; not caused by it. We may compare plural forms such as cases in Asturian or Valencian Catalan where palatalization has occurred without weakening of the consonant. In fact, the change -as > -es in feminine plurals is well attested in Mozarabic (the Romance varieties formerly spoken in Muslim Spain) and well represented in Andalusian toponyms (Galmés de Fuentes 1983:307-319). Thus, whereas in our hypothesis the contrast between open and close mid vowels is due to a historical process of vowel raising in word final position, the contrast between palatalized and non-palatalized variants of the low vowel could be due to a process of palatalization of /a/ in final unstressed closed syllables, like the one that has produced plurals in /-es/ for feminine words ending in /-a/ in Asturian, Valencian Catalan and, importantly, in Andalusian Mozarabic. The two phenomena could then be genetically independent and also without causal connection with the loss of /s/ and other final consonants.
The role of the loss of /s/ has been to give distinctive value to differences that were subphonemic.

The three phenomena that we have examined: the raising of word final mid vowels (libro > libro), the palatalization of the low vowel in final syllables closed by certain consonants (libras > librâ(s)), and the metaphonic influence of the final vowels on other vowels in the word are phenomena that have parallels in varieties of the Leonese type. There is no need to assume a direct influence of Leonese varieties on Eastern Andalusian, since none of the three phenomena is particularly unusual and they could have originated independently in the two areas. Nevertheless, it is certainly possible that there might have been a direct link. The focal area from which the Eastern Andalusian vocalism originally spread may have been populated by Leonese speakers who carried their phonetic habits to the area. We may note that there are some other noticeable coincidences between Leonese and Eastern Andalusian. The characteristic change from /l/ to /t/ after a tautosyllabic consonant found in Leonese and Galician-Portuguese is also found in Andalusia, reaching its highest intensity in Eastern Andalusia, specifically in the southern part of the province of Granada (ALEA = Alvar 1961-73, map 1721). This feature is taken to be a clear sign of Leonese colonization by Zamora Vicente (1970:325). In towns of the Alpujarras mountains of Granada which have such Leonese or Galician-sounding names as Ferreira (ALEA, point Gr 409) and Lanteira (ALEA, point Gr 410) this phenomenon is systematic and we find examples such as [diábro] diablo ‘devil’, (ALEA, map 1530), [bránko] blanco (ALEA, map 1574), [tábra] tabla ‘board’ (ALEA, map 24), etc. In our recording from an informant from Lanjarón (to be discussed later) we find, for instance, [deθa gradâbre] desagradable ‘disagreeable’, [teθbre] terrible, [limpoθbre] imposible ‘impossible’, together with the hypercorrected [éblo] Ebro ‘name of a river’. Other, less striking, coincidences that can be mentioned are the velarization of nasals, the deletion of intervocalic /d/, the presence of a laryngeal fricative /h/ instead of velar /x/ and also from Latin /f/ (cf. Rodríguez Castellano 1954, Penny 1978, Alvar 1977) and, indeed, the aspiration of /s/, which also occurs in some areas of the Leonese macrodialect (Penny 1978). These coincidences may point to a common origin also for the vowel phenomena with which we are concerned. Be that as it may, it is not necessary to assume a direct Leonese influence to maintain our hypothesis on the evolution of the Eastern Andalusian vocalism.

To sum up so far, the standard assumption on the origin of the contrast between close and open vowels in Eastern Andalusian has been that this phenomenon is a result of the opening of vowels in contact with [h], from original /s/. This assumption leaves a number of problems unresolved. We have attempted to address these problems by starting from the assumption that historically the first and main event was the raising of final mid vowels, as in Leonese. Final raising is independent from the loss of final /s/. When final /s/ was weakened, a vowel contrast in final position arose only in those areas that had an earlier process of final raising (or had preserved final high vowels).

Now we may go back to the question of why researchers have traditionally focused on the lowering that takes place in plurals and other forms with a “latent” final consonant and not on the raising of vowels in absolute word-final position.
We believe that a reason for this might be that vowel raising could very well be a receding phenomenon, for which there is some evidence, as we will see.

The most extensive acoustic study available of the Eastern Andalusian vowels is that in Sanders (1994). In this study a large corpus of utterances from three male college students from Granada was analyzed instrumentally. This study found a quite consistent alternation in vowel quality between singular and plural forms. The alternation affected the mid vowels in all positions, pretonic, tonic and final; and also the low vowel in word-final position. The average values for the mid vowels for all three informants are presented in Table 1.

If we compare these values with the average formant values for male speakers of standard Spanish given in Martínez Celadrán (1984:293-4), reproduced in the left columns in Table 2, or Quilis and Esgueva (1983:244), right columns in Table 2, we note that the Andalusian plural mid vowels are more open, but there is no evidence for any raising in the singular. The formant values of Eastern Andalusian singular vowels in Sanders (1994) are very similar to the standard Spanish values. These data thus appear difficult to reconcile with the hypothesis that we are defending in this paper.

We have made an hour-long recording of the oral poetry of AR, an elderly illiterate speaker from Lanjarón, in the Alpujarras mountains near Granada. The comparison of this corpus, which we have begun to analyze, with the results obtained in Sanders (1994) is rather revealing. In both cases there is a fairly systematic distinction in vowel quality depending on whether the word is singular or plural (i.e. whether it ends in a vowel or in a latent consonant), but the nature of the distinction is not the same. Both the college students in Sanders’ (1994) study and AR show generally lower variants in plural than in nonplural forms. In addition, however, AR’s final singular /e/ and /o/ (especially the latter) are significantly raised, which is not the case for the college students.

AR’s final /e/ and /o/ often approach the quality of high vowels. Compare the examples in Table 3, where first and second formant values for the final vowel are given, with those for the corresponding vowels in Tables 1 and 2.

The acoustic impression given by these examples in that of a very close mid or a high vowel. The raising is not always as extreme as in these selected examples, but when a comparison is made with the quality of final vowels in plural forms the existence of noticeable differences is apparent, especially in the case of final /o/, as mentioned. Incidentally, in our hour-long recording of AR, etymological syllable or word-final -s is never pronounced as such and the aspirated allophone [-h] has a very low incidence in word-final position. By far the most common pronunciation presents no trace of a final consonant. This is consistent with the situation reported in other studies of Eastern Andalusian. From AR’s text, we have scored 37 tokens of /o/ in absolute word-final position and 19 tokens in the final syllable of plurals (where a ‘latent’, although seldom realized final consonant can be postulated). The average formant values for final /o/ in final unstressed syllable for A.R. are the following: /-o#/10 F1 413.38 / F2 1163.71 (average of 37 tokens); plural F1 544.46 / F2 1155.20 (average of 19 tokens). A scatter graph where all scored tokens are represented is given in Graph 1.

As can be seen from Graph 1, in the plural final /-o/ presents, in general, higher values for the first formant, i.e. it is a lower vowel. All instances of very
raised /o/ (on the top part of the figure) correspond to non-plural items, whereas plural tokens concentrate on the bottom part of the figure. Graph 1 for our elderly speaker AR, can be compared with Graph 2, reproduced from Sanders (1994), which represents the values for final /o/ for AMB, one of the subjects in this study.

Both graphs are similar in the relative distribution of plural and non-plural tokens in the vowel space. But non-plural tokens for AMB are never as raised as for AR. The extreme raising that non-plural final /o/ can undergo in AR’s speech is absent from AMB’s data (notice that in Graph 2, no token has a F1 value lower than about 225 Hz). This is not a feature of his speech. The difference is apparent to a trained ear.

In non-plurals, AR shows a very pronounced raising of final mid vowels, which often approach high vowels in their quality. The younger speakers, on the other hand, do not present any raising (when compared with the values that are normal in Spanish). How are we to interpret these facts? According to our hypothesis the raising of final vowels is the element that creates the condition for ‘vowel doubling’. The data from AR, our elderly informant, show a great measure of raising and are consistent with our hypothesis. This would represent a more conservative dialectal pronunciation. In recent times, and for sociolinguistic reasons, there would have been a tendency to reduce the raising in the singular and, in order to maintain the opposition, to also increase the lowering or laxing in the plural. This is what has happened in the speech of the college students in Sanders’ (1994) study.

The raising of final vowels is nowadays highly stigmatized in Spain and is viewed as a trait of a rural, uneducated background (cf. Holmquist 1988). Zamora Vicente (1970:293) states that in Granada the raising of final mid vowels to high is characteristic of uncultured folk (“todas las clases poco cultas”). Moya Corral (1979:27) implies that raised final vowels are receding in the variety of Jaén, when he indicates that a very close [o] (which he sometimes transcribes with a superscript [u] over the [o]) in word-final position in the singular is especially frequent in the speech of his older informants. It does not seem unreasonable to suggest that the association of final raising with illiteracy has been a factor leading to the progressive dismissal of this linguistic feature.

On the other hand, laxing does not convey any negative associations. The different social valuation of these phenomena would have induced the reduction of raising; the opposition being now preserved in the speech of these speakers solely by means of a greater lowering of the lower mid vowels.

Whereas we cannot claim to have presented conclusive evidence that final raising is a receding feature, which presumably once had a more general distribution, the evidence that we do have is rather suggestive. The striking difference between the pronunciation of our elderly informant and that of the college students studied in Sanders (1994), together with the repeated reference in the literature to final raising as characteristic of older and/or uneducated speakers, are facts that point in this direction.

The final outcome would be that the raising of unstressed mid vowels in absolute final position which, in our hypothesis, was the original trigger of a sg/pl distinction based on vowel quality, appears to be on its way out, given its negative
connotations; whereas, lowering or laxing in the plural, a secondary development which does not carry the same social stigma, is establishing itself as the significant feature in the opposition.

TABLES AND GRAPHS

Table 1: Average vowel formant values in Sanders (1994:168)

<table>
<thead>
<tr>
<th></th>
<th>F1</th>
<th>F2</th>
<th>F1</th>
<th>F2</th>
</tr>
</thead>
<tbody>
<tr>
<td>singular</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/e/</td>
<td>pretonic</td>
<td>479.75</td>
<td>1830.53</td>
<td>588.66</td>
</tr>
<tr>
<td></td>
<td>tonic</td>
<td>476.07</td>
<td>1937.76</td>
<td>579.26</td>
</tr>
<tr>
<td></td>
<td>final</td>
<td>489.15</td>
<td>1867.41</td>
<td>568.06</td>
</tr>
<tr>
<td>/o/</td>
<td>pretonic</td>
<td>482.44</td>
<td>1077.49</td>
<td>578.97</td>
</tr>
<tr>
<td></td>
<td>tonic</td>
<td>504.96</td>
<td>1119.66</td>
<td>602.31</td>
</tr>
<tr>
<td></td>
<td>final</td>
<td>497.03</td>
<td>1099.99</td>
<td>564.49</td>
</tr>
</tbody>
</table>

Table 2: Average vowel formant values for standard Spanish

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>fi/</td>
<td>tonic</td>
<td>344</td>
<td>2202</td>
<td>268.28</td>
</tr>
<tr>
<td></td>
<td>atonic</td>
<td>347</td>
<td>2193</td>
<td>260.68</td>
</tr>
<tr>
<td>/e/</td>
<td>tonic</td>
<td>451</td>
<td>1921</td>
<td>449.71</td>
</tr>
<tr>
<td></td>
<td>atonic</td>
<td>451</td>
<td>1847</td>
<td>454.96</td>
</tr>
<tr>
<td>/a/</td>
<td>tonic</td>
<td>710</td>
<td>1230</td>
<td>665.68</td>
</tr>
<tr>
<td></td>
<td>atonic</td>
<td>650</td>
<td>1300</td>
<td>648.84</td>
</tr>
<tr>
<td>/o/</td>
<td>tonic</td>
<td>506</td>
<td>1037</td>
<td>475.8</td>
</tr>
<tr>
<td></td>
<td>atonic</td>
<td>496</td>
<td>1007</td>
<td>473.3</td>
</tr>
<tr>
<td>/u/</td>
<td>tonic</td>
<td>361</td>
<td>954</td>
<td>291.09</td>
</tr>
<tr>
<td></td>
<td>atonic</td>
<td>385</td>
<td>1008</td>
<td>283.5</td>
</tr>
</tbody>
</table>

Table 3: AR’s word-final mid vowels (selected examples)

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>noche [-e]</td>
<td>‘night’</td>
<td>F1 331.17</td>
<td>F2 2150.05</td>
</tr>
<tr>
<td>leche [-e]</td>
<td>‘milk’</td>
<td>F1 229.27</td>
<td>F2 2099.10</td>
</tr>
<tr>
<td>tiempo [-o]</td>
<td>‘time’</td>
<td>F1 387.21</td>
<td>F2 820.28</td>
</tr>
<tr>
<td>dinero [-o]</td>
<td>‘money’</td>
<td>F1 310.79</td>
<td>F2 1039.36</td>
</tr>
<tr>
<td>cuando [-o]</td>
<td>‘when’</td>
<td>F1 315.88</td>
<td>F2 1217.68</td>
</tr>
<tr>
<td>fuego [-o]</td>
<td>‘fire’</td>
<td>F1 361.74</td>
<td>F2 1192.21</td>
</tr>
</tbody>
</table>
Graph 1: final /o/ for AR (elderly speaker)

Graph 2: final /o/ for AMB (college student, from Sanders 1994)

NOTES

1 As mentioned, these are contrasting surface phones. Framing these vowel oppositions in terms of 'doubling of vowel phonemes', as some authors have done starting with Navarro Tomás (1939) and including Salvador (1977, 1987), is clearly unwarranted (cf. Zubizarreta 1979, Alarcos 1983, Cerdà Massó 1984, López Morales 1984, among others).

2 The influential Spanish phonetician Navarro Tomás (1957) recognizes the existence of open and close allophones for the mid vowels of standard Castilian Spanish. But the situation would be very different from that of Andalusian. To begin with, the allophones of Castilian, which would be conditioned by the phonetic environment, would not have the degree of
diferenciación que los cronistas incaicos han realizado en el periodo que nos ocupa. Segundo, otro dato no menos interesante es que en el siglo XIX se introduces un nuevo sistema de escritura, el que ya se conoce como “chancillería”, que permite una mejor comunicación entre los diferentes grupos lingüísticos.


Martínez Melgar, Antonia. 1986. Estudio experimental sobre un muestreo de vocalismo andaluz. Estudios de Fonética experimental (Univ. de Barcelona) 2.195-248


Yokuts as a target language in a shift from Miwok

Marvin Kramer
University of California, Berkeley

1. Introduction

The many lexical and grammatical resemblances between the Utian (Miwokan and Costanoan) and Yokutsan languages have led to various claims of relationship. These claims at present, however, are generally not accepted. I propose that the resemblances are due instead to contact, having occurred during the three millennia these languages have been immediate neighbors. In particular, while lexical resemblances are borrowings, resemblances in morphology are the result of language shift with interference. The interference is in Yokutsan, resulting from a shift by a group of Proto-Eastern Miwok speakers to Proto-Yokuts.

An outstanding feature of Utian, particularly Eastern Miwokan, and Yokutsan is templatic morphology. The templatic morphology of Utian and Yokutsan is extremely rare in the languages of the world, the most notable other similar example being Semitic. Finding it therefore in two neighboring languages generally considered to be unrelated suggests contact rather than chance. It is fortuitous that a rare and easily identified feature is a candidate for interference. It is fortuitous as well that there is an identifiable archaeological scenario which makes a plausible case for a social setting compatible with language shift.

2. Archaeological history of the Miwok and the Yokuts

Archaeological continuities through time suggest that Utian speakers arrived in the Sacramento Delta area circa 2400 BCE (Moratto 1984). The Yokuts appear in the Delta and northern San Joaquin Valley between 1500 and 1000 BCE. The Utian and Yokutsan people appear to have arrived in California at different times and from different directions, the situation for other Penutian families as well (Whistler 1977). A possible origin of the Yokuts is southern Oregon. The Utians may have been refugees from the drying Great Basin Pluvial Lakes. They were able to gain control of the rich San Francisco Bay marshlands, and eventually expand, splitting into Miwokan and Costanoan. By 500 BCE the Miwoks occupied the entire Delta and either blended with the Yokuts, or replaced them, forcing them south and east into the Sierra Nevada foothills.

About 400 AD a catastrophic climatological change altered cultural patterns throughout California and the Great Basin. A cool moist climate gave way abruptly to a period of intense warming and drought, which lasted until 1300 AD. The impact of the environmental change was greatest near ecotones, or biological 'edge' communities, and least deep within a life zone (Moratto 1978). The Delta was climatologically 'complacent', with minimal impact. Trade ended with the foothills, but continued with the coast. The basic social organization was not disrupted. The foothills, on the other hand, were climatologically 'sensitive', and suffered a considerable impact. The foothills were basically depopulated. Nuclear villages broke up and populations dispersed into small groups near available water. Political organization deteriorated. Trade ceased. Violence increased, as 50% of male burials show evidence of violent death (M 1978). Various Yokuts groups may have assimilated with neighbors. Linguistically, this period has been called the Yokuts 'bottleneck' (Whistler 1978 in M 1984); the linguistic diversity of the previous millennium disappeared and a single variety emerged, namely Proto-Yokuts. Interesting because the date corresponds, Proto-Yokuts has a lexicostatistical time depth claim of 15 centuries (Levy 1978 in M 1984).
It appears that highly stressed Yokuts communities were forced to compete with the Miwok for resources in the marshlands, where most Yokuts groups were linguistically assimilated into the presumably more prestigious Miwok speech community. But one group, the Proto-Yokuts speech community, was able to prevail linguistically and induce shift in a community of Proto-Eastern Miwok, or perhaps the later Proto-Sierra Miwok speakers. Callaghan (1980) estimates Proto-Sierra Miwok at roughly 300 - 900 AD. Such a shift would no doubt take place under duress and in a relatively short period of time, and probably also involve a proportionately large number of Miwok speakers. This accounts for the interference. This Yokuts variety would naturally be associated with resource availability, and so would be a logical candidate to survive the bottleneck.

3. Background

The Penutian Hypothesis, as originally proposed by Dixon & Kroeber (1919), included Wintun, Maidun, Costanoan, Miwokan and Yokutsan. It was subsequently expanded by Sapir, Swadesh, and Greenberg (Golla 1994) to include many languages of the western hemisphere. However, only Costanoan and Miwokan can now be shown to be related by the reconstruction of Proto-Utian (C 1983). Wintun, Maidun, and Yokutsan are generally considered to be unrelated to each other or to Utian (G 1994). Yet there are many phonological, morphological and syntactic resemblances between Utian (especially Eastern Miwokan) and Yokutsan. There are lexical resemblant sets as well (C 1994d), but no cognates have been proposed.

Structural resemblances and a lack or scarcity of vocabulary from the source language are characteristic of language shift with interference (Thomason & Kaufman 1988). Language shift is the replacement of the native language in a speech community, the source language, with another, the target language. Interference in the target language is the presence of source language features retained by the shifting population. Often interference is reinterpreted as a standard language feature, and is acquired by native speakers of the target language. Source language lexicon tends not to be retained, since a shifting population is motivated to learn the target language, and thereby at the very least to rel exify. Shifting speakers often do learn the target language quite well, but any resulting interference is a function of the relative size and social status of the shifting population, the speed of the shift, and access to the target language.

4. Phonological resemblances

Resemblances between Eastern Miwokan and Yokutsan which could be interpreted as interference include epenthesis and epenthetic vowel harmony, ghost segments, stem templates, and template conditioning suffixes. Interference, rather than borrowing, could explain the similarities in structure and phonological rules between these unrelated or very distantly related languages. Retained interference features would not be a complication of target language grammar for source language speakers, since they would already control those features. Structure and rule borrowing, on the other hand, is an unlikely explanation, because new constraints would complicate the grammar for borrowing speakers. In any case we would expect the borrowing of phonological rules to be accompanied by wholesale lexical borrowing. Through the borrowing of words, a familiarity with the source language develops, as does, over time, enough bilingualism to control the borrowed features (T & K 1988). There is no indication of any single period of intense lexical borrowing between Miwokan and Yokutsan, although there are examples, such as (1) and (2) below, suggesting retention or borrowing subsequent to shift. It appears now that there had been
considerable borrowing among all the California languages (Shipley 1980), which was mistaken for genetic relationship, in part leading to the establishment of the Penutian Hypothesis in the first place.

An explanation of relationship for the resemblances between Eastern Miwok and Yokutsan is unlikely, since it would assume that very particular rules and structures, some more characteristic of Eastern Miwok than of Utian, would survive in Yokutsan where cognates did not.

5. Templates in Southern Sierra Miwok and Yawelmani Yokuts

The Eastern Miwokan and Yokutsan languages have come to the attention of phonologists because of their templatic morphology phenomenon involving an analysis of blockar representation (Smith 1984, Archangeli 1983, et al.). The templates are representations of roots or stems in terms of consonants and vowels (the CV skeleton), or of prosodic structure (µ, σ, F), without reference to the identity of the consonants or vowels. Nouns and verbs in Eastern Miwokan (Broadbent 1960 for Southern Sierra Miwok) and verbs in Yokutsan (Newman 1944) occur in a number of templates. The Yawelman form in the example (1) appears to be a borrowing from Miwokan;

(1) SSMiwok CVCVCV- ʰiwaat- CVVCVC- ʰiwaat- 'to run'
    Yawelm. CVCVVC- ʰiwiit- CVCCC- ʰiwiit- 'to walk'

Both language families employ 1) various templatic shapes with conditioning suffixes and 2) default templates with nonconditioning suffixes. Suffixes that determine the template are conditioning. Nonconditioning suffixes do not specify a particular template, so the stem occurs in its default template;

(2) SSM Conditioning suffix -mₜ-, 'absent', requires CVCCV-
    ʰiwtₜ-mₜ- → ʰiwtₜamheettii 'let's run away' B.513
    Noncond. sx -mₜ-, 'reciprocal imp.', default in this case is CVCVVC-
    ʰiwaat-mₜ- → ʰiwaatimheettii 'let's run a race' B.514
    Ym Cond. sx-(?)in'ay, 'contemporaneous gerundial', requires CVCC-
    ʰiwtₜ-(?)in'ay → ʰiwtₜin'ay 'while walking' N.19:5
    Noncond. sx -iin, 'future', default in this case is CVCVVC-
    ʰiwiit-iin → ʰiweeten 'will walk' N.18:14

Yokutsan resembles Utian in the use of templates, but more closely resembles Eastern Miwokan in its use of default templates. In Lake Miwok, a Western Miwokan language, the use of default templates is not as robust, in that thematic verbal suffixes are conditioning and post thematic suffixes are not (C 1963). The Primary Stem in Mutsun Costanoan (Okrand 1977) suggests the default template in Miwokan, from which the others may be predicted, is Utian in origin.

Yokutsan templates resemble characteristically Eastern Miwokan templates, particularly CVCVVC-, which does not exist as a template in Lake Miwok (C 1994a). Yet resemblances in Yokutsan for the Proto-Miwok Light Stem C₁VC₂VC₃-, Geminate Stem C₁VC₂C₂VC₃-, and Cluster Stem C₁VC₂C₃V- (C 1992) are conspicuously absent. These stems, commonly referred to as Stems 2, 3, and 4 (Freeland 1951, Broadbent 1960, et al.) are hallmark Utian templates, existing in some form in all the Utian languages. An explanation for their absence in Yokutsan is given in 11.

Below is a general idea of the two templatic systems, using Southern Sierra Miwok (SSM) and Yawelmani (Ym) as representatives. These two languages have the advantage of not being immediate neighbors, and of having the most complete grammars. In addition, SSM is somewhat more similar phonemically (11) and syntactically to Yokutsan than other Eastern Miwokan languages. In Proto-Eastern Miwok, a southern dialect would be the most likely to encounter Yokutsan in 400
AD. For this reason the Smith (1984) table of templates (3) will be used, with fictitious data due to gaps in the attested data, even though Callaghan (1994c) has assembled a Central Sierra Miwok table with real data. Not included in the Smith table are irregular forms and forms where the suffix specifies no C₂, the Simplex Grades in Callaghan (1986). The boldface underlined entries in the SSM table are those that resemble the Ym;

(3) **Southern Sierra Miwok** (Smith 1984)

<table>
<thead>
<tr>
<th>Stem 2</th>
<th>Stem 3</th>
<th>Stem 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>polaat polat polat pollat polta</td>
<td>poolat polaat poltaa</td>
</tr>
<tr>
<td>II</td>
<td>kelit kelit kelit kelit kelit</td>
<td>kelit kelit kelit kelit</td>
</tr>
<tr>
<td>III</td>
<td>halih halih halih halih halih</td>
<td>haalih haalih halih</td>
</tr>
</tbody>
</table>

(4) **Yawelmani Yokuts** (Archangeli 1983)

<table>
<thead>
<tr>
<th>Stem 1</th>
<th>Stem 2</th>
<th>Stem 3</th>
<th>Stem 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>?amc'</td>
<td>?amc'</td>
<td>?amaac' approach</td>
</tr>
<tr>
<td>Ia</td>
<td>caaw</td>
<td>caaw</td>
<td>caaw caaw shout</td>
</tr>
<tr>
<td>II</td>
<td>diiyl</td>
<td>diiyl</td>
<td>diiyl diiyl guard</td>
</tr>
<tr>
<td>III</td>
<td>c'uum</td>
<td>c'uum</td>
<td>c'uum c'uumu destroy</td>
</tr>
<tr>
<td>IIIa</td>
<td>biniit</td>
<td>biniit</td>
<td>biniit biniit ask</td>
</tr>
<tr>
<td>IIIa</td>
<td>hoyo hoyo hoyo hoyo name</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In the Smith table, SSM uses the default vowel ♂ and consonant ? to fill in deficient templates, or templates lacking a V₂ or C₃. The default consonant is not always ?, it may be one of a set of default consonants (11) determined by the suffix. The default short vowel ♂ is also the epenthetic vowel (see 6), and will be supplied by epenthesis rules according to the template (6), which is also determined by the suffix. The boldface SSM forms represented without the default consonant and short vowel are strikingly similar to the Ym templates;

(5)    halih    haalih halih
       tiil    tiil tiil

Of the seven SSM template types, only I and II have three consonants and two vowels. III, IIIa, IV and IVa have two consonants, and IIa, IIIa, and IVa have one vowel. Historically the most common template probably involved two consonants and one vowel, as Callaghan's (1988, 1994b) Proto-Miwok and Proto-Eastern Miwok reconstructions show roughly half *CVC- or *CV₁CV₁-. These templates could have been expanded, as CVC- and CVCV- templates in the modern language are, by particular suffixes or default segments, see (11) and 9, which would supply the third consonant. The second vowel space could have been filled by vowel spread, as nearly one third of the roots in the modern languages have identical first and second vowels (B 1960, C 1987, et al.).

In a shift from PEM to PY, the only part of the template system to be retained in PY applied to those PY verb roots which happened to fit the common PEM roots with a single vowel, or with vowel spread to the second syllable. This would explain why only bi- and trilateral Ym verb roots where V₁ = V₂ are templatic. A significant number of Ym verbs and most nouns are simply suffixed without participating in templatic morphology. In SSM, all verbs and many nouns fit the templates (even borrowings from English and Spanish). This may have been true historically as well, although verb templates can often be reconstructed further.
back than noun templates (C 1994c). The introduction of templates into PY would undoubtedly add to its structural complexity. But selecting only a part of the templatic system would be a simplification on the part of PEM speakers, since simplification occurs in the source language, the language the shifting speakers control.

6. Default, epenthetic and ghost segments

Vowel epenthesis rules in SSM are the same in Ym. This contrasts with LM which does not have vowel epenthesis. The SSM default vowel for template expansion is the same as the epenthetic vowel, and harmonizes L > R by [+round]; i : u. The Ym epenthetic vowel is i : u, harmonizing L > R with [+round] a factor.

Neither language allows syllable initial or syllable final consonant clusters. In both languages consonants parse, by vowel epenthesis, rather than delete when in conflict with syllabification. Vowel epenthesis in both SSM and Ym is sensitive to the right edge of the template (for Ym see Zoll 1994). Considering consonants beyond the right edge of the template, the rule is;

(6) **Epenthesis**: In a series of extra-templatic consonants, the first is an onset, and subsequent consonants pair off as coda-onset;

underlying Template CCCC
surface Template C&C.C&C

Examples are given, though the template will be further defined in 8. The template is underlined;

(7) SSM hiwaat&m-hi- 'run a race' < hiwaat- + -mhi 'recip' B.514
    Ym hogoo.nlw.sel 'float' < hogoon- + -wsii1 'recip/rflx' Z.19

There are ghost consonants in both SSM and Ym. Ghost consonants surface only if they do not force epenthesis (Archangeli 1991);

(8) SSM ?ipii-pii-wiye 'uncles' < ?ipii-pii- + (?)wiye- 'pl' B.750
    SSM ?it?a-yya-wiye- 'sisters' < ?it?a-yya- + (?)wiye- 'pl' B.750
    Ym mazaa- -(h)niil > maxahnel 'procure-cons.pas.adj' A.52
    Ym yawaa.l- -(h)niil > yawalnel 'follow' (note NSM yoowu-’go’)

In Ym, suffixes with initial ghost consonants in effect often supply the C3 for biliteral roots. SSM, on the other hand, employs various default consonants to supply the third consonant in deficient templates (11). Like the ghosts in Ym and SSM, these are determined by the suffix;

(9) SSM liwa-?-pe- 'speaker' < liwawa- 'speak' + -pe- 'ag.' B.739
    SSM moli-y-ate- 'umbrella' < moli- 'shade' + -ate- 'inst. ag.' B.755

Particular suffixes also expand deficient templates in SSM, whereas most do not (Sloan 1991, et al.). The template expanding suffixes require CVC- (one vowel, see 11) or CVCCV-(two vowels). An expanded biliteral template is interpreted as a trilateral template. For example;

(10) SSM weel- 'fetch': weel-ki- 'fetch', weel-h- 'seek'
    SSM kaal- 'kick': kal-ya- 'kick all over', kalaa-ŋ- 'dance'

Epenthesis serves as a test for template expansion in SSM, since epenthesis is sensitive to the right edge of the template (see 9).

Ghost consonants in SSM differ from default consonants in that, lacking the epenthesis test for template expansion, they are not reinterpreted as part of the template. They too differ in not being restricted to following templates.

Table (11) compares the SSM default, CVC-(one vowel) template expanding and ghost consonants with the Ym ghost consonants. Interestingly, consonants affecting vowel quality match these sets;
(11) SSM default consonants for templates; -w -y -h -? 
SSM CVC-template expanding sfxs; -wa -la -ya -ki/-ku -h 
SSM ghost consonants; w l y k -? 
SSM cons. affecting vowel quality; w y k h -? 
Yawelmani ghost consonants; m l h -? 
Ym cons. affecting weak-strong ablaut; h -? 

By regular sound correspondence, the SSM suffix h is ̂ in other Eastern Miwokan languages (F 59.II.1). In this respect, SSM, more than other Eastern Miwokan languages, resembles the Yokutsan languages.

PEM CVC-template expanding suffixes could have been grammaticalized into default C₃ status by occurring frequently with particular suffixes, possibly as bipartite. SSM has several bipartite suffixes. One such suffix is -ee-ni-, 'discontinuous (in time) iterative', B.519. This suffix, as a further example, specifies ̂ as the default C₃, except for templates whose C₃ is h, in which case the h is replaced by y. Another suffix, -yee-ni, 'discontinuous (in space) iterative', B.521, always follows the vowel of Stem 4. It is not concerned with h;

(12) SSM tînį'-yee-ni- 'heel showing' < tînį'-h 'heel' B.519
SSM ̂ak-h-i-ye-ni- 'to bob up and down' B.521

I propose that the semantic difference between these two suffixes lies in the template, as suffixes often follow different templates with a difference of meaning. Thus an earlier suffix *-ee-ni- occurring frequently after the template expanding suffix *-ya > *-y /_/V would replace other expanding suffixes such as *-h (which also expands CVCVV-). The expanding suffix *-y would not occur after triliteral stems, and so could be reinterpreted as a default C₃ consonant. For templates other than Cluster (Stem 4), the syllable onset for *-ee-ni- would be either the stem C₃ or *-y. In this context *-y could be interpreted as a ghost: surfacing to supply an onset, but otherwise not, to avoid a complex onset or epenthesis. The interpretation of *-y as a ghost must have been fairly stable when the semantic distinction (time/space) brought in Stem 4, since-y was generally replaced as a default by the more common ̂. This suffix occurs again as -lee-ni-, 'discontinuous iterative', B.522, also after Stem 4, using the expanding suffix *-la > *-l /_/V.

This example implies that SSM ghost consonants developed from template expanding suffixes by way of default consonants, a view supported by the similarities in table (11). There is no such development implied for the ghost consonants in Ym, since there are no expanding suffixes or default consonants. This, I propose, is because ghost consonants transferred into PY as interference, having already developed in PM (there are ghost consonants in LM). This view is also supported by the similarities in table (11).

The most common default consonants in SSM are h and ̂. Ym h and ̂ do not occur as root final consonants. In shifting to PY, PEM speakers may have reinterpreted PY root final h and ̂ as default consonants of the suffix, thus eliminating them in that position in the PY root. h and ̂ are extremely rare as root final consonants in SSM as well, probably for the same reason.

7. Suffixes in SSM

SSM has non-conditioning and conditioning suffixes. Non-conditioning are concatenating to the default template, but conditioning suffixes put requirements on the preceding template. In addition, suffixes of both types may put a requirement on the preceding syllable, whether part of a template or not. The most common requirement is that the preceding syllable be heavy, namely σᵅμ.
Some suffixes do this by gerninating the preceding consonant (13a), or vowel (13b). Some do not specify, and will gerninate only when necessary (13c). Some require that the preceding syllable be closed by the suffix. Of these, some require that it be closed with the melody of the suffix (gernimate onset (13d), and some supply a different consonant (13e). These requirements will force epenthesis if necessary. There is another category of suffixes that is not specific as to the type of the heavy syllable, created by gerninating their onset if necessary (13f). They will not force epenthesis. And, parallel to the epenthessizing suffixes, some will supply a different consonant, avoiding epenthesis(13g). These are the ghost consonants. Suffixes not affecting the preceding syllable also pair ghost (13h) and regular (13i) consonants for their onsets. In (13) the conditioned syllable requirement is underlined. The ghost segments are in parentheses. The reference is to Broadbent (1960) with her representations;

| (13) | | | | |
|-----|-----|-----|-----|
| (a) CV-C | Cu-ma passive participle | B.763 | -u-ma- |
| (b) CV-V | ni | augmentative | B.761 | -ni |
| (c) CV-μ(1) | na | 'animal with a big . . .' | B.762 | -Hna |
| (d) CV-p | pa | dimin., distributive | B.739 | -ppa |
| (e) CV-i | ki | diminutive | B.743 | -ti |
| (f) CV-μ(2) | pute kind, species | B.740 | -PHute |
| (g) CV-μ(2) | wuye plural | B.750 | -wiye/ -wiye |
| (h) (y)aa | andative | B.505 | -aa / -yaa |
| (i) poksu reflexive | B.527 | -poksu |

The underlined syllables may be summarized as CV, CVV, or CVμ, and might be considered templates conditioned by the following suffix. The syllable types, CVC, CVV, or CVμ, are also the bimoraic syllable types in the stem templates (Crowhurst 1992, Sloan 1991, et al.). The suffix conditioned stem templates and the conditions put on syllables preceding some suffixes are obviously part of the same system. Examples of SSM conditioned syllables are underlined;

| (14) | | | | |
|-----|-----|-----|-----|
| (a) ?elut-ta- | 'a float' | ?elut- | ?elut- 'to float' | + :a 'ag.' | B.753 |
| (b) pi?ca-li- | 'jackrabbit' | pi?ca- | 'cottontail' | + :li | B.753 |
| (c) lawwaati-j | meti- 'several snakes' | lawwaati | +(:)meti(:) | B.760 |
| (d) hiwaat-i-tho- | 'running' | hiwaat- | -tho 'gerund' | B.742 |
| (e) ?ammu-n | -g to get hurt | ?ammu- | r-(n)me- 'pass.' | B.525 |
| (f) komta+poksu- | 'to hit oneself' | komta- | +poksu 'rflx' | B.527 |

I propose that SSM stem templates are like these conditioned syllables; the template includes one of the heavy syllable types, and extends up to but does not include the following consonant, which serves as an onset for the following syllable, forcing epenthesis if necessary. Stem templates also include monomoraic syllables, such as the syllable that happens to precede the suffix in (14). These syllables are also followed by an onset, since if they were followed by a consonant cluster, the first C would become a coda and the syllable would no longer be monomoraic. Thus for both monomoraic and bimoraic templatic syllables, the template and syllable right edge coincide.

Suffixes in Ym only condition stem templates, they put no conditions on other syllables. In this regard, Ym templates stand out from the rest of the system. They need special phonological treatment (different direction of association, Archangeli 1991; ALIGN-Template constraint, Zoll 1994). It is as if templates had been introduced from an alien system. In the proposed shift to PY, stem templates would be retained as a simplification of a more general system in PEM.
8. SSM templates and template expansion

Trochaic templates (σμμ,σμ) exclude the C after the heavy syllable, which must be the onset of the following light syllable. Trilliteral roots interpret the light syllable as templatic because it is part of the lexical representation (15a). Monosyllabic (σμμ) templates may be reinterpreted as bisyllabic (σμμ,σμ) templates if the second syllable qualifies. To qualify, it must be one of the template expanding suffixes (15b) or a default consonant (15c);

(15)(a) part of the lexical repr.;

\[
\begin{align*}
\text{hiw-} & \text{ wa} \text{-t-} \to \text{ hiw-} \text{ wa} \text{-t-} \text{ 'to run'} \\
\text{hi} & \text{ wa} \text{-t-} \to \text{ hi} \text{ wa} \text{-t-} \\
\text{wel-} & \text{ ki-} \to \text{ wel-} \text{ li} \text{-k-} \text{ 'to fetch'} \\
\text{wee} & \text{ li-ki-} \to \text{ wee} \text{ li} \text{-k-} \\
\text{tis} & \text{s-?} \to \text{ tis} \text{s?} \text{ - } \text{ 'hand'} \\
\text{tii} & \text{s-?} \to \text{ tii} \text{s?} \\
\end{align*}
\]

(b) one of the temp. exp. sfxs.;

(c) a default C3 cons.;

(d) not expanded;

\[
\begin{align*}
\text{yoh-meh-} & \to \text{ yoh-meh- } \text{ 'kill, indef.'} \\
\text{lambic templates (σμσμμ)} & \text{ exclude the C after the light syllable, which must be an onset of the following heavy syllable. Trilliteral roots (16a) interpret the heavy syllable as templatic as in (15a). Monosyllabic (σμ) templates may be reinterpreted as two syllable (σμσμμ) surface templates if the second syllable qualifies, under the conditions of (15b,c);}
\end{align*}
\]

(16) (a) lexical representation;

\[
\begin{align*}
\text{hi} & \text{ waa} \text{-t-} \to \text{ hi} \text{ waa} \text{-t-} \text{ 'to run'} \\
\text{ka} & \text{ laa-} \text{-} \eta \to \text{ ka} \text{ laa-} \eta \text{ 'to dance'} \\
\text{ha} & \text{ lee-} ? \to \text{ ha} \text{ lee-} ? \text{ 'wild animal'} \\
\end{align*}
\]

(b) template exp. sfx.;

(c) default C3;

(d) not expanded;

\[
\begin{align*}
\text{yoh-meh-} & \to \text{ yoh-meh- } \text{ 'kill, def.'} \\
\text{vo} & \text{ h-meh-} \to \text{ vo} \text{ h} \text{umeh-} \\
\end{align*}
\]

The SSM stem template has in effect been defined as not including the stem final consonant. But the closed syllable lambic template (Stem 2) and the closed monosyllabic template (15d) do not appear to fit this definition;

(17) trilit. hu.lep.-pa- 'a whistle' \< hu.lep- 'to whistle', -a- agt B.753 exp. li.wa-2-pe- 'speechmaker' \< li.wa- 'to speak', -pe- agt B.737

It has been observed (Freeland 1951, Crowhurst 1992, et al.) that Stem 2 is always followed by an onset consonant. If the suffix has no onset, gemination of C3 will supply it. This also holds for closed monosyllabic templates. I propose that this onset consonant, which appears only when it will not force ephenesis, is a ghost. Postulating closed templatic syllables with a ghost consonant;

(18) hu.lep.- (p)-a- li.wa- ? -(?) - pe- yoh.- (h)- meh

Thus when the ghost surfaces, closed templatic syllables are parallel to the syllable before suffix (13a), CVC-:,u- ma-. It is possible for one such suffix, -a-, 'agentive', B.753, that the gemination is not a ghost, since it also occurs following open templatic syllables, providing an example of a type of Simplex Grade stem template; mol.la.p-:a > mol.lappa-, 'mush making place'. However, all other vowel initial suffixes indicate that this gemination is a ghost, most efficiently represented as part of the closed templatic syllable (representing it as part of the following syllable also fits the shift scenario in 11).

(19) SSM hii.li.w- ay- 'whitefish' ho. po.n- n- ay 'trout' B.703

(20) Underlying representation of closed templatic syllables; CVC-; ()

Ghost consonants are essentially a hedging device on syllabification. They are a protective tactic, in this case protecting the coda consonant from becoming an onset and thereby preserving the right edge alignment of the template and the syllable. The trochaic templates Stem 3 and Stem 4 have a closed heavy syllable, which must be protected to maintain the template. Stem 3 specifies that V2
follows the closed syllable, and for Stem 4 the closed syllable is followed by C3. For template expansion (15):

\[(21)\]

\[
\begin{align*}
\text{Stem 3} & \quad \text{CVC. } \{\cdot\} - V_2 C_3 > \text{CVC.CV}_2 C_3 \\
\text{Stem 4} & \quad \text{CVC. } \{\cdot\} - C_3 V_2 > \text{CVC.C}_3 V_2
\end{align*}
\]

For example: Stem 3: \text{kow.}(w)a.t- \rightarrow \text{kow.wa.t-} \ 'to bump into'

Stem 4: \text{kow.}(w).ta- \rightarrow \text{kow.ta}

Interestingly, the \text{CVC.-} represented with ghost consonants (21) is the same \text{CVC.-} (one vowel) whose template expanding suffixes match the SSM and Ym ghost consonants (11).

9. Epenthesis as a test for template expansion in SSM

As mentioned in 6, epenthesis tests a suffix regarding template expansion. In example (22a) the suffix specifies Stem 3, the C3 qualifies (15a), so the template is expanded. Stem 3 licenses the default vowel as V2, allowing the ghost consonant to surface. The post-templatic consonant cluster forces epenthesis according to the pattern in (6);

\[(22a)\]

\[
\begin{align*}
\text{tek.(k).m- mma-} & \quad \text{to kick with toe- to excess} \\
\text{tek.k.} \text{m.m.m.a-} & \quad \text{one who always kicks} \quad \text{B.729}
\end{align*}
\]

In (22b), the suffix specifies Stem 3, the C3 qualifies (15c), so the template is expanded. Stem 3 licenses the default V2 and the default consonant which is supplied by the suffix (9). The ghost consonant surfaces, and as in (22a), the pattern of epenthesis locates the right edge of the template;

\[(22b)\]

\[
\begin{align*}
\text{kai.l.(l).} \cdot m- \text{mma-} & \quad \text{to kick with heel- to excess} \\
\text{kai.l.} \text{m.m.m.a-} & \quad \text{kicking horse} \quad \text{B.729}
\end{align*}
\]

In (22c), the suffix specifies Stem 3, but the suffix does not qualify (11), so the template does not expand. Stem 3 licenses a V2, so the ghost consonant surfaces. But this V2 is epenthetic, not default, since it is not followed by an expanding suffix;

\[(22c)\]

\[
\begin{align*}
\text{nii.(t).- } \cdot cc- \cdot ni- & \quad \text{to be quiet- static- can, ought} \\
\text{nii.t.} \text{c.c.c.ni-} & \quad \text{might be quiet} \quad \text{B.732}
\end{align*}
\]

\[
\text{*nii.t.} \text{c.c.c.ni-} \quad \text{(note Ym ninee- 'to quieten')} \]

10. Yawelmani underlying templates and surface forms

As already mentioned (6), SSM and Ym have the same epenthesis pattern, sensitive to the right edge of the template. Epenthesis, then, serves as a test for the template right edge in Ym as well as in SSM.

Ym templatic moras associate only to vowels (Archangeli 1983, Zoll 1994, et al.). Therefore the underlying representation of Ym templates contains no consonantal codas. There are no CVC syllables in Ym underlying templates;

\[(23)\]

\[
\begin{align*}
\text{Ym } \text{2a.m}- & \quad \text{2a.m}- \text{2a.m.aa.c- 'approach' (NSM ?amaac- 'wound')} \\
\text{Ym } \text{ca.w-} & \quad \text{ca.w-} \quad \text{ca.waa.- 'shout' (NSM ciwaat- 'chew')}
\end{align*}
\]

As in SSM, conditioning suffixes require particular templates, while nonconditioning suffixes follow default templates (2). Aside from this, the phonological shape of suffixes may alter these templates. The alteration consists of moving the syllable right edge to include a consonantal coda, with vowel shortening if necessary. This is because in Ym template alignment will be sacrificed if it prevents epenthesis (Z 1994). Template alignment will also be sacrificed so that a ghost consonant may surface. But ghost consonants will be sacrificed to prevent epenthesis. The nonconditioning suffix -hin, 'aorist', demonstrates the alteration on the default templates. In (24), template realignment prevents epenthesis;
(24) σμ μ du:b- -hin > du:b hun 'lead by hand' (NSM tuppu-'pull')
    σμ μ hii:x- -hin > hex hin 'be fat'
    σμ σμ μ bi:nii:t- -hin > bi:nii hin 'ask'

In (25), templates are saved from alteration since it would not prevent epenthesis anyway:
(25) σμ μ ho:gn- -hin > ho:gn hin 'float'
    σμ μ di:yl- -hin > de:yl hin 'guard' (NSM tii:li- 'tie someone up')

In (26), Ym onset initial suffixes, like -hin, do not alter σμ σμ μ bilateral templates;
(26) σμ σμ μ ni:nii- -hin > ni:nii hin 'become quiet'

In (27), cluster initial suffixes, like -wsiil, 'rflx., recip. cons. adj.', alter the alignment of bilateral σμ σμ μ templates to prevent epenthesis, and do not alter the triliteral since it would not prevent epenthesis anyway;
(27) i:ik'ii- -wsiil i:ik'ee selv 'tie' (NSM tii:kema- 'web')
    lu:k'uu- -wsiil lu:k'oo la:w.sool 'bury' (PIM -mši- 'recip')

In (28), bilateral σμ σμ μ template alignment is altered to save the ghost consonant, but this alone cannot prevent epenthesis in the triliteral, so the ghost is sacrificed. The suffix -(h)niil, 'consequent passive adjunctive', requires the σμ σμ μ template;
(28) ma:xaa- -(h)niil ma:xath nil 'procure'
    ya:waa- -(h)niil ya:wath nil 'follow' (NSM yoowu- 'go')

This analysis of Ym by Zoll (1994) allows a comparison of phonotactics in Ym and SSM. The SSM template always aligns itself with the right edge of the syllable. The surface forms of Ym templates also align with the right edge of the syllable. The SSM template is always followed by a single consonant which is an onset. The surface forms of Ym templates are also followed by a single consonant which is an onset.

1. Transfer of Miwok templates into Yokuts

In the proposed shift to PY, PEM speakers would bring with them the PEM open templatic syllables CV- and CVV-. What did not survive the shift was the PEM closed syllable CVC-, which included the extra-templatic ghost consonant (20). The loss of the closed templatic syllable might be accounted for by the elimination of the protective ghost, as a simplification of the source language. Without protection, coda consonants could become onsets, thus destroying the template-syllable right edge alignment. This would encourage the retention of open templatic syllables only, where alignment would be stable. Stem templates involving closed syllables, like SSM Stems 2, 3 and 4, would not transfer into PY.

The extra-templatic ghost consonant in SSM guarantees an extra-templatic onset, as in Stem 2. Lack of this ghost in Ym accounts for its lack of the closed syllable iambic, or Light (C 1992), template. The ghost allows for gemination, as in Stem 3, and its absence in Ym accounts for the lack of a Geminate template. This ghost also prevents epenthesis due to analogous reinterpretation of CVC. as CV.C in the Cluster template, Stem 4, which explains the virtual lack (12) of the Cluster template in Ym. The realignment of Ym surface templates to include CVC- may be the effect of the instability of templatic closed syllables during shift.

12. Induced glottalization as the residue of Stem 4 in Yokuts

There are several suffixes in Ym that exhibit a 'floating glottal stop' (Newman 1944,1:8). The various properties of the floating glottal stop have been given a uniform analysis in Zoll (1995).
An example is -(ʔ). .aa, 'continuative' (N.15:32);

(29) (a) caa.w- -(ʔ). .aa > caa.w'aa- shout 2nd post-voc sonorant
    (b) ?ii.ik- -(ʔ). .aa > ?el'.kaa- sing 2nd post-voc sonorant
    (c) maa.x- -(ʔ). .aa > max.ʔaa- procure otherwise, as ghost
    (d) ho.gn- -(ʔ). .aa > hog.naa- float otherwise, as ghost

Glottalized sonorants only occur post-vocalically (N.1:8). Newman suggests
that they are secondary phonemes historically.

The ghost ? in (29c,d), like most ghosts in Ym, matches the SSM default C3.

There is an interesting parallel to (29) in SSM;

(30) (a) tetye- -pu older half sister cf. teete- older sister B.738
    (b) ?amyi- -pu stepmother cf. ?ami- mother B.738
    (c) tunyi- -pu stepdaughter cf. tune- daughter B.738

The SSM suffix -pu has a default C3 which is -ʔ. This suffix in NSM has ? as
a C3; tuneʔpu- stepdaughter. If in SSM this suffix specified Stem 4, the C3
would appear after the second consonant; Stem 2 tu.niy-, Stem 4 tun.yi-. By far
? is the most common default C3 in SSM. Many suffixes are associated with ?/C3
/Stem 4. For suffixes like these borrowed into PY subsequent to shift, the C3
would be added to the stem of biliteral roots. For a sonorant C2, the ? could be
incorporated into a glottalized consonant, since there is no equivalent to Stem 4 in
Ym and so no need for a C3 in that place. C2 is a coda in Stem 4, which explains
why Ym glottalized sonorants are post-vocalic. Once established, the pattern
could have been analogized to triliteral roots as well. As for non-sonorants, the
SSM default C3 pattern continued.

As an example, the Ym suffix in (29a-d); -(ʔ). .aa, 'continuative', may be
the reflex of a borrowed PEM suffix, whose SSM reflex is; -a-, 'It is (bitter,
warm, dead, wild, etc.)' B.702, with ? as the C3 default.

13. An alternative analysis

Callaghan (1994c) offers an analysis of Sierra Miwokan which denies there is
epenthesis, and denies the templatic status of the iambic open syllable form
CVCV.C-. The epenthetic -i- is part of the underlying representation, deleted by
the rule V > O / _V_. The iambic open syllable is derived by CVCV-C(V) >
CVCVCV(C). Callaghan justifies these rules historically, particularly regarding
the UR status of -i-. PU *-i-, 'verbalizer', originally on some monosyllabic
consonant-final noun stems, spread to all consonant-final verb stems in PEM.

This analysis, by capturing history, strengthens the argument for
interference from PEM in PY. We see that these features developed in PEM,
ruling out the possibility that they are interference from PY in PEM. These
features developed to the point where either synchronic description, the Callaghan
analysis or that presented in this paper, would be adequate. Either interpretation
would be possible of course, since the history of a language is opaque to its
speakers. In the case of PEM, the most salient reinterpretation appears to have
involved epenthesis and the iambic open syllable template. This is the
interpretation PEM speakers would bring with them in the shift to PY.

References cited:
Archangeli, Diana. 1983. The root CV- template as a property of the affix:
evidence from Yawelmani. NLLT 1:347-384.
_____.1991. Syllabification and prosodic templates in Yawelmani. NLLT 9:231-
283.


CREOLE STUDIES AND HISTORICAL LINGUISTICS: RENEWING OUR VOWS

John McWhorter
Cornell University

1. INTRODUCTION

When creole studies came formally into existence in 1968 at the Conference on Pidginization and Creolization of Languages in Mona, Jamaica, one of the prime expectations resulting was that creole studies would vitally challenge models of language change, and thus that a new interface, as it were, between creole studies and historical linguistics would be particularly dynamic and rewarding.

However, a survey of the approaches to this interface in the literature over the past twenty-eight years unavoidably reveals that in the end, creole studies thus far has contributed rather little to our understanding of language change. In this paper, after making clear the grounds for this assessment, I will offer suggestions as to how we might make this interface more fruitful in the future, an undertaking which will include a fundamental shift in perspective regarding the nature of the interface in question.

2. THE STAMMBAUM VICTORY

The assertion here is hardly, of course, that historical linguistics has reaped no benefits at all from the work of creolists; we have seen an invaluable revision of the Neogrammarian Stammbaum model as the result of the examination of the parentage of pidgins and creoles. In fact, as conceived by the early creolists, the creole-historical interface was more or less a shorthand for the investigation of this particular Stammbaum issue. Where the Stammbaum model stipulated that all languages can be derived genetically from single ancestors, the fact that several languages make their contribution to all structural layers of pidgins and creoles made it clear that this model required expansion, in order to incorporate what had been designated as Mischsprache (mixed languages) by Hugo Schuchardt in the late nineteenth century (1884).

The revision of the Stammbaum paradigm was hardly instant. In the 1960s and 1970s, before the West African languages constituting the substrate of most Caribbean creoles had been confidently identified, before substantial fieldwork had been done on more than a couple of dozen or so pidgins or creoles, and at a time when a number of creolists were erstwhile Romance scholars, arguments were made for classifying pidgins and creoles simply as descendants of their lexifiers, although usually in more nuanced fashion than legend has hitherto depicted (e.g. Hall 1958:370, Chaudenson 1979, Posner 1983). However, over the past twenty years or so, the massive contribution of substrate languages to all structural levels in creoles has been well-documented, with some scholars going as far as classifying Caribbean creoles as West African languages simply relexified by European languages (Alleyne 1980, Lefebvre 1986). In view of such findings, the status of pidgins and creoles as mixed languages has become more or less a matter of general consensus; at this point it would seem that the only substantial resistance to this conception comes from a particular Francophone school of thought which is of rather marginal general influence (see below for
discussion of this work). Thomason & Kaufman's (1988) already classic treatment of this issue would seem to be the most authoritative and summary statement.

What motivates this paper, however, is that despite the specific intentions of the Mona creolists in their investigation of the creole-historical interface, from a more general perspective the *Stammbaum* issue constituted but one facet of a less theory-specific conception of the interface, and various creolists have seen it as worthwhile to explore other aspects thereof. It is here that it becomes clear that after the early *Stammbaum* victory, subsequent creole-historical encounters have presented a much lower yield.

3. OTHER APPROACHES TO THE CREOLE STUDIES-HISTORICAL LINGUISTICS INTERFACE

3.1. The Creole Continuum as Targeted Change

There can be seen to have been three main schools of creolist investigation which have aimed at forging ties with historical linguistics. The first is the study of creole continua. In the early 1970s, it was discovered that the wide-ranging variation that a given Caribbean creole displayed could be systematized via arrangement according to implicational hierarchies of features. Specifically, it was found that a given feature in a person's idiolect strictly implied the presence of certain others. For example, DeCamp (1971:355) selected six features representative of basilectal, or most conservative, Jamaican Creole English:

(1)

<table>
<thead>
<tr>
<th>IDIOLECT</th>
<th><em>nyam</em> &quot;eat&quot;</th>
<th><em>nana</em> &quot;granny&quot;</th>
<th><em>no ben</em> &quot;was not&quot;</th>
<th><em>pikni</em> &quot;child&quot;</th>
<th>/t/ for /θ/</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>1.</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>6.</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

In (1) we see his arrangement of seven idiolects, each displaying a different selection from among the six basilectal variables, in such a way that it becomes clear that there is pattern amidst the apparent chaos presented by a creole such as this one. Specifically, the presence of certain features automatically implies the presence of others. For example, note that a speaker who has *nyam* for eat, a particularly conservative feature in Caribbean English-based creoles, can be expected to also have all of the other conservative features, while a speaker without *nyam* but who has *nana* for grandmother can be expected to have all of the other features except *nyam*, rather than a smaller, and/or random, subset thereof. The proportion of basilectal features a given speaker controls is proportional to level of education, frequency of contact with standard English speakers, and motivation to imitate same -- in other words, creole continua were discovered to be linguistic indicators of sociological stratification.
The continuum studies inspired by DeCamp's work were couched in a conception of the continuum as representing a diachronic progression through time, the idea being that at the inception of a given colony, only the basilect was spoken, but that as former slaves and their descendants acquired more geographical mobility and economic opportunity, their increasing contact with whites led them to move their basilect towards the local standard. This "movement toward the target" was seen as offering a rich mine of data on what was designated "targeted language change", and spawned a fairly voluminous literature in the 1970s. One of the most eminent practitioners in this vein was Bickerton (1973, 1975, 1980), who conceived the powerful axiom that, to paraphrase, spontaneous change involves the recruitment of old forms in new functions while targeted change involves the recruitment of new forms in old functions. To illustrate, a typical spontaneous change was when Latin recruited forms of the verb habere to encode the future (amare habeo "I will love"), in the place of its eroding inflectional morphology for this purpose (i.e. amabo "I will love"). Here, an "old" form, the longstanding full verb habere, was recruited to fulfill a "new" function, one it had not served previously. As a contrast, Bickerton presented the use of did as an anterior marker in mesolectal Guyanese Creole English, in which did behaves not like standard did, but like the basilectal anterior marker bin. For example, with dynamic verbs bin encodes past-before-past, as in mi bin go "I had gone" (not "I went"), while with stative verbs it encodes the simple past, as in mi bin nuo "I knew" (not "I had known"). What is significant is that when used in the mesolect, did mirrors the behavior of bin: a did go "I had gone"; a did nuo "I knew". This behavior is distinct from the behavior of did in the standard, where it is restricted to inverted interrogative and emphatic sentences (Did I go?, I did know). Thus in its mesolectal reflex, did, under Bickerton's conception, was seen as constituting a form "new" (to the basilect) serving an "old" function, that of the basilectal bin.

As insightful as the diachronic approach to the continuum was, however, over the past several years there has been a profound paradigm shift in how creole continua are viewed. Historical investigation and documentation (e.g. Lalla & D'Costa 1990, Baker 1990) have made clear that the continua of lects on view in a given creole today actually have existed since the inception of the colony, rather than having emerged only over time. Thus whereas in 1971 the continuum was seen as having arisen as the result of latter-day socioeconomic change, in 1995 it is seen as having resulted from the varying degrees of contact with whites that slaves had even at the founding of the colonies -- field slaves would have had the least such contact, but house slaves and slaves residing in towns would have had much more.

The significance of the direction continuum studies have moved is that, while in itself representing the progress of scientific inquiry, it has the fortuitous result of essentially eliminating the continuum as an example of language change per se. Instead, the continuum of dialects represents various degrees of the acquisition of the standard. Mesolectal registers, for example, can no longer be seen as the result of a basilectal register moving towards a target over time; instead, a mesolectal register represents a degree of acquisition acquired at the inception of a colony, passed on to subsequent generations like any other language. More specifically, then, it must be stated that the work couched within the original conception of "targeted change" is now obsolete: Bickerton's axiom, for example, elegant though it is, can no longer be seen as having been formulated on the basis of an actual phenomenon. Rickford (1988:35) has suggested that despite the new
paradigm, we need not rule out that some actual spontaneous change still takes place when creole speakers move to different lects in the course of their lifetimes. This is certainly true, but given 1) the difficulty of distinguishing the created features from those simply acquired, 2) the fact that we can assume that acquisition vastly dominates creation in such contexts given the exigencies of communication, and that 3) the similarities between earlier texts and modern speech appear to confirm that generational transmission has kept creation at a relatively marginal level, it is unclear that maintenance of the earlier paradigm would yield significant benefits. In sum, then, all of the continuum literature in question must today be read through this lens, and leads us to look elsewhere for creolistic insights into language change.

3.2. Tok Pisin and Permeability to Mixture

In the meantime, in the late 1970s and 1980s Peter Mühlhäusler offered another perspective on the creole-historical linguistics interface in his work on Tok Pisin, the English-based creole of Papua New Guinea. One of Mühlhäusler’s many concerns was to identify constraints upon language mixture according to subsystem, his working hypothesis being that certain linguistic subsystems are less permeable to language mixture than others. Part of his intention in this work was to show how pidgins and creoles could be used to shed light on language change in contexts of contact.

Mühlhäusler’s analysis led to conclusions which are best summarized by this passage:

As a general principle, it can be postulated that the more arbitrary an area of grammar, the more readily can languages borrow from one another. With regard to the formation of developmental continua such as the pidgin-creole continuum, this implies that: 1) substratum influence will be most pronounced in the areas of lexical semantics; prosodic phonology, some segmental phonology [and] pragmatics whilst superstratum influence will be strongest in lexical form [and] segmental phonology; relatively independent of substratum or superstratum influences are syntax, inflectional morphology [and] derivational morphology [emphasis mine]. (Mühlhäusler 1980:36)

However, once again, further research over the past several years has, while lending us invaluable new insights on Tok Pisin and its history, had the unfortunate side effect of severely constraining the relevance of Tok Pisin itself to the language change issue, at least in the vein Mühlhäusler intended. While Mühlhäusler couched his hypothesis on the most fundamental level in traditional and intuitive conceptions of the relationship between arbitrariness of subsystem and susceptibility to transfer, his empirical base consisted of detailed comparison of Tok Pisin with the Austronesian language Tolai, based on the fact that this is the language spoken natively by the people who first adopted Tok Pisin in New Guinea. Mühlhäusler noted that despite the occasional teasing correspondence or two between Tok Pisin and Tolai, systematic comparison revealed a predominance of contrasts. For example, he noted that Tok Pisin and Tolai tend to display contrasting noun-adjective orderings despite the occasional apparently fortuitous match-up, as in (2), where we see that despite the correspondence in (2a) and (2b), most other cases contrast:
(2) Adjective placement in Tolai and Tok Pisin (adjectives in bold) (Mosel 1980:57-8):

<table>
<thead>
<tr>
<th>TOLAI</th>
<th>TOK PISIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) A tutana tuna. DEF man real</td>
<td>(b) Man tru. man real</td>
</tr>
<tr>
<td>A real man</td>
<td>A real man</td>
</tr>
<tr>
<td>(c) A mo na gamata. DEF ripe C apple.</td>
<td>(d) Banana mau. banana ripe</td>
</tr>
<tr>
<td>A ripe apple</td>
<td>A ripe banana</td>
</tr>
<tr>
<td>(e) A kaina pakana. DEF bad place</td>
<td>(f) Ples nogut. place bad</td>
</tr>
<tr>
<td>A bad place</td>
<td>A bad place</td>
</tr>
<tr>
<td>(g) A mal pua DEF clothes white</td>
<td>(h) Wetpela klos white clothes</td>
</tr>
<tr>
<td>White clothes</td>
<td>White clothes</td>
</tr>
</tbody>
</table>

However, various Pacific pidginists over the past seven years or so have decisively challenged these conclusions. For example, Keesing (1988) shows that Tok Pisin, rather than emerging and taking shape in New Guinea, actually emerged and developed decades before Mühlhäuser stipulated, as a widespread Melanesian Pidgin English developed starting in the early nineteenth century in the context of the whaling, sea cucumber, and sandalwood trades. As such, Tok Pisin can be assumed to have taken its fundamental form not among Tolai speakers in New Guinea, but much earlier among speakers of closely related Eastern Oceanic languages in Melanesia. This becomes particularly clear, for example, when we note how comprehensively and idiosyncratically similar Tok Pisin is to other Melanesian Pidgin English dialects such as Bislama, Solomon Islands Pijin, and Torres Strait Broken, which developed separately from the same early-pidgin progenitor as Tok Pisin -- the similarities demonstrate that the parent pidgin had expanded into a form much like its present-day progeny, since such close correspondences cannot have resulted from parallel development.

The crucial fact here is that it appears that Mühlhäuser chose the wrong substrate language upon which to base his conclusions about language mixture: while Tok Pisin indeed corresponds only erratically with Tolai except in terms of lexical contributions, it corresponds much more closely to Eastern Oceanic structure, a predictable result of it having taken shape in Melanesia rather than New Guinea. In (3) we see that where Tok Pisin and Tolai do not correspond, such as in the encoding of plural with the third person plural subject pronoun, Tok Pisin and a wide range of representative, closely-related Eastern Oceanic languages do:

(3) Correspondences between Tok Pisin and E. Oceanic as compared to Tolai:

<table>
<thead>
<tr>
<th>TOK PISIN</th>
<th>TOLAI</th>
<th>E. OCEANIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The result of these findings is that Mühlhäusler's theoretical ideas about language mixture, based upon an unintentional miscomparison, would appear to be largely invalidated, at least as far as they were derived from analysis of Tok Pisin. Whereas Mühlhäusler's comparison of Tok Pisin and Tolai naturally suggested language mixture in Tok Pisin occurring to starkly different degrees according to linguistic subsystem, the findings of scholars like Keesing (bolstered by Romaine 1992, et al.) instead demonstrate that language mixture permeates all systems of Tok Pisin to considerable degree.

Indeed, this has in general been shown to be true of all creoles, which can be classified as extremes along a cline of degree of language mixture. The fact is that the unusually extreme and wide-ranging language mixture in creoles hardly invalidates the basic notion that in less extreme cases of mixture, we can expect differences in permeability among subsystems less apparent in creoles themselves. However, it would seem that we would gain more insight into such issues through the examination not of a creole like Tok Pisin, which represents an extreme extent of mixture in all subsystems, but of instantiations closer to the middle of the restructuring cline, such as Yiddish, Andean Spanish (Lozano 1975), or Yawelmani Yokuts (see Kramer, this volume).

3.3. The French Superstratists

In the meantime, we see a final potential contribution to our interface in the form of the French superstratist school of creole origins (Chaudenson 1979, Fournier 1987, et al.). While most creolists agree that creoles are fundamentally mixed languages, many Francophone creolists hold to a belief that creoles constitute no challenge to basic Stammbaum theory, and that there is nothing in creole structure which cannot be attributed to nonstandard dialects of French undergoing ordinary language change unfettered by the conservatizing effects of literacy and standardization.

This view would seem, in the final analysis, to proceed from a perspective which, while philologically well-informed, neglects to engage substantially the decades of work outside of this school which would seem to offer rather comprehensive and irrefutable demonstrations that plantation creoles display a wealth of features difficult to subsume under the spontaneous change designation. Indeed, there are certain features of French-based creoles, for example, that appear exotic to the modern eye but in fact can be traced to regional French constructions. An example would be the progressive marker *ap*, used as in the Haitian Creole *m'ap pale* "I am speaking". This marker is derived from a regional French usage, still current, for example, in Québecois French as in *Le chat de Marie-Sylvia est après jouer* "Marie-Sylvia's cat is playing" (Tremblay 1978:108). Derivations like this are instructive, and indeed the contribution of regional dialects to creole structure remains underappreciated in English-based creole studies to this day, as becomes apparent when one encounters the richness of
evidence in occasional exceptions such as Hancock (1993). However, the fact is that regional dialect contributions, current or extinct, can hardly account for a great many other nonstandard features one finds in creoles. For example, it is highly unlikely that any regional French dialect has ever omitted the copula as the French-based creoles do in many contexts, as in the Haitian *Bouki o anba tab-la* "Bouki is under the table" (DeGraff 1992). Or, to extend the argument to creoles of other lexical bases, it is highly unlikely that any regional English dialect ever postposed positional adpositions as Sranan Creole English does (*Kofi de a tafa ondro* "Kofi is under the table"); a West African derivation, such as the Ewe *E le kpolo te* he is table under "He is under the table" (Westermann 1954:638) is much more likely given the provenience of slaves brought to Suriname during the period of formation crucial to Sranan’s development (McWhorter 1994).

Thus the superstratists would like creoles to be treated as an instantiation of a particular type of diachronic phenomenon, perhaps lightly brushed by contact with other languages. However, it would seem that the predominance of features prototypical of either West African contact, such as serial verbs and completive markers postposed to VP, or of radical reduction of a type foreign to spontaneous change, such as zero copulas where there were none before, render this frame of reference of limited application at best to a general theory of creole origin. Most importantly, it would seem to be outright dangerous to language change theory, given that its proponents purport that divergences as sharp as those between plantation creoles and their lexifiers are concomitants of spontaneous change, a claim belied by work in all other frameworks.

4. **RETHINKING THE INTERFACE**

Despite the unavoidably critical intent of the above discussions, this paper is written in a fundamentally constructive vein. The foregoing is intended solely to make clear the grounds for my sense that a fresh perspective must be applied to the creole-historical interface if there is to be such.

It is my contention that a partnership of creole studies with historical linguistics is necessary and of rich potential, but that if it is to be a vital partnership, then creolists will have to approach the issue in a somewhat different frame of mind than in the past. Specifically, there is most likely little of import discovered by creolists which will significantly reformulate language change theory as it is currently conceived; it is likely that the *Stammbaum* victory was the only such bomb to throw. Creole studies indeed has icons to topple elsewhere, such as the reformulation of generative theories of competence as promised by the early continuum scholars, and interactions with first- and second-language acquisition. However, the creole-historical interaction will be most fruitful at this point if treated more pacifically.

The justification for this response hardly springs from any devotion to bland even-handedness of the sort which tends, when overapplied, to impede true progress in scientific inquiry. Rather, the approach I have suggested is a concomitant of a more general realization over the decades that when viewed apart from their genesis and histories -- crucial to theory in themselves -- creoles are best treated as languages just like any other. Creolists have come to realize that there are no prototypically "creole" features, in that none of the features which cluster in many creoles are absent in other languages and even at times cluster in other languages as well (Chinese, with its analytic structure, serial verbs, etc. being a notable example). Moreover, contrary to an oft-cited axiom of
the past, creoles have not been observed to change more rapidly than other languages (we note, for example, the relative likeness of the earliest documentation of Sranan in 1718 and the modern language), and while pidgins indeed do develop rapidly to fulfill the needs of full language, situations in which this expansion can be or has been actually observed are rare to non-existent: the early pidgins we can observe are unlikely to expand (Gastarbeiterdeutsch), while the earlier stages of creoles and expanded pidgins are generally unrecoverable. In line with these realizations, then, in a rethought interface between creole studies and historical linguistics, the emphasis will be less on what creoles have to offer as a separate linguistic type than on what creoles have to offer given the particular traits -- out of the full set of traits found in all languages -- which happen to cluster in pidgins and creoles for various reasons. This is perhaps a little less titillating than the Trojan Horse approach of the old days, but it will also be the most constructive for all parties concerned.

4.1. Pidgins and Creoles as Sources of Data

Thus pidgins and creoles will contribute to historical linguistics in serving as sources of data for change in constructions typical of contact languages. For example, topic-comment constructions, particularly in which the subject is resumptive of the topic, are a typical trait of pidgins and creoles as the result of what appears to be a universal discourse strategy:

Saramaccan Creole English:
(4) Di nEn, hen da sükuma.
   the name it COP foam
   The name was "foam".

Tok Pisin:
(5) Nau wanfela masta bilong kampani em i-kisim mi.
   then ART white-man of company he PM-take me
   Then a white man from the company took me. (Hall 1966:149)

What is important here is that as it happens, cross-linguistically, topic-comment constructions tend to be the site of a wide range of reanalyses and grammaticalizations, resulting from the reanalysis of topic as subject, and pidgins and creoles are no exception. In Saramaccan Creole English, for example, this author has documented the emergence of the identificational copula from an erstwhile topic-comment construction, as shown in (6) and (7)) (see McWhorter 1993 for details):

Saramaccan (most likely as Guinea Coast Creole English) (reconstructed):
(6) Kofi, dati ði Gaama.
   Kofi that the chief
   Kofi, that's the chief.

(7) Development of the identificational copula in Saramaccan (McWhorter 1993):

<table>
<thead>
<tr>
<th>Stage A (reconstructed)</th>
<th>Stage B (modern)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Kofi] [da] [di Gaamá]</td>
<td>[Kofi da di Gaamá]</td>
</tr>
<tr>
<td>topic subject COP predicate</td>
<td>subject predicate</td>
</tr>
</tbody>
</table>

This author has elsewhere documented the development of an idiosyncratic allomorphy of the predicate negator, outlined in (8) (see McWhorter [forthcoming] for details):

(8)

<table>
<thead>
<tr>
<th>STAGE A</th>
<th>STAGE B</th>
<th>STAGE C</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reconstructed:</strong></td>
<td><strong>Reconstructed:</strong></td>
<td><strong>Modern:</strong></td>
</tr>
<tr>
<td>(11) Kófi, a ná wáká</td>
<td>(12) Kófi ḣe-NEG wáká</td>
<td>(13) Kófi NEG wáká</td>
</tr>
<tr>
<td>Kófi ̣ he doesn’t walk</td>
<td>Kófi he-NEG walk</td>
<td>Kófi doesn’t walk</td>
</tr>
<tr>
<td>[Kófi] [a ná wáká]</td>
<td>[Kófi] [á wáká]</td>
<td>[Kófi á wáká]</td>
</tr>
<tr>
<td>TOPIC COMMENT</td>
<td>TOPIC COMMENT</td>
<td>SUBJ-PRED</td>
</tr>
</tbody>
</table>

Elsewhere, as for example in Bislama Creole English, we see the grammaticalization of an erstwhile third person pronominal resumptive *i*, derived from *he*, as a paradigmatically generalized predicate marker used in all persons and numbers, as shown in (9), where it is used in the second person despite its third person singular etymology and original usage:

**Bislama Creole English:**

(9) a. Himi stap talen wanem? (Original parsing: Him, he stop talk what?)
   him he PROG talk what
   What is he saying? (Crowley 1989:186)

   b. O yufela i krangi.
   oh you-PL PM crazy
   Oh, you’re crazy. (Charpentier 1979:374)

Pidgins and creoles also display a very strong tendency to encode tense, mood and aspect via preverbal particles, generally with a rich and systematic range of combinatorial possibilities. This makes pidgins and creoles invaluable sources of data on the evolution of such markers over time, and by extension fine places to watch the development of inflectional morphology (although, crucially, pidgins and creoles are hardly alone in this). For example, Guinea-Bissau has reanalyzed the Portuguese verb *acabar* "to finish" as a past marker of unusually flexible syntactic movement, as shown in (10). A marker which appears to resemble its Ibero-Romance progenitor semantically in modifying VP as in (10a) surprises us in allowing accusative pronominals to intervene between the verb and itself as in (10b); furthermore, it even marks NP predicates as in (10c). Note also that this is not an SOV language, thus ruling out analyzing *ba* here as a sentence-final verb "to be" or the like:

**Guinea-Bissau Creole Portuguese:**

(10) a. I ten ba un minjer ki ten ba manga de konpadre.
   it have ANT a woman REL have PAST many of friend
   There was a woman who had a lot of friends. (Kihm 1980:234)
(10) b. I ka ta permiti-n ba pa n papyaa.
    it NEG IMP permit-me PAST COMP I speak
It didn't permit me to speak. (Kihm 1989:365)

(10) c. Ami i karpinteru ba.
    I PM carpenter PAST
I was a carpenter. (ibid.)

In the meantime, the curious habitual *be* of African-American Vernacular English (AAVE) has been traced back to an original concatenation *doz be* still on view in Gullah Creole English. In AAVE, *he be workin'* can only signify the habitual, and never, contrary to the folk perception, the progressive. This can be traced back to early Gullah, as seen in (1a), after which *doz* be eroded phonologically to simply *be*:

Stage 1:
(11) a. *He does be up and cut wood sometimes.*

Stage 2:
(11) b. *Sometimes you [z] be in the bed...*

Stage 3: (now also current in AAVE):
(11) c. *He ø be up and cut wood sometimes.* (Rickford 1986:270-1)

There are myriad such cases awaiting examination in the pidgins and creoles of the world.

4.2. Pidgins and Creoles and Grammaticalization

Furthermore, in general, creoles prove themselves to be rich sources of information on processes of grammaticalization. This is in part because, as new languages, they are often in the process of developing grammatical items whose origins in many regular languages is lost to time. For example, Bislama, which developed from Mealnesian Pidgin English which can be reconstructed to have had but two prepositions (*bilong* for possession and *long* elsewhere in all functions), is developing a rich collection of new prepositions out of serial verb constructions, as in the case of *kasem* "to get" in (12). *Kasem* began as a simple verb, as in (12a):

(12) a. Mi no *kasem* mane long bang.
    I NEG get money PREP bank
I didn't get money from the bank. (Crowley 1990:76)

What sparked the reanalysis of *kasem* as a preposition was its use in serial verb constructions. In Bislama, both verbs in serial construction are marked with the predicate marker *i*. *Kasem*, however, began to occur without it, suggesting that it was shedding the properties of verbhood:

(12) b. Sip ia i ron long Vila (*i) *kasem* Santo evri wik.
    ship this PM go PREP Vila get Santo every week
This ship goes from Vila to Santo every week. (ibid. 78)
Next arose semantic extensions, as well as occurrence within NP and PP, as in (12c):

(12) c. Graon long ples ia kasem taon  
land PREP place this to town  
The land from here to town... (ibid. 79)

Finally, today kasem can strand as in (12d), which is diagnostic of prepositions and impossible for verbs in the language:

(12) d. Wan mil nao, hem i sakem vatu kasem.  
one thousand now he PM contribute vatu to  
It was as much as 1,000 vatu that he contributed. (ibid. 85)

Creoles are also valuable to the investigation of grammaticalization because, despite the complaint common among creolists that historical records of creoles' earlier stages tend to be scanty, the fact is that creolists are comparatively well-off in this regard in comparison to many linguists working on other language families. Thus scholars of the Suriname creoles, for example, have almost three centuries of relatively copious documentation to work with, while most Africanists, Papuanists, Austronesianists, or Amerindianists could never hope for the same. This is an outgrowth of the fact that most creoles emerged in contexts dominated by Europeans, who had a stake in documenting the creoles for religious and communicative purposes. Thus creolists are hardly as blessed as scholars of the history of English or Russian, but a historical linguist can nevertheless take advantage of a creole's recent birth and its historical documentation to draw a great many discoveries, when combining the documentation with comparative and internal reconstruction.

For example, combining synchronic and diachronic analysis of a single creole, Sranan Creole English, one can trace the evolution of definite articles, a complementizer, a copula construction, a dative preposition, subordinating conjunctions, a place adverb, and a wide range of morphophonemic phenomena. Note illustrations of some of these in (13) and (14), with the processes of change sketched in more detail in (15), (16) and (17) (the latter strictly not a grammaticalization, but included to further illustrate the development of morphology as described in 4.1.):

Reconstructed Sranan c. 1690 (most forms attested):

(13) Dati man, dati man disi ben go tei mi.  
that man that man this ANT go tell me  
That man is the one who would have told me.

Modern Sranan:

(14) A man-dati na a man di bo tei mi.  
the man-that COP the man REL COND tell me  
That man is the one who would have told me.

(15) The erosion and redevelopement of distal demonstrative NP modification:
a. *dati* man "that man"
b. *da* man "the man"
c. *a* man "the man"
d. *a man-dati* "that man"

(16) The development of demonstrative pronoun to equative copula; proximal demonstrative pronoun to relativizer:

a. *Dati* man, *dati ô man disi*... "that man, that's the man who..."
b. *Da* man, *da man di*... 
c. *A man-dati da (da) man di*... "that man is the man who..."  
d. *A man-dati na a man di*...³

(17) The emergence of a monomorphemic conditional marker:

a. *ben go tei mi* "would have told me"
b. *ben o tei mi* 
c. *bo tei mi* 

4.3. A Two-Way Street

Finally, the creole-historical linguistics interface has always been couched in terms of what creole studies had to offer historical linguistics. However, at this juncture we need also consider what historical linguistics has to teach creole studies. The past ten years have seen a minor explosion in diachronic studies of creoles, and I think that this body of work would benefit immensely from close attention to the tenets and findings of historical linguists in general. Specifically, creolists must bring not only documentation to bear upon charting diachronic processes in creoles, but also comparative and internal reconstruction. The examples in (15) through (17), as well as (7) and (8), serve as illustrations. For example, the development of the Sranan and Saramaccan equative copulas is most rigorously and refutably charted via reference not only to historical documents, which themselves shed only partial light on the process, but via reference to the cross-linguistic literature on copula development, which reveals a strong tendency for copulas to develop from resumptive demonstrative pronouns in topic-comment constructions, and to principles of internal reconstruction, which reveal synchronic idiosyncracies which render the demonstrative derivation nearly inescapable (see McWhorter 1993). In addition, emerging creolists should seek thorough training in diachronic linguistics in order to most usefully fulfill the requirements of bringing its insights into creole studies.

5. CONCLUSION

In general, it is my view that our current aproach to the creole studies-historical linguistics interface is best informed by a bird's-eye perspective, via which we can see that the opening rounds of interaction produced a vital paradigm shift, but that the facts of the situation have since demonstrated that the initial catastrophically-oriented pitch has not lent itself to being maintained indefinitely -- as is the nature of most such interactions between two fundamentally sound theoretical orientations. Instead, our findings since that paradigm shift have shown us that the
interface can and must proceed, but will do so most fruitfully via a cooperation intended for the constructive improvement of both frameworks. The very existence of such an interaction is but one of many signs at this writing that display that creole studies, after almost three decades of exciting but inevitably somewhat breathless forays, is coming gracefully of age.

REFERENCES


Lima: Instituto de Estudios Peruanos.


FOOTNOTES

1. Bickerton’s (1981, 1984) Language Bioprogram Hypothesis is based upon the identification of the language of aged Hawaiian informants in the 1970s as a pidgin variety of the creole spoken by later descendants, but this analysis has been strongly challenged by Goodman (1985), Holm (1986), and McWhorter (1993, 1994). Meanwhile, Mühlhäuser in much of his work (e.g. 1980) designates certain varieties of Tok Pisin as "pidgin" varieties in contrast to "creole" varieties elsewhere, but the additional historical data cited above put these designations into severe question. In both the Bickerton and Mühlhäuser cases, we are left with the strong possibility that what the authors designate as "pidgin" registers
were actually second-language registers acquired of previously existent, more expanded reflexes.

2. For diachronic data on Sranan, see Arends (1989) for the most useful treatment, as well as Kramp (1983).

3. The extra da here is derived from a separate process, the bleaching of another instance of dati into definite article a via the intermediate form da, sketched above in (15). The extra dati is omitted in (1a) both for purposes of clarity, and because it can be presumed that at an early stage, determination had yet to be regularly marked overtly, with overt marking arising only via the bleaching and generalization of erstwhile demonstratives which had at first been used only for emphatic deictic designation (see Greenberg 1978, e.g., for a general conception of the development of determiners from demonstratives).
SOCIAL ISSUES IN HISTORICAL LINGUISTICS IN AFRICA

DEREK NURSE

MEMORIAL UNIVERSITY OF NEWFOUNDLAND

1. My purposes here are to outline (1) the way in which I think of the term “social issues in historical (and comparative) linguistics in Africa”, and (2) some of the lessons we have learned over the last few years and the directions in which I think we are moving.¹

I use the term “social issues in historical and comparative linguistics” very broadly to refer to any attempt to relate demonstrable linguistic events to real or assumed human or historical events. Historical linguistics has been characterized since its inception by a tendency to regard linguistic phenomena as if they occurred or existed independently of people. I think we are now trying to put people back into the formula. This is happening world wide but also in Africa. A forthcoming issue of the journal *Sprache und Geschichte in Afrika* consists entirely of some 15 articles on Language Contact and most are concerned, to some extent, with the human component.

What time period is being referred to in general? Historical linguistics effortlessly covers from yesterday back to the establishment of today’s language families. With the exception of Khoesan, that would probably not exceed 10,000BP in Africa. At that distant time remove, given today’s methods, historical linguistics will not yield us much worthwhile detail. In practice, historical and comparative linguistics in Africa have proved most useful in interpreting the events of approximately the last five millenia.

SOCIAL ISSUES IN HISTORICAL AND COMPARATIVE LINGUISTICS IN AFRICA

2. I would like to start with a case study. Recently T. Hinnebusch and I wrote a history of Swahili. It covers roughly the first fifteen centuries AD. We proceeded in the time-honored way using the Comparative Method: triangulation back from the present to the past, from current data to reconstruction of the past. Today Swahili is used across East Africa but it is known that prior to about 1800AD it was just a coastal phenomenon, stretching from southern Somalia down along the Kenyan coast, the Tanzanian coast, to central Mozambique; earlier it might have gone down to southern Mozambique. Our data came from all the old coastal dialects and we reconstructed aspects of Proto-Swahili (PSW: a partial lexicon, a C and V system, certain morphophonemic processes, the nominal system, and the verb paradigm, concentrating on tense and aspect categories). We did not stop at PSW but went a little further back in time. We did that for a number of reasons. Partly because we had comparable and comparative data for all of Swahili’s nearest relatives and why not use it? Partly because Swahili’s relatives are not very much farther along the continuum of variation evident in the Swahili dialects and so, chronologically, probably did not come into existence very long before the Swahili dialects. Partly because considering the relatives threw light on the development of Swahili itself. For example, reconstruction of suprasegmental features for PSW using only data from Swahili today would be difficult. Most Bantu languages are tonal, Proto-
Bantu was also tonal, it seems likely that PSW was too, but only two or three of today’s 15 or so Swahili dialects are tonal and the systems are vestigial so they don’t much help with reconstruction. But all of Swahili’s relatives are tonal so including them threw more light on earlier conditions in Swahili.

The phrase "It covers roughly the first 15 centuries AD" is used advisedly. Where it stops is clear enough - about 1500AD - partly because we assume that change after that point was relatively small and partly because that marks the entry of the Swahili into history. The Portuguese arrived at that time, described what they saw in some detail, and events after that are less interesting and challenging to an historical linguist. We think that we started around the start of the Christian era because we were able to tentatively link our linguistic picture to a very few known or assumed archaeological facts. The co-author and myself disagreed on a few matters. One was how we viewed the limit of our job. Both of us thought of ourselves as comparative and historical linguists, that is, as comparativists we were interested in moving backwards to reconstruction and as historical linguists we were interested in charting the forward movement of language development. But whereas my co-author viewed his role as more or less stopping with that, I was also interested in linking purely linguistic evolution to real or assumed non-linguistic, external, events. I would go further, and say that I am actively interested in using linguistics as a tool for reconstructing prehistory. I would regard as incomplete linguistic explanations that do not take into account relevant non-linguistic variables. Although it would be an overstatement to claim that all linguistic change is externally driven - in most cases, for example, palatalization happens without any external cause - some kinds of linguistic change, and not just lexical, relate to events in the external world.

Because of this interest, a few years ago, another co-author, a historian, T. Spear, and myself produced a book on the history of the Swahili people. What we attempted to do there was put together a history of the coastal Swahili from ca. 800 to 1500AD.

This raises the question of how to do this: how does one write the history of a language and a community whose earliest written records date from only some three centuries ago? That could be expressed in more general terms. Africa is said to be home to some 1500 languages, which is said to represent about 25% of the world’s languages. If we exclude languages spoken north of the Sahara such as Coptic or Arabic, none of those African languages were written down by indigenous people or by outsiders more than a very few centuries ago. Or if they were, no records have come down to us. That could also be couched more generally. We are used to work with languages such as a few in Europe, the Middle East, and India, a very few in Asia, and perhaps one or two in Central America, that have a written tradition or written documents that go back a millenium or more. How many are they? Out of the world’s 6000 languages, only a handful go back two millenia or longer and a few dozen exceed one millenium. So writing a history covering a millenium or two or three for a language whose written records go back only a century or two or three has implications for most of the world’s languages.

Parallel to the history of Swahili is the history of the coastal Swahili-speaking community. If one wanted to write a history of, say, the country or the nation that speaks French, that is, France or the French, where would one start and how would one set about it? One might decide to go back as far as, say, 800AD, because that is about when French documents start, when Charles the Great thrived. A linguist might find it hard to tell where Late Latin finished and French started, so might claim that France is an extension of Rome and that it is necessary
to go back to, 300BC, which is about when Latin documents started. In either case the writer would have access to, firstly, a long, continuous, and fairly substantial set of written documents by the community itself, secondly, another set of documents, by outsiders, which would be less detailed or substantial but nevertheless informative, especially for the earlier period, and thirdly, a fairly systematic and detailed archaeological record covering a long period. One would not have to rely on oral traditions to any extent, one would not have to rely on historical linguistics because the kinds of things they reveal are known already. We all know that Latin and the French replaced Celtic and the Gauls and that Latin grew into French.

None of that would be true for the Swahili-speaking community. Written literature in Swahili goes back just some three centuries, is locally based, and deals mainly with religious themes. It is not much help with uncovering two millennia of history. Accounts by outsiders are mainly also of local areas, and are geographically and chronologically spasmodic. Although they span nearly two thousand years, there are frequent gaps of several centuries, they refer only to a few places (Muqdishu, the Green Island, Rhaps, the Comoro Islands) that are often many miles apart and sometimes by names that no longer exist, and the detail they provide is haphazard and does not afford anything like a complete picture. The archaeological coverage is also incomplete and somewhat haphazard. The amount of money available to dig holes on the East African coast is paltry compared to that for other parts of the world. A small group of individuals has done splendid work over the last thirty or forty years and built up a basic archaeological profile of Swahili history, but I think all would say they are just picking at the surface of the task. More is undone than done.

Most other African communities are poorly served compared to Swahili and the Swahili. Most have no written records at all or have records that are even shorter than Swahili’s. They are unmentioned by outsiders: as various outsiders sailed down the East African littoral from southern Arabia during the first fifteen hundred years AD, they must have passed dozens of discrete communities but not a single one is mentioned by name. Similarly, while an archaeologist on the Swahili coast can be reasonably sure of the general correlation between linguistic and archaeological profile, that is not true of most other places in Africa or the world. While archaeologists can be fairly sure in some cases of the correlation between, say, potsherds and language family a millennium or two ago, they are usually far from being able to identify with certainty the individual communities within a language family responsible for the sherds.

A source not much used by communities in the world that do have lengthy written records is oral traditions. Many African communities have - or did until they broke down in the twentieth century - oral traditions passed on from generation to generation, traditions often covering the events of the last several centuries. For the Swahili these were either traditions of individual towns or traditions of clans, or apocryphal stories, inspired by religion, of how they came from southern Arabia in general and Mecca in particular. For other societies in Africa, the traditions cover groups larger than clans. All are of limited use, partly because they explicitly cover only local areas or local groups, and especially because as historical accounts they are only reliable back for a few centuries (many observers would say maybe five centuries before present). Some cover events before that but then they are no longer “objective” accounts of history, nor are they intended to be.

Given this general situation, historical and comparative linguistics becomes more prominent as a tool for history. Distinguishing language families and their
subgroups is useful because it provides a framework on which historians can hang details of early events. It implies the existence of earlier communities which historians can elaborate. Consideration of the geographical distribution of language communities and language families gives broad hints about the probable earlier movements of peoples. Reconstruction of parts of earlier languages - mainly vocabulary - affords historians details of earlier culture and its possible sources. Analysis of which components of individual languages or groups of languages are inherited and which derive from contact with other peoples often tells historians about early events they only suspected or did not even suspect. I am constantly amazed how often historians accept statements by linguists about historical and comparative matters, and often without questioning them because they do not feel confident enough to deal with the methodology.

That is why historical and comparative linguistics is useful for historians in Africa and elsewhere, particularly for periods of history before the advent of written records. There are limits to its role. This is not Martha’s Vineyard, or the Lower East Side, or Panama City, or Norwich, or Turkish communities in Germany, or the mobile women of Belfast, or a systematic dialect survey. We can never go into the linguistic communities that have covered most of the world during most of its history and conduct investigations into the progress of vowels or socioeconomic status or attitudes to language, nor can we repeat the investigations a generation later just to make sure change is on schedule. Social issues in historical linguistics in most of the world’s languages are not the issues of contemporary sociolinguistics.

I would like to turn now to some of the lessons we have learned and have picked just a few out of a long possible list. Some are unique to Africa, some are familiar from work elsewhere in the world over the last two or three decades.

SOME LESSONS

3. IN PRACTICE ALMOST ANYTHING CAN BE TRANSFERRED IN CONTACT SITUATIONS. In what follows the general term “transfer” is preferred to “borrow/loan” to refer to any situation where linguistic material has moved from one community to another. Borrowing/loaning is a subform of transferring, usually resulting from transfer due to lengthy geographical adjacency. It used to be said there was a hierarchy of features that could be borrowed from one language to another, an order in which they were likely to be borrowed. Cultural vocabulary would be borrowed before core vocabulary, content words before function words. Vocabulary in general was borrowed before phonemic units, followed by phonetic processes, followed by phonological processes, followed by morphological processes in a certain order, followed by syntactic processes, marked features were likely to be eliminated in favor of unmarked features, and so on.

Languages which did not fit into this hierarchy were regarded as exceptional. One such language is Ma’a in Tanzania. It has a largely Bantu grammar and a largely Cushitic vocabulary. Over the last few decades many have tried to unravel the mystery of how this came about. The explanations are various but all used to agree that this was a strange case.²

We have started to learn in the last decade that it is not so strange. The process started back in the 1970s with studies such as that by Gumperz (1971) of the case
of Kupwar village in India, where several language communities have lived closely together for five or six centuries. Each language in the area has developed a local variant, and all these local variants share a similar or identical surface structure. The main differences are lexical. The result is that a person there slots one set of lexis into the surface structure if they want to communicate with people A and exclude people B, but if they want to exclude A and include B, they slot in a second set of lexis and so on. This is an effective and economical strategy and a model which can be usefully applied elsewhere.

Then came Thomason and Kaufman’s book (1988, henceforth TK) which recognizes that historical collisions between languages are common and tries to set up a typology of what happens in different sets of circumstances. Ma’a is there and does not look quite so exceptional.

Most recently there appeared a book edited by Bakker and Mous (1994) called simply “Mixed Languages”. Mixed languages, ostracized for many decades, have come in from the cold. It is becoming acceptable again to speak of mixed languages, as they appear more often. It contains 15 case studies of relatively unknown languages from all over the world. Ma’a is there and starts to look almost mainstream. One finds what one looks for. If one looks for regular and exceptionless sound change one will find it. If one looks for cases where what is most obviously transferred is vocabulary one will find them. If one starts to look for a different set of priorities, if one looks for mixed languages, if one starts to look for numerous cases of things other than vocabulary being transferred, they will be there. Highly marked features such as clicks or ejectives can be transferred, articulation habits can be changed, inflectional morphemes can be transferred, TA categories and systems, basic word order, and much else, can be transferred.

It might be objected that this new scenario is based on recipient languages which are small, languages with relatively few speakers - which would be true - and on donor languages which are large or larger, and so is not a model usable elsewhere. I think this is a non-objection. First, most language communities during most of the world’s history have been relatively small: consider Britain which has a population today of just over 50 million, but had just over one million inhabitants a thousand years ago at the time of the Norman invasion - what if we went back another millennium or five or ten? So these contact phenomena in a place such as Africa, where most communities were small until recently or are still small, are typical of most language communities during most of the world’s linguistic development over tens of thousands of years. But second, from the work of Labov and others in big, very crowded, American cities, it is clear that change, once established, is spreading across large communities almost before our eyes. Size of community is not too significant. Or maybe the distinction between large and small communities is misleading. Any large community - the USA, or Pennsylvania - consists of a row of small communities. An individual or group of individuals regard themselves as members of several interlocking communities - a smaller local community, and some number of larger communities. Language is a badge of group membership. How quickly and how far change spreads reflects, inter alia, how speakers feel about membership.

In summary it may be true that there is a probability hierarchy of features that can be transferred. But it is not as fixed as we once thought and the incidence of features other than lexicon being transferred is higher than we once thought.

One question posed at the BLS meeting was: can the outcome of contact between two, or more, languages be predicted on the basis of linguistic inputs alone? The predictive power of any model depends on how good the model is. Most of
our current models are based on an inadequate analysis of partially described or partially understood languages and situations. When we get to the point where our linguistic models are adequate, will be able to predict the outcome of language contact? This is like trying to predict what English or any other language will look like three centuries from now. In purely linguistic terms, we can talk of probabilities. We cannot talk of good probabilities yet, because no one has put the contact question in those terms or sought an answer. TK, for example, examine many cases of contact and make generalisations. That is different from trying to develop a predictive model. I think a probability model with good predictive value could be developed if we focused on that.

The obvious obstacle, however, to predicting the outcome of linguistic outcome is the non-linguistic component, the actual circumstances under which language or dialect communities meet. We know generally what the variables are: relative size of the communities, degree of bi- or multi-lingualism/-dialectism, language use, prestige of the languages involved, length of contact, and so on. Any particular situation contains its own particular mix of the linguistic and the non-linguistic variables and at present we are some distance from being able to predict the total outcome with any degree of certainty.

4. TRANSFER; BORROWING, LANGUAGE SHIFT, PIDGINS, CREOLES. Traditionally, through much of the twentieth century, the main contact model was that which led to “borrowing”, although notions such as substratum implied the possibility of shift. What was not inherited was borrowed but, since the historical field was dominated by Indo-Europeanists, explanations relying on borrowing - other than of vocabulary - were regarded as dubious. The advent of pidgin and creole studies in the 1960s and the discovery of a growing number of cases unlikely to result from borrowing made these assumptions increasingly untenable.

TK outline three or four types of contact situation: those that lead to borrowing, to language shift, to pidgins, and maybe to instant creoles. Since they say that after some time pidgins and instant creoles will probably be indistinguishable, that reduces the task of the historical linguist to looking for traces of three only.

TK raised at least two exciting possibilities - that of finding new cases of pidgins or creoles in areas other than those well known (West Africa, the Caribbean, Papua New Guinea, etc), and that of finding numerous cases of language shift. After all, during the decades prior to TK, new creoles had been found in Africa. It was known that the main dialect or form of several languages was or was until recently a pidgin and that many well known languages had pidgin varieties. Cases of language shift had been documented.

However, it seems to me that by and large the number of cases in Africa uncovered since TK and pointing unambiguously to pidginization, creolization, or language shift has been disappointingly small. In my own project, mentioned above, where I actively looked for such cases, I found it hard to identify any clearly. As far as I am aware, the literature in general on Africa in recent years has produced, and continues to produce, new cases of contact but explanations pointing to language shift or pidginization as the exclusive or main mechanism have been few.

If these observations are correct, then why is this? It may be that TK is a recent work, and that not enough time has passed for researchers to unearth new cases, especially in a continent where active field work is not expanding. Or it may be that we are looking at the wrong languages. In the case of my own investigation, for
example, I was aware of Bantu languages whose genetic affiliation had been clearly
distorted and thought that examining them would lead to uncovering evidence of
new historical situations. It did but it did not lead to clear cases of pidgins, creoles,
or language shift. Or it may be that our general models are still not powerful enough
and thus we still have to know the historical circumstances in individual cases in
order to identify the correct transfer process - but we usually do not know them.

Or it may be that "borrowing" is simply more frequent than other forms of
contact, or at least to other forms of contact that leave a trace. Language shift has
occurred often in this world but for it to leave an identifiable trace, at least two
conditions have to be met. One is that it is not enough for individuals or small
groups to give up their language because typically such an action leaves no trace as
the people involved sink anonymously into the mass of the host community. A
large coherent group is a prerequisite. The second is that the shifting group must
stay together as a cohesive entity for some time after the shift and retain something
of the linguistic features they have brought with them. Thus, in North America, the
Jewish community in New York and the Afro-American community are large and
retain distinctive linguistic traces whereas the descendants of countless other
millions of immigrants who moved as individuals or in families lost their former
speech without trace. Likewise, pidgins are few because they derive from very
particular circumstances which are/were relatively uncommon. By contrast,
"borrowing" derives mainly from two - language or dialect - communities being
side by side for a long time, this is a relatively common situation, thus borrowing is
the commonest form of transfer.

Further, as pidgins develop into creoles and as creole systems become
elaborated, it becomes increasingly hard or even impossible to identify their traces
after the passage of some time, perhaps several centuries. Similarly with shifted
languages. When a community shifts its primary language, the result will most
often be a "dialect" form of the new language. As long as the historical
circumstances are remembered, and as long as some record or form of the original
language are kept, the origins of the new "dialect" will be recognized. But once the
memory and the original language are gone, the new "dialect" will be wrongly
thought to have derived genetically from an old protolanguage, just as the other
dialects of the language.

J. McWhorter suggested (p.c.) that rather than looking for traces of fully blown
pidgins or creoles, it would make more sense to look for traces of pidginization or
creolization as processes, by identifying sets of features. In principle this makes
sense but suffers from the practical drawback that the sets of features that
characterize the different models often overlap and after a few millenia are hard to
identify. One way to distinguish, for example, the results of borrowing from those
of language shift is they follow different chronologies of transfer: for instance,
structural transfer occurs rather late in the borrowing hierarchy but early in language
shift. But since the historical circumstances surrounding most of the world's
languages are unknown, and since they do not have written records, the chronology
of transfer is a matter of conjecture. It is tricky to deduce from the linguistic facts.
Finally, it may happen that several processes are involved. The small community at
Brava in Somalia, which speaks Mwiini, the most northerly mainland Swahili
dialect, has been surrounded by the much larger Somali-speaking community for a
millennium or so. During that time, it repeatedly absorbed Somali-speakers - thus
language shift - and was influenced by the surrounding community - so borrowing.
There is at least a possibility that in earlier times it was used as an areal trade
language - thus pidginized (Nurse 1991).
This perhaps overemphasizes the difficulties involved and sounds too pessimistic about the chances of success. Success is possible but it requires careful analysis of the linguistic features of a given contact situation and of as much as can be discovered about the real or assumed historical sociolinguistic circumstances.

A corollary to this is that our current models for analyzing historical language contact are inadequate and still evolving. 3

5. CHANGE CAN SPREAD QUICKLY. We know now from the work of Labov and others that sound change can spread quickly across a monolingual community. A sound change or a set of changes, once established, can sweep across a community or communities in a generation or two.

Over the past few years I have been investigating various small communities in East Africa, communities speaking languages (Ilwana, Daiso, Sonjo, Mwiini) known to have undergone quite massive changes under the influence of neighboring languages. In some of these, not all, the period within which the changes took place can be established. In these languages changes can also apparently move quickly.

For two of these languages (Ilwana, Daiso) it can be reasonably shown that massive changes occurred in at most three centuries, changes so massive that if speakers of the two stages of the languages could be brought face to face they would hardly understand each other. While three centuries is of course more than Labov’s “a generation or two”, it is also possible that these changes took less than 300 years. It is clear that they could not have started before a certain point or finished after another certain point and that these points were 300 years apart but the changes may have been accomplished in part of that period.

Even if they did take as long as some part of three centuries, the total set of changes is much greater than that of Labov’s vowel shifts. They include a radical enlargement of the phonemic inventory, a significant change of articulation habits, of phonological habits, a radical restructuring of the inherited TA categories using inherited morphology, and a host of smaller changes. In both these languages a large majority of their vocabulary has been borrowed (in one case over 50%, in the other 75%). All these changes push the languages involved in the direction of the surrounding donor language(s).

The changes also apparently differ from those of Labov and his colleagues in that they were externally induced by a “foreign” language.

It is apparently unimportant whether the source of the change is internal or external. Changes from both sources, or maybe change from any source, or maybe just “any change”, can move quickly. In historical linguistics we do not necessarily have to assume that great changes took a great time. One thing that does need to be established is how far, across how large a community, or across how many adjacent communities, a change or a set of changes could spread quickly before running out of steam. A change that spread 100 miles in Western Europe or in post-Contact America would most often affect dialects of the same language whereas the same distance in Africa or pre-Contact North America would cross language boundaries.

6. THE TRANSFER OF MARKED FEATURES. An issue often raised in the literature is the fate of marked features in language contact. TK refer often to this and give a summary of the debate. A central difficulty is how to define the notion of
“marked” feature. Marked features are normally said to get lost in language contact. Consider two cases.

The first is Ilwana, spoken in north eastern Kenya, which at some point during the course of the three centuries just referred to, introduced two new features: a set of ejective consonants and a distinction between implosive and explosive stops. The second is the Bantu languages of South Africa, which at some point introduced clicks into their systems, where none had existed before.

From the viewpoint of the typologist, both these cases involve the introduction of sets of new sounds often considered typologically marked. Consider them rather from the point of view of the people speaking the language, people who had not heard of typology. In the one case, a community is surrounded by communities speaking languages with ejectives and having a distinction between implosive and non-implosive consonants. Not only that, there is good reason to think that the Ilwana became bilingual in the major - and prestigious - donor language. In the case of South African Bantu, two scenarios have been suggested. The traditional one is the offspring of Bantu fathers and Khoesan mothers being brought up by Khoesan mothers or nannies. Although the offspring eventually spoke the languages of their fathers, it is not surprising that they were much influenced lexically and phonetically by what they heard all around them as young children and also in later years as the same mothers or nannies would have continued to speak Khoesan. A newer suggestion (Herbert 1993) has (presumably) Khoesan-speaking women practising hlonipha in Nguni-speaking households: hlonipha is respect through avoidance, and in this case women substituted clicks in forbidden words or syllables. Both scenarios involve bilingualism over a lengthy period - Herbert speaks of “three to five centuries” of intense Bantu-Khoesan interaction.

In both cases, speakers were simply adjusting to their linguistic, their typological, surroundings. Whether or not these involved “marked” features, they were the local norm.

In thinking about all this, I remembered a conversation I had in the 1970s with a baker-poet in Lamu, Kenya. He spoke what was once a prestige dialect of Swahili, now encroached on by Standard Swahili. I was doing field work on Swahili dialects in the area and was rather naive. I expected linguistic purity, I wanted pure dialect forms. He allowed me to record some of his poetry and analyze it. When I found Standard, non-dialect, forms intruding, I reproached him with this and wanted to change it. He did not want it changed. He acknowledged the intrusion of Standard forms but said that his purpose was communication. He wanted to communicate with the people of his town, said they all increasingly spoken like that, and so be it. The mixture was becoming the local norm. I think he would have chuckled at some of our views of typology.

7. FAMILY TREES: WHAT IS SWAHILI (HISTORICALLY)? Today’s Standard Swahili is a twentieth century development initially based on Unguja, a dialect of Zanzibar Island and a form of southern Swahili, one of the two historical branches of Swahili. Swahili in turn is one of the five or six branches of Sabaki, in turn one of the branches of the North-East Coast (NEC) Bantu languages. Three nodes are involved: Proto-NEC, -Sabaki, and -Swahili. Within a group of related languages, a new branch is most reliably defined by some set of (non-lexical, phonological) innovations. We traced the line that leads down though these nodes to today’s Unguja and found that that line is hardly ever defined by its own innovations, it is most often defined by other parts of the group(s) innovating and by Swahili and southern Swahili having stood still. Both are historically largely
defined by what others did linguistically, not by what they did. Is this just an accident, or does it need an explanation?

A branching in a family tree is normally taken to represent a language splitting into two or more, but how would it be interpreted in human terms, as a community dividing? Or rather, if a community is already divided in some non-linguistic ways, how does subsequent linguistic change fit in? One possible but speculative interpretation is via language as part of self identity. Part of how a community identifies itself as a community is by sharing some set of linguistic features, either internal innovations or transfers. This act of self-identification is not what induces the change in the first place but once the change has started, the community adopts it, as it were. Of all the NEC and Sabaki groups, that which spoke Swahili was demonstrably the most “successful”. Could that be taken to mean “the most at ease with itself, the most content”? If so, then perhaps it is a case of “let others redefine themselves by innovation, we stand here”? Or maybe there is no connection between phonological conservatism and non-linguistic events?

Another possible, more concrete and more plausible explanation might have to do with geographical isolation. Original Swahili settlements were all or nearly all on islands or tucked away by the shore, remote from larger mainland population centers. Change(s) which started on the mainland didn’t reach into Swahili. Did geographical isolation simply lead to linguistic insulation?

8. CORRELATING LINGUISTIC AND ARCHAEOLOGICAL PROFILES. An approach that has been used in Africa in uncovering details of earlier historical periods is the juxtaposition of linguistic and archaeological distributions. Plotting the geographical distribution of linguistic groups and archaeological strata on a map sometimes shows remarkable congruity between the two. If such congruence can be found, we can tentatively assume identity between the linguistic group(s) and the people who left behind their material remains. Progress can then be made. Linguistically, we can talk about spread, either of people or of language, and we can reconstruct details of earlier culture through reconstruction of vocabulary. This can complement the much greater detail provided by archaeologists. Combined use of the two may enable us to date linguistic events.

In today’s circumstances this methodology may prove to be appealing. In several parts of the world we see ethnopolitical change. What were large political blocks are being replaced by local nationalisms. These local nationalisms are assertive and in search of an identity, of which a central part is their own history. Often local memories are short and this local history has no written record and has to be constructed. Linguistic and archaeological evidence provides a way forward.

In Africa, this approach has had some striking successes. One is with language families in general in Africa. Despite disagreement about many details and sometimes about major issues, there are some solid achievements in identifying earlier archaeological-linguistic communities at the level of family. Another is with Bantu in general, where there is almost a general one-to-one geographical correspondence between the overall domain of Bantu languages and certain archaeological distributions. A third is with Swahili. The present and former distribution of Swahili-speaking settlements along the East African coast correlates well with archaeological evidence for a culture with a specific building style, economic mode, maritime connection, and eventually religion. That allowed us to link the two and thus reveal details of past Swahili society. There have also been some striking failures. The history of the Bantu-speaking peoples has been characterized by well meant misjudgements. Guthrie’s linguistic work in the 1950s
and 1960s was used by historians.⁴ as the basis for positing prehistoric migrations. Other linguistic work done in the 1960s and 1970s was then used by other historians to support another model of Bantu dispersal.⁵ In both cases the linguistic picture and the historical interpretation have been largely discredited. Most recently, a new model of dispersal is based on a third set of linguistic data.⁶ Common features run through these and other juxtapositions. One is the quality and nature of the linguistic material. This consisted most often of classifications based on lexical measurement alone, and on incomplete geographical coverage. A second is the archaeological coverage, which also often suffered from incomplete geographical coverage. A third was that the relationship between different ceramic types, often central to the archaeological interpretations, appears to have been not always well understood. The fourth and major problem area was in interpreting the relationship between linguistic and archaeological evidence. At the macro-level, the correlation between linguistic and archaeological distributions has worked reasonably well. But how to interpret whether this related to the movements of people or languages? Should a split in a linguistic tree necessarily be interpreted as involving a movement of people? At the micro-level, how to relate pot types to peoples? There is no reason why ceramic and linguistic distributions should be coextensive. And how to interpret the history of sites which have a long settlement history and were archaeologically and maybe linguistically mixed?

Despite the caveats above, these are in fact exciting times for historical linguists concerned with the human component. New cases appear frequently, new ideas are evolving and need testing. Old languages are disappearing and need recording. There is much to be done.

FOOTNOTES

1 Comments by V. Bubenik, L. Hyman, J. McWhorter, H. Muzale, H. Paddock, and M. Silverstein have caused me to modify the original text. To them my thanks are due.
2 See Brenzinger, M (1987) and Mous (forthcoming)
3 The work done in Africa by C. Myers-Scotton deserves more attention than I have given it here, for example.
4 Oliver 1966.
6 Vansina 1995.

REFERENCES


Mous, Maarten. Forthcoming. Was there ever a Southern Cushitic language Ma’a?. To appear in the proceedings of the 3rd Conference on Cushitic and Omotic Languages.


-------. 1993. Inheritance, contact, and change in Ilwana. Memorial University, ms.

-------. 1993. Inheritance, contact, and change in Daiso. Memorial University, ms.


Internal vs. external factors in socio-historical explanations of change: a fruitless dichotomy?

Suzanne Romaine
Merton College, University of Oxford

INTRODUCTION

1. Stephen Jay Gould (1984:7) wrote that 'dichotomy is the usual pathway to vulgarization. We take a complex web of arguments and divide it into two polarized positions'. On this occasion Gould was writing about what he called the 'false antithesis between nature and nurture', but his remarks are equally applicable to the dichotomy which concerns me here: namely, that between internally vs. externally motivated linguistic change. This and other dichotomies have been part of the structuralist legacy of Saussure inherited by contemporary linguistics.

An early manifestation of frustration with this traditional dichotomy can be found in Weinreich, Labov and Herzog's (1968) discussion of empirical foundations for a theory of language change, where they emphasized the need to use the external history of a language in historical reconstruction and to abandon the dichotomy between synchrony and diachrony. This led to a view of language change embedded in both linguistic and social structure. We might conveniently cite their article as the birth of socio-historical linguistics although the authors do not use that term. Their so-called 'actuation problem', however, is the very heart of the matter since it deals with the motivation or causation of change. Nevertheless, one can still say that while attention is paid to external factors constraining and facilitating the implementation of change, they are still seen as separate from internal factors.

A recent manifestation of this can be found in Labov's three volume Principles of Linguistic Change (not yet completed), where he divides up the volumes as follows: Volume 1 deals with internal factors; Volume 2 with social factors, and Volume 3, cognitive factors. He (1994:1) comments that the separation of internal from external factors may not 'seem practical to those who view language as a unified whole where tout se tient, or to those who believe that every feature of language has a social aspect.... it is reasonable to ask whether internal factors can be successfully separated from social factors'. He then goes on to argue that the rationale for the separate discussion of internal and external factors is their respective independence of one another.

Speaking of the results obtained from multivariate
models of analysis used to investigate change, it appears to Labov that ‘if an internal factor is dropped or changed, changes appear in other internal factors, but the external factors remain unchanged; if an external factor is dropped or changed, other external factors change, but the internal factors remain as they were’ (1994:3). He adds that the ‘internal factors are normally independent of each other, while the external factors are heavily interactive’. Although Labov (1994:4) is careful to claim he is not offering a ‘theory of language change’ in the three volumes, he nevertheless, ventures that to ‘explain a finding about linguistic change will mean to find its causes in a domain outside of linguistics’.

In just these few remarks we can see the simplistic way in which most sociolinguists have treated the ‘social’ as a set of separate and independent categories or factors such as social class, sex, style, etc., which can be correlated with bits of language. These correlations reveal little about the dynamic processes that bring them about in the absence of some more general theory about social behavior.¹

With nearly 30 years of empirical research within this paradigm, it is perhaps time for a reassessment of how successfully we have dealt with integrating the internal and external into our explanations of language change. To do so, I will look at approaches which have focussed on one or the other of the two poles, as well as those which aim for some degree of synthesis. While there are problems with all of them, I will suggest some ways forward.

SYSTEM-INTERNAL EXPLANATIONS FOR CHANGE

2. I will begin by looking at approaches which rule out the contribution of system-external factors. With respect to the distinction between internal and external factors, Saussure permitted no compromise (1959:22):

’I believe that the study of external linguistic phenomena is most fruitful; but to say that we cannot understand the internal linguistic organism without studying external phenomena is wrong... One must always distinguish what is internal and what is external. In each instance one can determine the nature of the phenomenon by applying this rule: everything that changes the system in any way is internal’.

While I believe Saussure is right that one CAN study what he calls the ‘internal linguistic organism’ without studying external linguistic phenomena, the result is not satisfactory. As Meillet (1916:66) pointed out in a review of Saussure’s Cours, by separating linguistic change from the external conditions on which it depended, Saussure reduced it to an abstraction which was
essentially inexplicable. If languages were not used, there would be no reason for them to change. More importantly, however, Saussure's injunction that we must distinguish between internal and external using the criterion that whatever changes the system in any way is by default internal, begs the question of what one means by 'system' and how one identifies a change in it. It also reflects a traditional bias against external factors which are seen to be inherently messy; only internal factors were systematic and therefore worthy of treatment.

The predominance of the generative paradigm in both mainstream and historical linguistics has been partly to blame for the fact that external factors are still generally appealed to only as a last resort for anything that could not be explained internally. Generative approaches have conceived of change variously as radical restructuring triggered by opacity (see e.g. Lightfoot 1979) or more recently as parameter resetting (see e.g. Lightfoot 1991), with the major role given to children's acquisition as the source of structural change. Parameter resettings constitute change in individual grammars. Lightfoot attempts to figure out what factors triggered a particular resetting by looking at the surface data before and after the reparametrization. The problems he encounters are made more complicated by the incomplete nature of the written historical record and its unsuitability for answering with any decisiveness questions about what the input to learners must have been at a given point in time.

The dichotomy between internal and external motivation also informs the thinking of contemporary historical linguists who are not generativists, such as Lass (forthcoming, 6), who has quite recently reiterated the separation between internal and external when he writes 'we can (and indeed must) say quite properly that the structural history of a language ('linguistic history' in the strict sense) is quite independent in principle of its social history. The story of a language 'itself' must be carefully distinguished from the story of its changing uses, users, and social context - just as the changes themselves (as results) must be distinguished from the mechanisms by which they came about (e.g. lexical and social diffusion). The two are related in subtle and complex ways, but the relation is never 'causal' in any philosophically respectable sense.' He later says (forthcoming, 7) that it is a 'vulgar error to talk about the 'social causation' of changes in linguistic structure'. Lass (1980:121) claims that where speaker-based 'external' (i.e. social) factors have been appealed to in historical linguistics, these have been
'superficial and otiose', often involving ad hoc appeals to 'unanalyzed social categories' such as prestige.

In a different discussion of endogeny (language internal origin) and exogeny in varieties of English, he (1990:48) further elevates the dichotomy to an axiom when he says whenever there is a choice 'between (demonstrable residue) and (putative) contact-influence, the former is the more parsimonious and hence preferred account'. Thus, he concludes that Hiberno-English is a 'perfectly normal first-language internally evolved variety, with only marginal contact effects'. In a similar argument about the putative influence of Afrikaans on South African English, he and Wright (1986:202) observe that the normal condition for a language-system is to be a self-contained structural network. Its degree of permeability to outside influences is inversely related to the 'structuredness' of the subsystem involved.

While it is always easy to SUSPECT borrowing if we know there has been contact between two languages, it is much more difficult to prove that it is actually responsible for a given syntactic pattern or construction. It is always possible that the form in question evolved independently. Part of the problem in coming to grips with this issue is that we have yet to develop a rigorous theory of language contact which allows us to make reliable predictions about what can be borrowed under what circumstances. The traditional position has been that while the lexicon is highly receptive to borrowing, syntax is less so. This too is an oversimplification.

Thomason and Kaufman (1988) have proposed a methodological criterion for establishing external causation of language change, namely, that if a language has undergone structural interference in one subsystem then it will have undergone interference from the same source in others as well. They believe that interference can occur in all subsystems of a language. The social context rather than the structure of the languages determines the extent and direction of interference. Lightfoot (1979:380) too provides an additional criterion; to prove that external influence produced a syntactic change, we must prove that the innovating forms were first used by bilingual speakers.

Such discussions often simplify the issue since contact does not manifest itself solely in terms of transfer of features from one language to another. Other effects of contact may be simplification or regularization of irregularity in paradigms, reduction of variants or options, overgeneralization, or an acceleration of change already underway. Sufficient studies exist to make it clear that the phenomenon of
language contact cannot be separated from the issue of decreasing usage. Where restriction in use occurs, the frequency of irregular and marked forms falls below a critical threshold and makes it less likely that these elements will be acquired by younger speakers. Dorian’s (1993) discussion of the complexity of the factors which must be taken into account before one can meaningfully assess the question of internal vs. external motivation sounds a cautionary note.

It might be argued that I have muddied the waters by bringing language contact into the picture under the domain of external factors motivating change, but this reflects a more general lack of clarity in the literature as to what constitutes internal vs. external factors. Traugott (1994:492) points out in her review of Gerritsen and Stein's (1992) edited volume entitled Internal and External Factors in Syntactic Change that the papers therein make but a meager contribution to a theory of internal and external factors. She identifies as the reason for this the artificial dichotomy between the two sets of factors which has been part of the ideology of linguistics since Saussure. While Gerritsen and Stein urge a reassessment of the dichotomy, the problem is ultimately given short shrift.

Certainly what counts as internal or external depends on a particular theory. Stockwell and Minkova’s paper (On the role of prosodic features in syntactic change), for example, does not consider the ‘external’ dimension at all, if what we mean by that is the consideration of factors outside the abstract language system internalized by speakers. What they take to be ‘external’ is the implication of the prosodic level in syntactic change. From a larger perspective of course both prosody and syntax are still system-internal by comparison, say, with the impact of social processes such as standardization and literacy on a language more generally.

As the papers in this volume demonstrate, the distinction between external and internal is far from straightforward and is to some extent muddled by many other terms often used in the explanation of change, e.g. ‘natural’, ‘marked’, ‘autonomous’, ‘cognitive’, ‘social’, etc. Moreover, one could argue that the borderline between what is regarded as internal and external is continually shifting. The role of an innate universal grammar is being progressively reduced by more sophisticated work in pragmatics and cognition with the result that a number of abstract and fundamental grammatical principles are being shown to derive from more general processing constraints and pragmatic considerations. While processing principles are often
thought of as no less innate than grammatical ones, they are generally regarded as system-external by generativists. As Gerritsen and Stein (1992:8) point out, the dichotomy between internal and external factors itself rests on another dichotomy, namely that between the language as abstract structural system on the one hand and its use on the other. Traugott concludes her review by saying that the question of what is internal and what is external will remain unresolved until we reach a better understanding of factors related to use or pragmatics and of how to theorize continua.

SYSTEM-EXTERNAL EXPLANATIONS FOR CHANGE

3. A few linguists have migrated to the opposite pole of the dichotomy between internal and external factors in their search for motivations for change. I mentioned Meillet (1906:17), for instance, who argued that since language is a social institution, linguistics must be a social science and 'the only variable element to which one may appeal in order to account for a linguistic change is social change'. More recently, Milroy (1992) has proposed a speaker-based model of change, in which there is no system-internal motivation for change. All change is socially motivated, introduced by acts of speaking on the part of individual speakers, and then diffused through social networks of various types. Milroy (1993:216) states quite bluntly that 'change would not take place if the speakers did not in some way agree that it should take place'. This apparently conscious intentionality distinguishes Milroy's position from that of the Neogrammarians, who believed in the randomness of sound change. Labov (1994:604) too finds the Neogrammarian view 'essentially correct. ...structure is a largely mechanical system, out of the reach of conscious recognition or adjustment by its users'.

The problem for Milroy is to explain how these speaker-based changes get into the abstract structure we call language. Weinreich et al. do not have this problem since they make no distinction between a change and its propagation, nor between a change itself and an analysis of its mechanism. Milroy, however, invokes a distinction between speaker innovation and language change. Thus, speakers innovate, while languages change. More specifically, innovation is an act of the speaker capable of influencing linguistic structure. Yet, innovations occur all the time and many will not enter the language system at all. Once they do enter the system, however, they follow the predictable paths through social and linguistic structures, which, thanks to the research program articulated by Weinreich et al. (1968), we now know a great deal about. The question then
becomes: under what social conditions do innovations become change? There is another question of course, but it is ignored. Milroy says 'we are not asking how spontaneous innovations arise'. He goes on instead to say that linguistic change is a social phenomenon that comes about for reasons of marking social identity, stylistic differentiation, etc. If it does not carry these social meanings, then it is not a linguistic change.

Milroy's answer to the question about the social conditions appropriate for change is found in network structure. Members of dense, multiplex networks impose norms of behavior which restrict change. In dense and multiplex networks everyone knows everyone else in a range of capacities. Integration in such a group involves an acceptance of its norms. A closeknit network thus functions as a conservative force resisting pressures for changes originating outside the network. Where ties are loose-knit, people are susceptible to change. The innovators themselves are individuals weakly linked to a network and likely to be in a position to contract weak ties.

Milroy extrapolates from the findings of his work in Belfast to make claims about rates of change in whole languages, predicting, for example, that linguistic change progresses most slowly in tightly knit communities which have little contact with the outside world. He offers the conservatism of Icelandic as a case in point. However, that is true only for grammar and lexis. Another factor he does not mention which I feel is relevant is that Iceland has the longest unbroken tradition of writing in all the Scandinavian countries. Iceland also shows much greater purism in its policy towards foreign borrowing. Indeed, grammatical and lexical conservatism became a matter of national pride and deliberate policy in the 19th century as a symbol of resistance to Danish domination.

Milroy has conceded that relative geographic isolation may have as much to do with the uniformity of Icelandic as the preservation of close knit network structures. 'Relative' is a key qualifier here in considering the respective contributions of isolation on the one hand and network structure on the other. Compared to Faroese, its closest relative, Icelandic evidently has far less phonological dialect variation (see Arnason 1980:60), and has preserved more of the proto-Scandinavian morphological system. Both Faroese and Icelandic, however, are certainly more conservative than the mainland Scandinavian languages. By definition, change in a totally isolated language has to be endogenous, but this still raises the question within Milroy's framework of where within these isolated tight-
knit communities weak network ties exist. If we accept that it is speakers who introduce innovations which may lead to change, then contact between speakers is necessary for change to spread.

Recent evidence from Dorian's (1994) work among Scottish-Gaelic speaking enclaves on the east coast of Sutherland make certain of Milroy's assumptions about the normativeness of dense, multiplex networks and their capacity to resist change questionable. The three Gaelic-speaking villages in question are exceptionally homogeneous. Networks are virtually all kin-based, so individuals interact in the capacity of kin, friends and neighbors. Yet Dorian found thirty some cross-generational intra-village differences which did not correlate with the usual external variables of socio-economic status, network, sex, and style, and are largely ignored by the villagers. She (1994:633) concluded that a 'profusion of variant forms can be tolerated within a small community over a long period, without discernible movement toward the reduction of variants and also without the development of differences in the social evaluation of most variants'. This profusion of variant forms might be equated with innovations which do not spread from one speaker to another and therefore do not make their way into the language system. For some reason, these variants do not attract social attention and are not then invested with the identity marking functions Milroy specifically links with change.

The explanation for what Dorian calls 'personal pattern variation' does not lie in the fact that Gaelic is dying in these communities since these particular instances are not tied to age or proficiency either. There are other cases of variation of the familiar type, related, for instance, to geography, which the villagers are keenly aware of even though they are no more or less salient from a linguistic point of view. We know from other studies that the rate of change in languages spoken by small linguistic groups can be very rapid, and is potentially much greater than in languages spoken by large industrial societies. Certainly there is rapid change going on with respect to other features of the language in these villages. These may be attributed more generally to the process of restriction.

Trudgill (1992) has elaborated the dimension of relative isolation in addressing the question of whether different types of societies produce different types of language structure and language change. Taking the Scandinavian languages again as examples, he wonders why Faroese has remained more conservative in its grammatical structure than Norwegian. As Trudgill points out, the Faroe islands are obviously more remote and isolated than
Norway. This gives rise to the hypothesis that varieties in contact change more quickly than those in isolation. Moreover, the kinds of changes that take place in contact languages often make them more simplified, regular and analytical in structure—pidgins being an extreme case. Yet similar trends can be seen across whole families of languages such as Indo-European. Conversely, other types of changes, such as fast speech rules, which make phonology more opaque to second language learners, are less likely to become institutionalized in high contact languages.

These ideas are, of course, not new. Trudgill would have found it useful to consult Thurston's (1987) study of language change in northwestern New Britain, Papua New Guinea, where the author tries to validate the existence of two kinds of language change, one gradual and the other abrupt, motivated by different social conditions. Like Trudgill, Thurston believes that centuries of ingroup use in small communities leads to the acquisition of complexity. Some of this complexity may manifest itself in the personal pattern type of variation Dorian found. Conversely, languages spoken by strangers to one another are simplified to a degree corresponding to the extent of their speakers' shared knowledge. The most radical change tends to occur in languages undergoing intense contact. It is also such languages which tend to display an unusual amount of what Grace (1992) calls 'aberrant' change, i.e. not easily explicable in terms of traditional comparative historical reconstruction.

These ideas receive some support from quite a different quarter when we look at the geographical distribution of diversity over time. In a far-reaching attempt to quantify typological structures and correlate them with patterns of migration, Nichols (1992) has identified 'spread zones', characterized by low structural diversity, innovating center and conservative periphery (e.g. Australia, W. European, central Oceania, Mesoamerica). Residual zones are high in genetic density (the ratio of genetic lineages per square mile) and structural diversity (amount of disparity exhibited by a language or group of languages). The island of New Guinea is an example; it alone contains about as many language families as the entire Old World.

The basic notions of historical linguistics apply well to spread zones because they have been worked out with reference to spreading language families such as Indo-European, and therefore less well to 'aberrant' languages. Spreading languages replace or absorb their predecessors and cause an increase in uniformity. There is a tendency for unmarked features to spread in situations of high diversity and language contact. In
this way a lot of genetic density has been lost in Europe through the spread of Indo-European. Western Europe is now dominated by language stocks that originated east of the Urals: namely, Indo-European, Turkic and Uralic. Only Basque and the indigenous Caucasian languages continue pre-Indo-European lineages. Basque is a good example of an extreme relic population and language. The Basques underwent the least genetic admixture with the invading Indo-European-speaking peoples. Today the Basques are both genetically and linguistically distinct. In the peripheral parts of the world, western Europe, central and southern Australia, eastern North America and parts of South America, we have extreme and idiosyncratic profiles.

Diversity appears to stabilize over time. There is clear evidence for west to east order as a major predictor of the frequencies of language structures. Australia represents the eastern extreme and constitutes the first split after the Old World from the tree of typological divergence. This is compatible with hypotheses of a west to east expansion which went first from Africa and the Near East to the tropical Pacific. Western structures gradually expanded eastward and later entrants to the New World showed more Western admixture than earlier entrants. A reduction in diversity is a product of modern times, more specifically of social and political factors such as the rise of complex economies, urbanization, etc., and wide spread of languages driven by economic and political prestige. In Nichols’ scenario the greatest degree of complexity is in areas of high diversity and contact. The Pacific emerges as an ancient center of typological differentiation.

CONCLUSION

4. What can we conclude from my brief and rather sweeping examination of the problems in separating the internal from the external in our search for the motivations for change? Labov (1994:5) has commented: ‘a set of propositions that relate general findings about language change to general properties of human beings or of human societies will certainly deserve to be called a theory of language change’. I believe we may be on the verge of sketching the outlines of such a theory if we can integrate some of the strands of research I have mentioned here. A number of things are required in order to do this. At present, socio-historical reconstruction is constrained by the fact that the external histories of most languages are unknown. Horvath and Wexler (1994:264) argue that the uncovering of these histories should be a major desideratum of historical linguistics. Presumably more communities were historically of the focussed type with high density networks in which people interacted
mainly with others whom they had known most of their lives, generally shared the same knowledge, values, etc., but not necessarily the same speech repertoires. There is an urgent need for the comparative study of language variation and change in small scale societies which would allow us to contrast relatively isolated communities with those under intense contact. I cannot stress too much the urgency of this task since the world's small languages are disappearing, and with them the kind of complexity that develops in such intimate communities.

FOOTNOTES
1. Students of gender have certainly questioned the nature of the relationship between two common social variables, namely, sex and social class. After reanalyzing a number of sociolinguistic patterns, Coates (1986) concluded that gender quite often provided a better account of the variation than did social class. Williams (1992) offers a more thorough-going critique of bias within sociolinguistics towards social theory based on Parsonian structural functionalism, emphasizing individualism, identity, consensus and mobility. The belief that change is functionally adaptive is also part and parcel of the Parsonian legacy.

REFERENCES


The declension of ethnonyms in English
Kevin Tuite
Université de Montréal

1. Introduction. The morphology of the English language has undergone a
tremendous erosion in the course of the past millenium. In Anglo-Saxon, as is well
known, nominals were partitioned into numerous classes and subclasses according
to gender and declensional characteristics. Very few traces of this formal diversity
remain in the modern language, and what does remain is stored in the lexicon as
morphological irregularity. Yet it cannot be said that declensional diversity of a
systematic sort has disappeared from English. In this paper I will discuss one class
of English nouns which presents an interesting variety in terms of its inflection for
number; moreover, the morphological properties are correlated with certain other
characteristics, some formal and some semantic. The nouns to be analyzed here are
ethnonyms, that is, designations of persons in terms of their ethnicity or place of
residence. Contemporary English ethnonyms can be divided into four principal
groups in terms of their declensional properties:

2.1. Type I (e.g. ‘German’). The first group comprises those ethnonyms with
the same declensional properties as most common nouns. Many of these are
derived by the suffix -an or -ian (Mayan, Italian, Russian, Acadian), but a large
number are morphologically unanalyzable in English (Czech, Finn, Arab, Gypsy).
Type I ethnonyms have these characteristics:

(a) A regular singular/plural opposition:
    A German is ... / The Germans are ... / Five Germans are ...
(b) The /'s/ genitive is used:
    A German’s sense of humor / The Germans’ sense of humor

2.2. Type II (e.g. ‘Irish’). The ethnonyms of the second group are usually
considered to be adjectives utilized in some sense as nouns. Type II comprises the
following nine ethnonyms formed with the suffix of Germanic origin variously
Type II ethnonyms are characterized by these properties:

(a) Only plural-collective usage; no formal singular.
    *An Irish is ... / ?Five Irish are ... / The Irish are ...
(b) The /'s/ genitive is not possible:
    *The Irish’s sense of humor [cp. adjective: The Irish sense of humor.]

The formal class of Type II ethnonyms does not include those adjectives in -ish
formed from Type I ethnonym, which cannot function as nouns, except when
referring to a language: e.g. Polish, Danish, Finnish (cp. Pole, Dane, Finn, etc.)

2.3. Type III. The third group comprises those ethnonyms with formally
identical singular and plural forms. I have further divided it into two subgroups:

2.3.1. Type IIIa (e.g. ‘Japanese’). This group is formed exclusively of
ethnonyms formed by the suffix -ese of Romance origin. According to the OED,
" -ese forms derivatives from names of countries (chiefly after Romanic
prototypes), and from some foreign (never English) towns, as *Milanese, Viennese, Pekinese, Cantonese.*” For Type III ethnonyms:

(a) Singular and plural are formally identical:

*A Japanese is … / The Japanese are … / Five Japanese are …

(b) The /-s/ genitive is not possible:

*The Japanese’s sense of humor [cp.: The Japanese sense of humor.]

### 2.3.2. Type IIIb (e.g. ‘Kirghiz’)

Certain ethnonyms ending in sibilants have the same formal characteristics as the ‘Japanese’ type above: *Cheremis, Chuvash, Laz, Manx, Swiss, Talys, Tungus.* Some dictionaries do allow regular plurals for certain of these words, but they seem awkward to this writer (e.g. the *American Heritage Dictionary* [1971] and *Webster’s New Riverside University Dictionary* [1984] give these alternate plural forms: *Kirghiz, Kirghizes; Tungus, Tunguses; Chuvash, Chuvashes*). Examples of -es plurals can be found here and there, e.g.:

> When amongst the Abchases the shepherds in spring …
> [J. Frazer *Golden Bough* (abridged version); p. 618]

Most contemporary writers I have consulted, however, avoid the marked plural forms of Type IIIb ethnonyms:

*Delegates of the Chechens, Ingush and Balkars …*
[Walter Kolarz *Russia and her colonies* Archon Books, 1967; p. 189]

### 2.4. Type IV (e.g. ‘Dyirbal’)

The last group to be distinguished is composed of ethnonyms which can be declined according to either the Type I or Type III model, i.e. with or without an overt plural suffix. The dictionaries list dozens of ethnic designations allowing this alternation, e.g. *Ainu, Arawak, Baluchi, Chickasaw, Fox, Mandingo, Zulu.*

(a) Plural either formally identical to singular, or formed regularly:

*A Dyirbal is … / The Dyirbal are … / The Dyirbals are …

Five Dyirbal are … / Five Dyirbals are …

(b) The /-s/ genitive is used:

*The Dyirbal’s sense of humor.* (refers to one Dyirbal or the whole group)

The formal properties of Type IV ethnonyms resemble those of nouns denoting certain wild animals, which — especially when one is referring to them as the prey of hunters or fishers — allow the use of the Ø-plural (*eel, elk, partridge, pheasant, rabbit, quail*, etc.): “We ’bagged’ two brindled gnu, four water-bog …” [†7 Speke Discov. Nile 36 (1863)]. Not infrequently we come across both regular and unmarked plurals of these nouns within the same text, as in the following:

*Praying, among the Havasupai, is much like speech-making … A formal speech has never been recorded and analyzed among the Havasupais or Hualapais.*
In almost all cases Type IV ethnyonyms refer to autochthonous peoples of Africa, Asia, Australia and the Americas, a fact which has not gone unnoticed by the grammarians:

The uninflected plural is especially common with the names of uncivilized or less civilized peoples: the Iroquois, Navaho, Ojibwa, Omaha, Blackfoot, Duala, Bantu, Swahili, etc. [George Curme A grammar of the English language, vol II §43.3 (NY: Heath & Co 1935)]

Names of uncivilized peoples are often used unchanged in the plural: the Eskimo, Bateke, Batungo ... [O. Jespersen A modern English grammar on historical principles, vol II, §11.58 (Heidelberg: Carl Winter, 1922)]

In this paper I will discuss the history of this last class of English ethnyonyms, and offer some hypotheses concerning the course of their evolution. The declension of Type IV ethnyonyms reflects a complicated mix of factors, not all of which I can claim to have understood. What I will argue is that in order to account for the sharp increase in the usage of uninflected plurals of the "names of uncivilized or less civilized peoples" in the last century, we must examine both (1) the writer's conception of the Object of ethnographic study, and of his or her role as trained observer of this Object; and (2) how writers position their texts within a valued discursive tradition through the (conscious or unconscious) appropriation of textual surface features indexing that tradition.

3. The declension of the names of non-Western peoples. The remainder of the paper will be given over to an examination of the declension of the ethnyonyms of non-Western peoples, in particular the indigenous nations of North America. The general conclusions I draw appear to apply equally well to the ethnyonyms of African, Asian, Oceanic and South American peoples.

3.1. The early accounts of aboriginal North Americans [16th-19th cc.]. Ethnyonyms of indigenous American peoples begin to appear in English-language texts in the 16th century. Up to the end of the 19th century these nouns, with rare and scattered exceptions, are declined like Type I ethnyonyms:

... the lande amonge the Esquimawes of the Grande Bay
[†Hakluyt Disc. Western Planting xiii.88 (1584)]

... with the Wunnashowatuckoogs and Wusquowhananawkits, who are the further most Neepnet men [Roger Williams Correspondence (Aug 1636)]

The few examples of non-declined plural ethnyonyms are most often accompanied by others which are overtly pluralized. (It may be the case that the writers were using native forms known by them to have plural reference). The examples below are the sole exceptions in each text to a consistent use of -s plural ethnyonyms.

The Indians of the south of this territory, they call Edu, Eduu or Edues; the
general name for themselves is Monqui, or Monquis ...  
[Miguel Venegas A natural and civil history of California (orig. in Spanish; Eng. trans.
published in 1759; reprint: Ann Arbor: University Microfilms, 1966), p 55]

Native American ethnonyms took regular plurals in the Chronicles of the Pilgrim
Fathers of the Colony of Plymouth from 1602 to 1625 [ed. by Alexander Young.
Boston: Little & Brown, 1841]; The correspondence of Roger Williams (vol 1:
1629-1653) [ed. G. W. LaFantasie. Providence: Brown U Press, 1988]; and the
series of journals published in the series Early Western Travels, 1748-1846 [series
ed. Reuben G. Thwaites. Cleveland: A. H. Clarke Co., 1904]. The vast majority of
the ethnonyms recorded in the journals of the Lewis and Clark expedition
(1804-1806) take plurals in -(e)s (e.g. the employments used by the Chinooks
Clatsops Cath-lah-mahs Kil-a-mox &c in hunting (Jan 15, 1806)). The exceptions
are Assinniboin, Osage and Saukee, which alternate between Type I and Type IV
decisions; and Cataka, Staetaen and Kanenavich, which always take the -Ø
plural. The almost exclusive use of -(e)s plurals for native American ethnonyms
continues right up through the writings of John Wesley Powell in the last third of
the 19th century (the only variation noted being in the declension of Ute):

They call themselves Nu-mes, Nu-intz, Nu-mas, Nu-mos ... We will call them
Numas [p 37]; ...they are known as Sho-sho-nees, Bannocks, Ute, Pai-Utes,
Mo-quis, Chem-a-hue-vas ...  
[J. W. Powell Anthropology of the Numa; D.D. Fowler & C.S. Fowler, eds.

3.2. Ethnonyms of non-Western peoples in encyclopedias [19th-20th cc.].
The pre-20th-century works I examined for the declension of ethnonyms represent
a variety of styles: diaries, letters, official reports. In order to keep the variable of
genre constant, a sample of encyclopedia articles from the 19th and early 20th
centuries was selected. What emerged was an increase with the passage of time in
the proportion of -Ø plurals for the designations of African, Asian and American
peoples. Note in particular the pattern of usage in the three editions of the
Encyclopedia Britannica which were sampled: the (merely orthographically?) overt
plural Esquimaux in the 8th edition gives way to the -Ø plural Eskimo in the 10th;
the Zulus-Xosas of the 10th edition are replaced by the Zulu-Xosa in the 11th.

(1833) The Penny Cyclopaedia of the Society for the Diffusion of Useful Knowledge
[London: Charles Knight]; entries AFRICA and AMERICA:
-(e)s plurals are used throughout, with only one exception.
Examples: the country of the Wanketzens or Wanketchies; the King of the Foulahs; the
Tuuricks [pluralized even though the root itself is plural in Arabic]; the territories of the
Ashantees; &c. But: (p 182) the Mackoua ... are described...

entries AFRICA, AMERICA and INDIAN TERRITORY:
-(e)s plurals used throughout, with occasional exceptions.
Examples: Esquimaux; the Tschutschoi <maybe a Russianism?>, the Samoiedes and the Laplanders; the Osages, Missouri, Kansas; the Natches; the Tuarec; the Kabyles, or Kabaily, of Algeria; the Feloops ... and the Ashanti

(1877) The Condensed American Cyclopedia [NY: D. Appleton]; entry 'AMERICAN INDIANS' -es pluralis used throughout, without exceptions.
The Mandans and Chinooks; the Pequots and Narragansetts; the Foxes and Miamis

(1902) Encyclopaedia Britannica 10th edition; entries 'AFRICA' and 'AMERICA':
Both -es and -s pluralis used throughout.
The eastern Eskimo are dolichocephalic ... and the Aleuts brachycephalic; Christianity has made some progress among the Waganda ... and the Zulus-Xosas, Basutos and Bechuanas; Songhay, Hausas, Baghirmi, Kanuri, Mabas

(1910-11) Encyclopaedia Britannica 11th edition; entry 'AFRICA' (The 'AMERICA' article is mostly unchanged from the 10th edition):
Almost exclusively -s pluralis.
Beehive huts are found among the Zulu-Xosa and Herero; the Tuareg, Tibbu, Bedouins and Bushmen

3.3. Ethnonyms of non-Western peoples in the writings of professional ethnographers [late 19th-20th cc.]. It is the opinion of many that the field of anthropology as we now know it in North America began to form around the turn of the century, and the individual frequently credited with being the 'shaper of American professional anthropology' [D. Hymes Language in culture & society, p 7] was Franz Boas (1858-1942). The German-born and educated Boas played a crucial role in furthering the collection and scholarly study of indigenous American languages, and trained many of the major figures in American anthropology. Boas wrote in both German and English, and in his published works in both languages he overwhelmingly preferred -s pluralis for Type IV ethnonyms (and even for the name of at least one European ethnic group); e.g.:

The Eskimo inhabit ... [The Central Eskimo (Smithsonian Inst, 1884-5)]

... the Magyar of Europe, who have retained their old language ...
[Introduction Handbook of American Indian languages (Washington: Government Printing Office, 1911); Vol 1, p 9.]

Der Einfluss der sozialen Gliederung der Kwakiutl auf deren Kultur

In cases of tribal names derived from English words, Boas's tendency was to avoid the nominal use altogether and use the ethnonym as an adjective:

...from the Hare Indians and from the Chippewayan
...among the Chilcotin, who live northwest of the Thompson Indians
[Race, language and culture (NY: Free Press, 1940); p. 417]
The usage preferred by Boas continued — and continues — to predominate in scholarly writing, as well as in works intended for a wider readership. Among the other ethnographers who contributed to the first and second volumes of the Handbook of American Indian languages edited by Boas, the majority (P. E. Goddard, J. R. Swanton, R. B. Dixon, L. Frachtenberg and W. Bogoras) used Ø-plurals without exception for Type IV ethnonyms. The twenty-volume series The North American Indian by Edward S. Curtis [1907-1930; reprint: NY: Johnson Reprint, 1970] employs Ø-plurals exclusively. Of the 44 titles in the series Indians of North America [series editor F. W. Porter III; NY: Chelsea House] in which plural tribal names appear, only four have -(e)s plurals.10

As the author Bernard de Voto complained in the preface to his book Across the wide Missouri (Boston: Houghton Mifflin, 1947), this linguistic practice has become a stereotypical feature of the writing style of professional ethnologists:

If anyone reads this book who has read earlier works of mine in which Indians figure, he will observe that herein I have emancipated myself from a pedantry of my betters. I was brought up to respect learned men and few scholars impress me so profoundly as the ethnologists. So heretofore, mutely enduring the discomfort natural to a literary man, I have used the spellings of Indian plurals that their guild oaths impose. It seems that when you bring Indian tribal names into English from the mother tongue, the plural has the same form as the singular. The ethnologist’s medicine has commanded him in a vision to stand on those plurals, even when they are clearly English words, and neither logic nor a decent sensitiveness of style will move him to violate the sacred teaching. Up to now I have followed his precept, forcing myself to write not only “twenty-one Arapaho” but also “thirty-eight Crow” and even, God help us, “one hundred and two Blackfoot” ...

4. Further observations concerning the declension of Type IV ethnonyms. The widespread use of Ø-plurals with Type IV ethnonyms finds no parallel in the declension of common nouns borrowed in recent times from non-Western languages. Words such as fez, dhow, quetzal, coyote, wallaby, cacique and sachem are regularly declined with plurals in -(e)s. Except in the case of ‘hunting/fishing’ Ø-plurals (moose, caribou, muskellunge), nouns borrowed from aboriginal languages take regular plurals. In this section I will consider two factors, associated with the appropriateness of a scholarly writing style by professional ethnologists, to account for the special treatment of the ethnonyms of “uncivilized peoples.”

4.1. Ø-plurals and collectivity. In his study of grammatical number in English, W. Hirtle [1982:20] sees the use of Ø-plurals with nouns designating both ethnic groups and animal species as a linguistic means of representing individuals as “partaking of the collectivity.” He writes, “according to some grammarians, an example like these Micmacs evokes a number of individuals, whereas these Micmac emphasizes more strongly the notion of tribe member, of the collectivity. In the case of animals something similar can be observed: where most people would speak of, say, two bears to characterize two animals, hunters, naturalists, conservationists, etc. — “cognoscenti”, as one writer terms such speakers — would be more likely to use a zero plural ... This use of the zero plural is characteristic of precisely those speakers who are most cognizant of the species as an entity with its own characteristic traits of behavior” [loc. cit.]. Further on in the same monograph Hirtle again speaks “of words naming members of tribes and
other ethnic groups: 2000 Eskimo or these western Carrier, but an Eskimo or a western Carrier ... One interesting question concerning these nouns is what distinguishes them from nouns of nationality (e.g. Canadian, Dane, Brazilian), which are not found in the internal plural [= Ø-plural — KT]. The distinction would seem to involve the impression expressed by Eskimo, etc. of an inherent, innate characteristic linking one individual with others in an ethnic group. Nationality words, on the other hand, suggest a more accidental relation of a geographical or a political character. However the problem requires further investigation before even a tentative explanation can be offered” [1982:68].

The key points to take note of in Hirtle’s observations are, first, the linking of the declensions of Type IV ethnonyms and game animals as reflecting the same underlying semantics (emphasis on collectivity or species), and, second, the claim that the use of Ø-plurals is characteristic of “cognoscenti” in particular.

4.2. Ø-plurals and the cognoscenti. In a discussion of the declension of animal names in English, Allan [1976: 102] observes that the “hunting/fishing” Ø-plurals are also frequent “in reports of animals observed in nature reserves particularly by game rangers and cognoscenti.” He illustrates the contrast between the usage of experts and non-experts with the following sentences (his acceptability judgments):

We bagged three elephant that day.
We observed three elephant in the game park.
?We saw three elephant in the game park.
?*We saw three elephant at the zoo.

Commenting on Allan’s data, Hirtle [1982: 57] adds that “the third sentence would sound pretentious if spoken by the average tourist, simply because he would not have the knowledge of the species permitting him to see the individuals as animated by it. By the same token, the second sentence is quite acceptable because the use of observed implies a more cognizant speaker.”

The conclusion one would draw from Allan’s and Hirtle’s work is that the Ø-plurals of Type IV ethnonyms would be favored by those whose relation to non-Western ethnic groups corresponds to that of “game rangers and cognoscenti” to animals. This prediction is borne out by the data presented here — and more explicitly in the remarks by de Voto cited above. I do not think it a coincidence that a sharp increase in the use of Ø-plurals occurred at the same time as the professionalization of American anthropology. It appears, nonetheless, that there is more than the perception of an “innate characteristic linking one individual with others in an ethnic group” behind the linguistic behavior of social scientists; one should consider the possible influence of certain non-English plurals used by writers of an earlier era.

4.3. Ø-plurals and Latinate plurals. Until relatively recently, familiarity with the classical languages was widespread, almost universal, among those Europeans who had access to the medium of writing. Latin and Greek sources provided not only the informational background for many works written in the vernacular languages, but also many surface features of these texts. The presence of the classical languages could range from scattered quotations or borrowings within a
clearly vernacular matrix to something approaching the opposite extreme (e.g. the Lindisfarne gospels, in Latin with Old English interlinear glosses).

In most cases proper names of Latin and Greek origin were assimilated to English morphological patterns in texts of the 16th to 18th centuries (e.g. *the writings of Plato* [and not ‘of Platonis’]; *Render, therefore, unto Caesar the things which are Caesar’s* [Mt. 22:21 (AV); rather than ‘Cæsari ... Cæsarios/Cæsaris’]. The exceptions are varied. In some instances an author has simply lifted a name or phrase from a source and left it in its original form, perhaps because of the assumption that a reader of that period would be more likely to know it from classical rather than vernacular texts. Here are two examples:

*the Pope’s Belvedere in Rome, as pleasing as those horti pensiles in Babylon ...*  
[Robert Burton *Anatomy of melancholy* Pt 2, Sec 2, Mem 4 (1621)]

*The two Apollinarii [= Apollinarius and his son — KT] were fain, as a man may say, to coin all the seven liberal sciences out of the Bible ...*  
[John Milton *Areopagitica* (1644)]

The majority of ethnonyms taken from classical-language sources are likewise assimilated into English, sometimes through derivational suffixation (*Etrusci ⇒ Etruscans, Sarmatae ⇒ Sarmatians, &c.*), sometimes without (*Gothi ⇒ Goths*). In many instances, however, the ethonym is taken unchanged from Latin, even though the ethnonyms in its immediate context may have English -(e)s plurals:

*those cantons of Switzers, Rheti [cp Latin R(h)æti — KT], Grisons, Walloons ...*  
[R. Burton *Anatomy of melancholy* ‘Democritus to the reader’ (1621)]

*The Goths, the Vandals, the Gepidae, the Burgundians, the Alemanni, wasted each other’s strength by destructive hostilities ... They granted a settlement to several colonies of the Carpi, the Bastarnae, and the Sarmatians*  
[E. Gibbon *Decline and fall of the Roman Empire* §13 (1776)]

In some cases the nonassimilated declension may have been resorted to because the author had no assimilated ethonym ready to hand. At the same time, could it have been the case that at least some writers employed Latinate declensions to add an aura of prestige to a text primarily written in the vulgar tongue?

Consider the various German renderings of the New Testament. In Luther’s translation of 1522, most proper names are declined as in Latin; e.g. accusative: ... *vnd legten die hende an Jhesum* [Mt 26:50; cp Vulgate *et manus injecerunt in Jesum*]; dative: ... *vnd vberantworten in dem Landpfleger Pontio Pilato* [Mt 27:2; Vulgate ... *et tradiderunt Pontio Pilato præsidi*]; vocative: *Weissage vns, Christe* ... [Mt 26:68; Vulgate *Prophetiza nobis, Christe*].¹¹ In some later translations, including the revision of Luther’s text in modern orthography, the Latin declensions are preserved only for the names of the central figure (*Jesus, Christus*). The names of mortals, by contrast, are either undecorated, or morphologically assimilated into German (e.g. masculine genitive in -*s*). Some examples: revised Luther ... *und sprach zu Petrus* [Mt 16:23; cp Luther’s original ... *vnd sprach zu Petro*]; revised Luther ... *und überantworteten ihn dem Landpfleger Pontius Pilatus* [Mt 27:2; cp
Luther's original above. 12 One is tempted to see here a shift in the way the prestige believed to be conferred by a classical language was distributed. If Luther used Latin, and occasionally Greek, declensions to highlight the sacredness of the text as a whole, the later editors distributed them more locally, to emphasize the sacredness of a particular figure as against all of the other characters mentioned.

Let us return to English literature. Some of the works cited above, such as Gibbon's *Decline and fall*, were considered essential reading for members of the privileged classes until this century, and still enjoy high repute (as indicated, for example, by their enshrinement in the Britannica's *Great Books of the Western World* series). The use of more-or-less unassimilated bits of Latin and Greek can be seen, naturally enough, as a sign of familiarity with a body of knowledge considered prestigious by many in our culture. This practice of appropriating surface features from texts composed in one language, and employing them strategically within the matrix of writings in another language, is not limited to the importation of Latin and Greek lexemes. Words from non-Western languages have also been appropriated in this fashion, for the same purpose. This influence could operate at at least two levels: the morphological and the phonological.

[a] *Morphological*: Just as many 16th-19th century English writers made use of nouns inflected as in Latin or Greek in their vernacular-language works, some later writers have employed the declensional patterns of a non-Western language for proper nouns set in an English text. T. E. Lawrence employs Arabic singulars and plurals for most of the tribal and clan names mentioned in his *Seven pillars of wisdom* (1926) — e.g. sg. *an Ageyli*, pl. *the Ageyl*; sg. *a Howeiti*, pl. *the Howeitat* — as does the German writer Werner Munzinger in a book on East African tribes (*Ostafrikanische Studien* [Schaffhausen, 1864; reprint NY: Johnson Reprints, 1967]); e.g. sg. *der Bedui*, pl. *die Beduan* "Bedouins". In using Latinate plurals, of course, Gibbon, Milton, Burton et al. were tacitly indexing the body of classical learning shared by them and their intended audience; Lawrence may have been doing something of the same (with a much smaller circle of intended readers), though I cannot be sure of this. In any event, one would imagine that this type of usage would be resorted to less often with ethnonyms derived directly from non-Western languages (rather than through classical intermediaries, as in Gibbon's case), except in special contexts (professional journal or conference) where the audience is likely to know something of the language in question.

[b] *Phonological*: Most often it is the case that the declension in English texts of an ethnonym derived from a native word takes no account of its morphological alternations in the language of origin. Thus we have ethnonyms derived from what were originally plurals (e.g. *Bantu* < Swahili *ba*- 'plural prefix' + *ntu* 'man'; cp Swahili singular *mu*-*ntu* 'a man'; likewise *Inuit* < absolutive plural of Eskimo *inuk* 'a man'), which can function as singulars in English (*a Bantu, the Bantu(s); an Inuit, the Inuit(s)*).

On the other hand, there is evidence that certain authors who may in fact have little knowledge of the source languages inflect ethnonyms in a manner which implies a degree of imitation of Gibbon et al. at the phonological level. In a number of texts from the past century, the authors, in presenting lists of ethnonyms of
peoples from the same geographical and cultural regions, specifically avoid adding the -(c)s plural to ethnonyms ending in a final vowel, especially -i, e.g.:

Songhay, Hausas, Baghirmi, Kanuri, Mabas ...
Bejas, Somali, Gallas, Turkana, Masai ...
[Encyclopaedia Britannica 10th ed., 1902: article ‘AFRICA’ (vol XXV, pp 139ff)]

The Gilyaks, the Aino and the Goldi are all of the opinion that ...
[J. Frazer Golden Bough; p. 597 of abridged version (1922)]

Avars, Andi, Akhvakhs, Bagulals, Botlikhs, Chamalals, Godoberi, Karata, Tindi, Dido, Bezheita, Khunzals, Khvarshi, Archi...

One plausible explanation for this behavior is the subliminal influence of vowel-final plurals in much of the learned vocabulary of Greek, Latin or Romance origin — foci, loci, literati, temp; varia, strata, flora — but perhaps more particularly the Lateinian plurals in -i or -e of ethnonyms13 cited in the widely-read works (widely-read in certain circles, at least) of 16th-19th century historians, philologists and humanists.

If this conjecture is substantially correct, it would follow that these ethnonyms could be considered a special subgroup of Type IIIb: just as the final sibilants of Chuvash, Cheremis, Abkhaz &c inhibit the addition of a sibilant English plural suffix, so the final vowels of exotic-sounding ethnonyms recall the Lateinian plural forms of ethnonyms in works produced within a prestigious scholarly tradition, and likewise discourage the suffixation of -s. To the Sprachgefühl of an initiated reader (but not too initiated!), these words sound plural.

5. Summary and conclusion. The professionalisation of American anthropology (associated with Franz Boas in particular) at the turn of the century is accompanied by a marked increase in the use of Ø-plurals of the “names of uncivilized peoples” in ethnographic literature. Why would one seek a link between declensional preference and the professionalization of ethnology?

The linguists Allan and Hirtle have noted an association between the Ø-pluralization of names of animals and ethnic groups, and a certain expertise or connoisseurship claimed by the speaker: We bagged two brindled gnu and four water-boc; ... five years of fieldwork among the Akhvakhs ... &c.). The non-availability of a distinct plural form is also linked with collectivity (mass nouns and the like), and, so the explanation goes, it is the “cognoscentsi” — ethnologists, zoologists, hunters — who are most prone to use this linguistic device to represent individuals as “partaking of the collectivity” [Hirtle, p. 20].

Without denying that Allan and Hirtle are in fact on to something, I propose here another factor suggested by a historical examination of ethnographic writing: the (largely unconscious) indexing of the learned literature of an earlier era through the avoidance of -(c)s plurals for “exotic” ethnonyms. In previous centuries it was common practice for writers within the humanist tradition to import many surface
features of Greek and Latin texts into their vernacular-language writings. In the case of ethnonyms in particular, the names of peoples unfamiliar to the writer save through classical-language sources were frequently taken unchanged from the Latin, even though the ethnonyms in its immediate context might have English -(e)s plurals. My hypothesis is that this usage has become part of the stereotype of the scholarly writings of previous centuries, and thus endowed with a certain aura of erudition and prestige, especially in the eyes of those who adopt a distinctly ‘academic’ style of expression. When de Voto characterized the Ø-plural as a usage imposed by the “guild oaths” of professional ethnology, he was hardly exaggerating. Finally, it is a curious fact that many scholarly writers selectively avoid adding the -(e)s plural to “exotic” ethnonyms ending in a final vowel, a reflection, I suspect, of the subliminal influence of vowel-final plurals in much of the learned vocabulary of Greek, Latin or Romance origin.

References and notes.

The author wishes to acknowledge the helpful comments on an earlier version of this paper offered by Modj-ta-ba Sadria (Tokyo) and Michael Silverstein (Chicago), who are, of course, absolved from responsibility for its contents. Silverstein raised the important question of syntactic context: is there any statistical correlation between Ø-plurals and genericizing contexts, such as [Quant N of N]NP : a band of Cherokee(s) ? At present I do not have enough facts at my fingertips to respond to this query, but I agree that it is worth looking into. Thanks also to those who commented on the paper after the oral presentation at BLS: Johanna Nichols, Suzanne Romaine, Dan Slobin, Rebecca Wheeler, among others.

The OED states that Irish, Dutch, etc. are “elliptical uses of the adjective.” Likewise Allen [1980: 558], who has “accounted for the human reference of such NPs as the police, the English, the poor, etc., by, in effect, postulating a head noun with the semantics of ‘person/people’; this does not appear on the surface of these NPs as a lexical item, but remains understood.”

Not counting the use of the non-articulated singular to denote the language (Irish is not easy to learn). The usage of quantifiers before Type II ethnonyms is considered awkward in Modern English, but there is some difference among speakers concerning the degree of unacceptability. Thus Jespersen [1922: vol II, §11.54]: “The use of these adjectives after numerals and similar adjectives [i.e. quantifiers — KT] is not quite natural nowadays.” That being said, he provides several counter-examples from texts: six thousand British, some English that I know . . . See also Allen [1980: 558]
The status of Spanish, British and Scotch/Scottish is not so clear. All have Type I counterparts (Spaniard, Briton, Scot) and lack derivatives in -man/woman, but are nonetheless frequently employed as plural-collective nouns. For this latter reason they will be classified as Type II ethnonyms.

The sequence of linguistic changes leading to the Modern English treatment of Type II and III ethnonyms cannot be reconstructed with certainty. I suspect that phonology played an important role in the process: all of these ethnonyms end in sibilants (Irish, Japanese, Kirghiz), which may have blocked the use of the -(e)s plural suffix. I offer the following three-stage scenario for the reader’s consideration:

(i) Middle English ethnonyms in -(i)sc/sh were pluralized with an unstressed -g (Saxons and Englishe), which was lost by the beginning of the Modern English period.

(ii) While some speakers extended the now-dominant -(e)s plural to these nouns as well (Irishes), others continued to treat them as elliptical uses of the adjective, with an understood plural head noun meaning ‘people’ (the Irish [people] are ...) [Allen 1980:558]. Type III ethnonyms took regular plurals when they first appeared in English in the 16th century (the Genoese; Chinesas and Japaneesas [OED]), but shifted to Ø-plurals in the 17th and 18th centuries.

(iii) The preference for Ø-plurals with Type III ethnonyms could have been encouraged by the collective plural use of Type II ethnonyms, which are likewise sibilant-final. Note as well that -(e)s plurals were sporadically omitted in other sibilant final nouns in Early Modern English: Are there balance here to weigh the flesh? [Merchant of Venice IV.i.255; cited by Jespersen 1982:175].

6 Some Type I ethnonyms have parallel Type IV forms: e.g. the Mayans / the Maya(s).

7 The symbol [†] indicates a citation from the Oxford English Dictionary (OED).

8 This title, it must be pointed out, was added by the editors, who use the Type IV declension consistently for American ethnonyms, e.g. several groups of Southern Paiute ... Shoshoni dialects were spoken by the Western Shoshoni ... (pp 6-7).

9 The declension of non-Western ethnonyms in German merits a separate study. The oldest examples I have come across (in a far-from-thorough search) occur in the writings of J. G. Herder from the 1770’s, e.g. zehn Jahr unter den Abenakiern, ... die abstehenden Ohren der Pevas und Amikuanes. The earliest undecorated ethnonyms I have noted are in W. von Humboldt’s Über die Verschiedenheit des menschlichen Sprachbaues [1836]; e.g. Denn die Samang, welche einen Theil der Gebirge derselben bewohnen, sind ...

10 Among the authors contributing to the series, R. S. Grumet (The Lenapes, 1989) uses the -(e)s plural with an unusual consistency; e.g. (p 25) ... such diverse peoples as the Cheyennes and Blackfoots of the Great Plains. J.R. Wunder (The Kiowa, 1989) employs the Ø-plural throughout, except when an ethnonym is quantified by a numeral; e.g. (p 54) a war party of 1500 Kiowas, Comanches, Cheyennes and Arapahos...

11 On Luther’s treatment of loanwords and proper names, see Carl Francke Grundzüge der Sprachsprache Luthers [Hildesheim: G. Olms Verlag, 1973]; Pt II, §109.

12 Correspondingly, the occasional lapsus calami in Luther’s rendering of the oblique cases of the Saviour’s names is changed to fit the new pattern: Christus Geist [Romans 8:9] becomes Christi Geist in the revised Lutheran Bible.

13 This latter explanation seems all the more likely in view of the fact that other vowel-final words borrowed from non-Western languages take regular plurals.
Variation In Modern Dutch D-Weakening: A Historical Perspective
Nelleke Van Deusen-Scholl
University Of California, Berkeley

1. INTRODUCTION. In modern Standard Dutch an optional rule exists by which d may be weakened in intervocalic position, followed by either a glide insertion across morpheme boundaries or a word contraction. Thus, goed ‘good’ can become goeié, brede ‘broad’ becomes breeé, and leder ‘leather’ becomes leer. (see example 1). These forms are subject to considerable (regional and social) dialectal, stylistic (esp. casual speaking style), and individual variation. The phenomenon, however, is not a recent one but can be traced back to perhaps as early as the 11th century (van Loon 1986:141). The earliest evidence shows up in old place names (cf. 12th century Nerisca for earlier Netherisca) and in the earliest Middle Dutch (e.g. woensdag<Wodanesdag ‘Wednesday’) and is well-documented by the 14th century. The change appears to have spread from the South to the North of the Low Countries, and Middle Dutch and early Modern Dutch texts already show a significant amount of dialectal and stylistic variation.

Historically, a number of sociolinguistic factors played an important role in this process, while synchronic variation is to a great extent constrained by historical factors. Thus far, the majority of the research devoted to this issue has consisted of either phonological or historical approaches (Cohen 1959; Franck 1910; Goossens 1974; van Loey 1957; van Loon 1986; Schönfeld 1947). Neither approach, however, captures the sociolinguistic reality of modern Dutch d-weakening. In this paper, I will present an analysis of synchronic variation in Dutch d-weakening which takes into account the historical processes which have affected it. Understanding the principles involved in the historical development of the problem may provide insight into the seemingly random patterns of synchronic variation and may explain the choices open to native speakers with respect to the degree of acceptability of these forms.

2. ZONNEVELD’S (1978) ANALYSIS. Most recently, Zonneveld (1975, 1976, 1978, 1981) and Smith (1973, 1975) have considered the phonological constraints of d-weakening. Zonneveld’s (1978) dissertation presents the most detailed phonological analysis thus far. He proposes a two-rule solution to the problem: (1) a weakening process, which occurs across suffixal boundaries (e.g. goed-goëie ‘good’) and (2) a contraction rule which occurs intramorphemically (e.g. broederoer ‘brother’). He distinguishes three classes within the d-weakening process.

Class I undergoes weakening and gliding across suffixal boundaries. Between a stressed tense vowel or diphthong and an unstressed morpheme of the shape -e [ə], -en [on], -er [ər], -ig [ay], d is weakened to a palatal or labial glide (Zonneveld 1978:27). Some examples are:

(1) a. attributive -e: brede - breeé ‘wide’
  b. infinitive -en: ríjden - ríjen ‘to drive’
  c. noun pl. -en: kruíden - kruíen ‘spices’
  d. comparative -er: breder - breeër ‘wider’
  e. derived adj. -ig: goedig - goëig ‘good-natured’

In Class II, d is deleted intramorphemically, accompanied by contraction of the following schwa:

(2) leder - leer ‘leather’
    ader - aer ‘vein’
Zonneveld (1978:37) describes this rule informally as "-de is deleted after long vowels and diphthongs in word-final position or before the consonants g, k, m, l, or r." He perceives this process as a contraction rather than a weakening process and therefore argues for two different rules for intervocalic d-weakening.

Class III contains a small number of exceptions to the above two classes. These forms may undergo:

(3) a. both the weakening process of class I and the contraction of Class II, resulting in three possible realizations of a word:
   bodem - boiem - boom 'bottom'
   kade - kaai - ka 'quay'
   b. the Class II contraction across morpheme boundaries:
   bladeren - blaren 'leaves'
   c. the weakening of Class I intramorphemically:
   poeder - poeier 'powder'

D-weakening thus appears to take place in the environment following long vowels (the front vowels [i:] and [e:]), and the back vowels [u:], [o:], and [a:] and diphthongs and before an unstressed morpheme.

Zonneveld (1978:22) only briefly refers to the sociolinguistic factors which affect this phenomenon by noting that "[t]here is considerable dialectal and even, so it seems, personal variation as regards the acceptability or nonacceptability of these forms." His analysis, however, as any of the historical approaches before him, oversimplifies the complexities of individual and stylistic variation. For instance, in his 1981 article, Zonneveld (1978:28) simply notes that "d-weakening is triggered by casual style of speech," ignoring the considerable variation among native speakers with respect to the acceptability of these forms. Before addressing this issue, I will present an overview of the historical background of d-weakening.

3.1 HISTORICAL BACKGROUND. Historically, d-weakening appears to have spread in the Low Countries from the South, now approximately the southwest of Belgium, to the North. Schönfeld (1947:34) suggests that a possible explanation for the phenomenon might be found in the influence of French of Westflanders. In French, d became syncopated (4a) but remained intact after l (4b):

(4) a. Lat. audire > Fr. oîre 'to hear'
   Similarly, West-Flemish maintained a distinction between, on the one hand, forms such as broer (<broeder 'brother'), rïjen (<rijden 'ride, drive'), zaal (<zadel 'saddle') and, on the other hand, those forms which subsequently underwent l-vocalization, such as houden (cf. Ger. halten 'hold'), oude (cf. Ger. alt 'old'), and koude (cf. Ger. kalt 'cold'). Schönfeld hypothesizes that when d-syncope spread from West-Flanders northwards, l- vocalization had already taken place, so that the difference between forms containing the l and the other forms does not exist elsewhere.

Although Schönfeld (1947:67) suggests that the d-deletion process may have started as early as the 11th century, an opinion based on the hypercorrection Odeka for the 10th century place name Hoica, clear examples can be found in the 14th century in the provinces (North- and South-) Holland and Utrecht. Some early examples are:

(5) Mi.D. roesien hoymaker gerechtigheien
   Mod.D. roeden hoedemaker gerechtigheden 'rods'
   'hatmaker'
   'justices'
Franck (1910: 102) reports the following examples from 14th century manuscripts:

(6) 
\[
\begin{array}{ccc}
ghecroeten & < & ghecroden \\
sint & < & sident \\
vlerc & < & vlederic & \text{‘wing’}
\end{array}
\]

According to Schönfeld (1947:67), the oldest and most prevalent forms of d-syncope are those where d is followed by vowel + consonant (preferably l or r) + syllable; for example:

(7) 
\[
\begin{array}{ccc}
\text{E. Mod.D.} & \text{vlerc} & < & \text{vlederik} \\
\text{Bokele} & \text{Bodokenlo} & \text{‘wing’} \\
\text{seware} & \text{sedeware} & \text{‘seaweed’} \\
\text{goeliyc} & \text{goedelic} & \text{‘good-natured’}
\end{array}
\]

Schönfeld (1947: 67) suggests that d-syncope must have spread northward and subsequently in eastward direction via the Brabant dialect. D-syncope appears earlier and with much greater frequency in that particular dialect. He hypothesizes that the change moved north to the province Utrecht, which influenced the northern part of South-Holland, which in turn influenced the North-Holland dialects, specifically that of Amsterdam. To the east the change appears to have reached de Betuwe, which is a region in the river valley just east of the province Utrecht. This northward direction of d-syncope is supported by historical facts, particularly the fall of Antwerp which occurred in 1585 and caused many southerners to flee to the north, especially Amsterdam. The geographic distribution of d-deletion can be seen in placenames, such as the following cited by Schönfeld (1947: 35):

(8) 
\[
\begin{array}{ccc}
\text{Limburg} & \text{Venray} & \text{from} & \text{Venrade} \\
\text{N. Brabant} & \text{Gijzenrooi} & \text{Leverooi} & \text{Gijzenrode} \\
\text{Betuwe} & \text{Wadenooien} & \text{‘Wayenoyen} & \text{Ravenwaai} \\
\end{array}
\]

Similarly, d-weakening appears to be more widespread in the literary works of Southern writers than of Northerners.

3.2 SCHÖNFELD’S ANALYSIS. Schönfeld (1947) provides the following historical analysis of those forms that have undergone d-syncope by early modern Dutch.

(9) 
\[
\begin{array}{ccc}
a. & \text{Rule I. Syncope:} \\
\text{\begin{itemize}
\item with loss of syllable: word-final d- assimilated to the stem syllable:}
\item \text{reu} & < & \text{Mi.Du. reude/rode} & \text{‘male dog’}
\item \text{sloe} & < & \text{slede} & \text{‘gutter’}
\item \text{Leerdam} & < & \text{Mi.Du. Lederdamme} & \text{place name}
\item \text{beul} & < & \text{bodel} & \text{‘executioner’}
\item \text{roer} & < & \text{roeder} & \text{‘rudder’}
\item \text{kuieren} & < & \text{Mi.Du. koderen} & \text{‘to stroll’}
\item \text{opruien} & < & \text{opruiden} & \text{‘to incite’}
\item \text{vouwen} & < & \text{vouden} & \text{‘to fold’}
\end{itemize}}
\end{array}
\]
b. Rule II. Glide-formation from [d] to [y]:

- ooievaar <Mi.Du. odevaarde ‘stork’
- rooien <Mi.Du. roden ‘to dig up’

Schönfeld distinguishes Rule II from the glide formation rule under I, in that the latter is a case of homorganic glide formation, while in Rule II a front glide follows the back vowels [o:] and [a:] and is the result of d-weakening.

Many of the early syncopated forms have become stabilized in modern Dutch, while others are synchronically only acceptable with d or remain open to variation. A variety of phonological, dialectal, semantic, and sociolinguistic factors are responsible for this. A first consideration is dialect mixture. Several of the Middle Dutch dialects did not allow d-weakening, and the varying political and cultural influence of such dialects resulted in variations in the standard language. Particularly also city versus country distribution resulted, with d-weakening seen as a low-status rural variant. Second, the influence of the written language and the competition between socially more and less prestigious varieties, and formal and colloquial registers were responsible for later spelling pronunciations which in turn resulted in restoration or retention of the d in more formal or less common forms. Examples of this in modern Dutch are aanbieden ‘to offer’, gebieden ‘to order’, verluiden ‘to be rumored’, strijden ‘to combat’, etc. Third, there exists a large group of early hypercorrections which occurred in the period from Middle Dutch to early modern Dutch. Te Winkel (1901:94) mentions two environments which were particularly conducive to hypercorrections:

- d-insertion after Middle Dutch ĭ:
  - geschieden <Mi.Du ghesiëne ‘to happen’
  - vlieden <Mi.Du vliëne ‘to flee’
  - kastijden <Mi.Du castiën ‘to punish’
- epenthetic d between l, n, or r and (ə)r:
  - kelder ‘cellar’
  - selderij ‘celery’
  - Hendrik ‘Henry’

3.3 EARLY MODERN DUTCH TO PRESENT. From early Modern Dutch to the present, the forms which had undergone earlier d-weakening either stabilized in d-less form or were subsequently restored, with the result that d-weakening, in many cases, became possible again. Thus, three patterns can be seen.

3.3.1 STABILIZED D-LESS FORMS. A great number of words became stabilized in their d-less forms. All the examples listed under Schönfeld’s analysis in 3.2 above, and many others, have remained that way into the modern period. These forms appear to be the result of a completed sound change. Te Winkel (1901:93) gives the following examples:

- a. contraction
  - bleu <*bleude (OS bloði) ‘shy’
  - gedwee <MHG getwedic ‘submissive’
  - veer <veder (OLF fethera) ‘feather’
  - vleermuis <OHG feldarmus ‘bat’
- b. glide formation
  - rooien <Mi.Du. roden ‘to dig up’
  - uitrooien <Mi.Du. roeden ‘eradicate’
  - ooievaar <Mi.Du. odevaarde ‘stork’
3.3.2 RESTORED FORMS - NO SYNCHRONIC WEAKENING. A small number of forms which had undergone d-weakening by early modern Dutch were subsequently restored and do not allow synchronic weakening. Two examples cited by Schönfeld are:

(12) Mi.Du. vadem > Early Mod.Du. vaam > Mod. Dutch vader ‘fathom’
        vaar vader ‘father’

3.3.3 RESTORED FORMS - SYNCHRONICALLY VARIABLE. In quite a large number of words, the d was restored subsequent to its weakening, which then created a renewed environment for d-weakening. This type of hypercorrection differs from the one described above in that it puts a deleted d back where it previously was rather than inserting it in an analogous environment. Thus a cyclical pattern of d-weakening, restoration, and synchronic variability occurs. Among the forms cited by Schönfeld, the following fall into this pattern:

(13) Mi.Du. lieden > early Mod.Du. lui >Mod.Du. lieden/lui ‘people’
        slede slede/slee ‘sled, sleigh’

3.4 SEMANTIC CHANGES. The availability of dual forms also facilitated a number of semantic changes. As noted earlier, style and register played historically as significant a role as they do synchronically. A number of forms exist where these stylistic options have become lexicalized, splitting along stylistic lines. For example, the word moeder ‘mother’, which cannot be weakened in any way synchronically, shows up in weakened form in a number of colloquial expressions. Following are some examples of this type of split, where the d-less forms denote the lower register:

(14) a. compounds
    Standard Dutch
    but colloquial/dialectal:
          moeder ‘mother’
          bestemoer ‘granny’
          moervos ‘female fox’
          parelmoer ‘mother-of-pearl’
          vader ‘father’
          bestevaar ‘granddad’
          praatvaar ‘chatterbox’
    b. adjectives:
    Standard Dutch
    but colloquial:
          dode ‘dead’
          op z’n dooie gemak ‘at one’s leisure’
          op z’n dooie akkertje ‘at one’s leisure’
          op z’n dooie eentje ‘all by himself’
    c. nouns:
    Standard Dutch
    but denigrating:
          weduw ‘widow’
          weewutje ‘merry little widow’

Stylistic differences in the other direction, i.e. higher register, appear to account for the preservation of d in the following forms:

(15) broer ‘brother’ but: broederschap ‘fraternity’
     broedermoord ‘fratricide’

In a number of cases, it is impossible to detect any such pattern and the d-containing and d-less forms appear to have split haphazardly to accommodate various semantic uses.

(16) neer ‘down’
     nederig ‘humble’
     nederlaag ‘defeat’
     Nederland ‘the Netherlands’
however: Neerlandistiek ‘Dutch studies’
neerbuigend ‘condescending’

In some instances, stylistic differences led to semantic differentiation between the forms containing d and the syncopated forms:

(17) boedel ‘property’ boel ‘a lot’
broeder ‘monk/male nurse’ broer ‘brother’
vergaderen ‘hold a meeting’ vergaren ‘collect, hoard’
tijdel ‘vain’ ijl ‘thin (air)’
teder ‘tender’ teer ‘delicate’

4. SYNCHRONIC SOCIOLINGUISTIC FACTORS AFFECTING D-WEAKENING. Though stylistic choice appears to be the primary motivating factor for d-weakening, it does not present itself as a simple dichotomy between casual and formal styles. Additional (socio)linguistic and functional variables play a role as well.

4.1 STYLE. As Zonneveld (1981:28; 1978: 100) noted, d-weakening occurs most frequently in casual style. However, considerable individual variation is present among speakers of Western Standard Dutch with respect to the acceptability or non-acceptability of the weakened forms. For a limited number of forms the weakened form has become the preferred or even required variant, while other forms remain heavily stigmatized. A small number of high-frequency individual words have begun to be accepted in all styles of speech, even the most formal, with their d-containing counterparts being regarded as overly formal and stilted in spoken Dutch. Examples are goeie ‘good’ (especially its compounds goeiemorgen ‘good morning’, goeiemiddag ‘good afternoon’, etc.) ouwe ‘old’, and kouwe ‘cold’. These are becoming increasingly acceptable in informal written styles as well, suggesting perhaps a renewal of the cycle noted above.

On the other hand, many forms remain highly stigmatized, even in the most casual styles; forms such as zouwen instead of zouden ‘should, houwen instead of houden ‘hold’ tend to be labelled as nonstandard by most people, while many forms are ruled as unacceptable in most situations (e.g. *brooien for broden ‘breads, *ouwerdom for ouderdom ‘old age’). Yet another small set is labeled as belonging to children’s style or ‘childish’ (especially bloeden-bloeiien ‘to bleed’ and verkleden-verkleëen ‘to dress up’). Thus, there appear to be a continuum of options stylistically, ranging from acceptable/preferred to acceptable in casual style to stigmatized and/or unacceptable.

4.2 SOCIOECONOMIC CLASS. While d-weakening in general tends to be socially stigmatized to some extent and stratified along socioeconomic lines (with the exceptions noted above), it also appears to have a significant amount of covert prestige among (upper) middle class, college-educated speakers of Standard Dutch. Particularly among males from western college towns (e.g. Leiden Utrecht, Delft), the extent of acceptability seems to be much greater than among other groups. Conversely, lower middle class hypercorrection is common as well. The examples in (17) show d-insertion in environments which appear to have undergone d-deletion.

(18) *breiden for breien ‘to knit’
*blaude for blauwe ‘blue’
*kauwen for kauwen ‘to chew’
*beeldhouden for beeldhouwen ‘to sculpt’
4.3 REGISTER. An important variable in d-weakening is register, particularly since Dutch has a wide gap between its spoken and written language (sprektaal versus schrijftaal). A large percentage of Zonneveld’s class II words has become differentiated in spoken and written form, where the d-containing variants (for example, leder, slede, lade) occur only in formal written contexts (e.g. advertisements) and are generally unacceptable or highly marked in spoken form (spelling pronunciations), while their d-less counterparts are used in all styles of speaking and informal written contexts.

Less frequently used words which are felt to belong to higher registers also tend to resist d-weakening. Thus, synchronically, the following forms are generally considered unacceptable:

(19) *versmaaien - versmadden ‘to slight’
*geschieën - geschieden ‘to happen’
*vermijen - vermijden ‘to avoid’
*beduijen - beduiden ‘to indicate’

4.4 DIALECT. Dialectal variation, which played historically a major role in the distribution of d-weakening, remains an important factor. The West-Flemish dialects of Blegium and the Brabant dialect of the Netherlands seem to be the most tolerant toward d-less variants.

4.5 FUNCTIONAL FACTORS. Avoidance of homonyms appears to be of only minor significance in blocking the application of the weakening rule. While a number of examples can be found (e.g. waden ‘to wade’ cannot be weakened to waaïen ‘to blow’; goden ‘gods’ cannot become gooien ‘to throw’), an equal number of counterexamples is readily available (rijden ‘to drive’ becomes rijen, also ‘rows’; luider ‘louder’ can become luier, also ‘diaper’).

4.6 WORD CLASSES. Among the Class I forms, nouns (plural ending -en) appear to allow d-weakening less freely than adjectives or verbs. Of the examples presented by Zonneveld (1981:28), none of the nouns were acceptable among the ten native speakers I consulted (e.g. broden -*brooien ‘breads’, hoeden-*hoeien ‘hats’, kruiden-*kruien ‘herbs’, treden-treeën ‘steps’).

5.0 CONCLUSION. Even though modern Dutch d-weakening is generally perceived to be an innovation resulting in stylistic variation, the process itself originated quite early on in the history of the Dutch language. After seven or eight centuries of repeating patterns of change (dialect mixture, stylistic differentiation, hypercorrection, spelling pronunciation, etc.), a stable sociolinguistic variable (Labov, 1994; 1972) has developed which synchronically operates under two phonological conditions (weakening and contraction). However, while phonological analyses, such as Zonneveld’s (1978), may indicate the phonological constraints of d-weakening, they cannot account for the seemingly random and highly idiosyncratic patterns employed by speakers of modern Standard Dutch. A sociolinguistic approach may provide some insights, yet without a historical perspective a great many puzzling exceptions remain. Historical patterns of d-weakening intersect with synchronic ones and produce a complex array of partially completed changes which may block further variation or trigger subsequent hypercorrections. The historical development of the d-weakening process, then, should not be seen as an example of a change in progress, but rather as the creation of a highly complex, widespread, and purely sociolinguistic marker.
REFERENCES

**BERKELEY LINGUISTICS SOCIETY**
2337 Dwinelle Hall, University of California, Berkeley, CA 94720 USA
Telephone: (510) 642-5808, (510) 642-2757; e-mail: bls@garnet.berkeley.edu

### Book Prices*

<table>
<thead>
<tr>
<th>Parasession &amp; Special Session (S) Topics</th>
<th>Individuals</th>
<th>Institutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLS-1 (1975)</td>
<td>$10.00</td>
<td>$15.00</td>
</tr>
<tr>
<td>BLS-2 (1976)</td>
<td>$10.00</td>
<td>$15.00</td>
</tr>
<tr>
<td>BLS-3 (1977)</td>
<td>$10.00</td>
<td>$15.00</td>
</tr>
<tr>
<td>BLS-4 (1978)</td>
<td>$10.00</td>
<td>$15.00</td>
</tr>
<tr>
<td>BLS-5 (1979)</td>
<td>$10.00</td>
<td>$15.00</td>
</tr>
<tr>
<td>BLS-6 (1980)</td>
<td>$10.00</td>
<td>$15.00</td>
</tr>
<tr>
<td>BLS-7 (1981)</td>
<td>$10.00</td>
<td>$15.00</td>
</tr>
<tr>
<td>BLS-8 (1982)</td>
<td>$14.00</td>
<td>$20.00</td>
</tr>
<tr>
<td>BLS-9 (1983)</td>
<td>$10.00</td>
<td>$15.00</td>
</tr>
<tr>
<td>BLS-10 (1984)</td>
<td>$14.00</td>
<td>$20.00</td>
</tr>
<tr>
<td>BLS-11 (1985)</td>
<td>$14.00</td>
<td>$20.00</td>
</tr>
<tr>
<td>BLS-12 (1986)</td>
<td>$14.00</td>
<td>$20.00</td>
</tr>
<tr>
<td>BLS-13 (1987)</td>
<td>$14.00</td>
<td>$20.00</td>
</tr>
<tr>
<td>BLS-14 (1988)</td>
<td>$14.00</td>
<td>$20.00</td>
</tr>
<tr>
<td>BLS-15 (1989)</td>
<td>$14.00</td>
<td>$20.00</td>
</tr>
<tr>
<td>BLS-16 (1990)</td>
<td>$16.00</td>
<td>$20.00</td>
</tr>
<tr>
<td>BLS-16S (1990)</td>
<td>$8.00</td>
<td>$12.00</td>
</tr>
<tr>
<td>BLS-17 (1991)</td>
<td>$16.00</td>
<td>$20.00</td>
</tr>
<tr>
<td>BLS-17S (1991)</td>
<td>$8.00</td>
<td>$12.00</td>
</tr>
<tr>
<td>BLS-18 (1992)</td>
<td>$16.00</td>
<td>$20.00</td>
</tr>
<tr>
<td>BLS-18S (1992)</td>
<td>$8.00</td>
<td>$12.00</td>
</tr>
<tr>
<td>BLS-19 (1993)</td>
<td>$16.00</td>
<td>$20.00</td>
</tr>
<tr>
<td>BLS-19S (1993)</td>
<td>$8.00</td>
<td>$12.00</td>
</tr>
<tr>
<td>BLS-20 (1994)</td>
<td>$16.00</td>
<td>$20.00</td>
</tr>
<tr>
<td>BLS-20S (1994)</td>
<td>$8.00</td>
<td>$12.00</td>
</tr>
<tr>
<td>BLS-21 (1995)</td>
<td>$16.00</td>
<td>$20.00</td>
</tr>
<tr>
<td>BLS-21S (1995)</td>
<td>$8.00</td>
<td>$12.00</td>
</tr>
<tr>
<td>Author Index to BLS 1-16S</td>
<td></td>
<td>$2.00 (includes shipping)</td>
</tr>
</tbody>
</table>

### California Sales Tax

8.50% for residents of San Francisco County  
8.25% for residents of Alameda, Contra Costa, Los Angeles, San Benito, San Mateo, and Santa Cruz counties  
7.75% for residents of Del Norte, Fresno, Imperial, Inyo, Madera, Orange, Riverside, Sacramento, San Bernardino, San Joaquin, Santa Barbara, and Santa Clara counties  
7.50% for residents of Sonoma county  
7.00% for residents of San Diego county  
6.50% for residents of Monterey county  
7.25% for all other California residents

### Shipping and Handling

<table>
<thead>
<tr>
<th></th>
<th>Within USA</th>
<th>Outside USA</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLS 1,2,6,7,9,11,12,13,14,15,16S,17S,18S,19S,20S,21S</td>
<td>$2.00</td>
<td>$3.00</td>
</tr>
<tr>
<td>BLS 3,4,5,8,10,16,17,18,19,20,21</td>
<td>$2.50</td>
<td>$4.00</td>
</tr>
</tbody>
</table>

ORDERS MUST BE ACCOMPANIED BY PAYMENT  
Make checks and money orders payable to the Berkeley Linguistics Society