

Proceedings of the Twenty-Fourth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Phonetics and Phonological Universals (1998)

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-FOURTH ANNUAL MEETING
OF THE
BERKELEY LINGUISTICS SOCIETY
February 14-16, 1998**

**GENERAL SESSION
and
PARASESSION
ON
PHONETICS AND PHONOLOGICAL
UNIVERSALS**

Edited by
Benjamin K. Bergen
Madelaine C. Plauché
and
Ashlee C. Bailey

**Berkeley Linguistics Society
Berkeley, California, USA**



**PROCEEDINGS OF THE TWENTY-FOURTH ANNUAL MEETING
OF THE
BERKELEY LINGUISTICS SOCIETY**

February 14-16, 1998

**GENERAL SESSION
and
PARASESSION
ON
PHONETICS AND PHONOLOGICAL
UNIVERSALS**

edited by

Benjamin K. Bergen

Madelaine C. Plauché

and

Ashlee C. Bailey

Berkeley Linguistics Society

All papers copyright © 1998 by the Berkeley Linguistics Society, Inc.
1203 Dwinelle Hall
University of California
Berkeley, California 94720-2650 USA

All rights reserved.

ISSN 0363-2946

LCCN 76-640143

Printed by Braun-Brumfield, Inc.
P.O. Box 1203
Ann Arbor, Michigan 48108 USA

TABLE OF CONTENTS

Table of contents	iii
Preface	vii

GENERAL SESSION

Indirect grammatical pressure driving language change MARK D. ARNOLD	2
Why children omit function morphemes: Metric vs. syntactic structure MISHA BECKER	13
Person references and references to people in Brazilian Sign Language NORINE BERENZ	24
Language and responsibility: Perceptions of NATO expansion PIOTR CAP	32
The tough construction is neither extraction nor control YOON-SUK CHUNG	44
Reduced relatives: Lexical constraint-based analysis HANA FILIP	57
The discontinuous intonation contour: A case for rethinking intonation contour/intonation phrase isomorphy PHIL GAINES	69
The convergence of 'similarities' and making the best of probabilistic evidence GWANG-YOON GOH	81
The phonetics and phonology of non-modal vowels: A cross-linguistic perspective MATTHEW GORDON	93
Location and direction in Waikuruan languages VERÓNICA GRONDONA	106
Linguistic systems and social models: A case study of Japanese YOKO HASEGAWA	117

Phonetic assessment of tone spreading SUNG-A KIM	129
Motivating morpho-syntactic changes in Turkic subordination JAKLIN KORNFILT	139
The case of the noun phrase in the Finnish adpositional phrase: A discourse-pragmatic account RITVA LAURY	150
<i>Anar</i> 'go', <i>donar</i> 'give', and <i>posar</i> 'put': Shift verbs in Catalan ROSER MORANTE AND GLÒRIA VÁZQUEZ	161
Linguistics and the 'linguistic turn': Language, reality, and knowledge EVE NG	173
On Mohawk ghost vowels: Audibility vs. visibility GRAŻYNA ROWICKA	184
A new look at Japanese conversational styles: (In-)direct speech and turn management SHIE SATO AND SHIGEKU OKAMOTO	195
Discourse-based solutions to quantitative problems in sociolinguistics: The case of men's and women's speech in an Indo-Guyanese village JACK SIDNELL	207
How 'give' and 'receive' provide structure for more abstract notions: The case of benefactives, adversatives, causatives, and passives TOMOKO YAMASHITA SMITH	219
Evidence for the distinction between resultative and consequential serial verbs OSAMUYIMEN THOMPSON STEWART	232
Telescope anaphora JANINE TOOLE	244
The natural classes of two-handed signs JANINE TOOLE AND LINDA UYECHI	256

<i>Bueno</i> : A Spanish interactive discourse marker CATHERINE TRAVIS	268
A theory of agreement and disagreement STEPHEN WECHSLER AND LARISA ZLATIĆ	280
Lexical and grammatical universals as a key to conceptual structures ANNA WIERZBICKA	292

PARASESSION: PHONETICS AND PHONOLOGICAL UNIVERSALS

Testing opposing phonetic structural principles: Polarization and gestural economy VICTORIA B. ANDERSON	309
Perceptual confusions and phonological change: How confused is the listener? PATRICE SPEETER BEDDOR AND RENA ARENS KRAKOW	320
Functional bases of phonological universals: A connectionist approach MARC F. JOANISSE AND MARK S. SEIDENBERG	335
Perception in Optimality Theory: The frugality of the base PAUL LASSETTRE AND PATRICIA DONEGAN	346
Stress effects on CV coarticulation and their implications for phonology JOO-KYEONG LEE	356
Why make life hard? Resolutions to problems of rare and difficult sound types IAN MADDIESON	367
Glottalized sonorants: A phonetic universal? MADELAINE C. PLAUCHÉ, ROSEMARY BEAM DE AZCONA, RUNGPAT ROENGPITYA, AND WILLIAM F. WEIGEL	381
Vowel height: Reconsidering distinctive features DON SALTING	391
Phonological universals: Trilling, voicing, and frication MARIA-JOSEP SOLÉ	403



The following papers were presented at the conference but do not appear in this volume:

Northeast Caucasian as evidence for early Indo-European branching

JOHANNA NICHOLS

Comparing linguistic structure with visual structure

LEN TALMY

Nasalized fricatives? An aerodynamic and acoustic study

JOHN J. OHALA

Words and rules

STEPHEN PINKER

**SPECIAL SESSION:
INDO-EUROPEAN SUBGROUPING AND INTERNAL RELATIONS
FOUND IN BLS 24S**

The linguistic position of Lepontic JOSEPH F. ESKA	2
Adjarian's Law, the Glottalic Theory, and the position of Armenian ANDREW GARRETT	12
The dialectal position of Anatolian within Indo-European H. CRAIG MELCHERT	24
Some consequences of a new proposal for subgrouping the IE family DON RINGE	32

PREFACE

We are delighted to present the proceedings of the Twenty-Fourth Annual Meeting of the Berkeley Linguistics Society. We would like to thank the organizers of BLS conferences past for their insightful advice and the organizers of BLS 25 for so swiftly and energetically taking up the substantial responsibility involved in organizing a conference. We would also be remiss if we didn't thank all of the volunteers who made BLS 24 possible, especially Steve Chang and Lily Liaw. We hope you enjoy the volume as much as we enjoyed preparing it.

BLS 24 Editors
Benjamin K. Bergen
Madelaine C. Plauché
Ashlee C. Bailey

GENERAL SESSION

Indirect Grammatical Pressure Driving Language Change

Mark D. Arnold
University of Maryland

1. Introduction

While many aspects of language change are relatively unexplainable—for example, why some words from a given era disappeared while others became more widely used—other instances of language change can provide insight about the abstract structure of language as well as the language acquisition device (see Lightfoot 1979, 1991). In addition, even though many instances of change can be seen to be the direct consequence of specific sociolinguistic influences (e.g. invasion and political domination by speakers of a non-indigenous language), there are other instances of change for which no such obvious external cause is determinable. It is the latter type of change which I focus on here.

The particular details I will address concern the loss of V-raising in English. The specific aim of the paper is to present new evidence from the Penn-Helsinki Parsed Corpus of Middle English which supports Arnold's (1995, 1997) proposal that the loss of V-raising is related to the (innovation and) spread of preposition stranding, exceptional case marking (ECM), and the possibility of deleting the sentential complementizer *that*.

The broader aim of the paper is to suggest, by way of illustration, that diachronic studies must begin to explore models of change which are far more subtle than those commonly invoked. In other words, the problem of explaining the loss of V-raising provides an excellent example of explanations which are too simplistic in the analysis of cause/effect relationships in language change.

The paper proceeds as follows. Section 2 opens with two preliminary examples of the direct type of cause/effect explanation and then turns to the overly simplistic analyses of the loss of V-raising. Section 3 provides a brief overview of Arnold's proposal, and the new evidence supporting the analysis is presented in section 4.

2. Direct causes for language change

2.1 Examples from A-movement

While any explanation for language change will—at some point in the view back through time—face the impossible task of explaining the background changes which led up to the change under study, there is nonetheless ample opportunity to gain enlightenment from changes which occurred once a given set of details was in place. That is to say, even though we might not know why certain details developed at time T, that does not detract from the insights we might uncover by investigating how the details changed from T to T+1.

In more concrete terms, the history of English provides an excellent opportunity to study the consequences of the collapse of rich morphological paradigms: Old English (OE) had full declensions for both nominal and verbal inflections, and these morphological paradigms began collapsing by early Middle English. Roughly speaking, the nominal system was essentially flattened by the 13th century, and the verbal paradigm was all but lost by the 16th. In terms of explaining these changes in the morphological systems, we have essentially nothing but speculation to offer for why the OE morphological system collapsed; nevertheless, given the collapse, we can take advantage of the history of English to study how other aspects of grammars change when the morphological system simplifies.

A clear ramification of the loss of morphology is that children will lose a direct trigger for various details of the grammar they are acquiring. In other words, whatever else morphology does for a grammatical system, it certainly provides children with overt triggering experiences for the acquisition process. To the extent that morphology provides an overt, phonetically realized presentation of the implementation of various abstract mechanisms of grammar, morphology provides children with a direct cue for those abstract mechanisms. However, if the morphological paradigm collapses, then the acquisition device has no direct trigger for those mechanisms and is left to some other means to determine the presence and/or influence of those mechanisms in the grammar being acquired.

A clear case of such a scenario can be found in the innovation of indirect object passives (see Lightfoot 1981): as long as English had a morphological distinction between Accusative and Dative case, moving the indirect object to subject position (in a passive construction) would violate the Case filter (Chomsky 1981). Given the fact that the morphological paradigm presented children with clear evidence for the Accusative/Dative distinction, the acquisition device would direct evidence to hypothesize an inherent abstract case for indirect objects, and indirect object passives would be impossible; however, once the morphological paradigm was lost, the acquisition device had no necessary reason to posit a distinction between the abstract cases for internal objects, and thus indirect objects could move to subject position without violating the Case filter.

The important point in the above example is the directness of the cause for the change: when the morphological paradigm collapsed, the acquisition device no longer had evidence to assume a distinction in the abstract case system, and thus the possibility for the innovation of indirect passives emerged (though the situation is a bit more complicated as noted in section 3). A similar example concerns P-stranding from A-movement (see van Kemenade 1987): as long as objects of prepositions were morphologically distinct from objects of verbs, the acquisition device had cause to posit an abstract distinction between verbal and prepositional objects, and thus moving the object of a preposition to subject position would violate the Case filter. However, once the morphological paradigm no longer provided a direct trigger for a distinction in the abstract case system, the acquisition device was free to analyze the case of prepositional objects on par with verbal objects, and thus P-stranding in passive constructions (so-called pseudo-passives) became possible.

In both of these cases, the syntactic change is taken to be the direct consequence of the loss of the morphological trigger for a particular detail of the abstract mechanisms of grammar. This type of direct analysis has also been used to explain the loss of V-raising in English, a point to which we now turn.

2.2 Direct cause(s) for the loss of V-raising

Given that English's verbal inflectional paradigm had a history similar to the nominal paradigm, Roberts (1993) and Rohrbacher (1994) provide essentially the same type of explanation for the loss of V-raising as Lightfoot and van Kemenade provided for the novel passive forms. Though the details of the morphology-based analyses differ, both face significant empirical and theoretical problems, discussed below. Likewise, Watanabe's (1993) proposal, though not stated in terms of the collapse of the morphological paradigm, is similar to the morphology approaches in that it assumes that the loss of V-raising is directly linked to the internal structure of the acquisition device. Like the morphology-based approaches, Watanabe's analysis faces serious empirical and theoretical problems.

As for the morphology-based approaches, the assumption in both studies is that a V head is raised to the Infl head in languages with rich verbal morphology; the corollary assumption is that such rich verbal paradigms will be a direct trigger for the acquisition device to hypothesize V-raising. The crucial assumption for these proposals is that the trigger for the acquisition of V-raising is the richness of the morphological paradigm—not syntactic word order effects such as subject/verb inversion.

Given the assumption about the central role of the verbal morphology for the acquisition of V-raising, both authors engaged in cross-linguistic studies of verbal inflectional paradigms in languages with and without V-raising. Of passing interest is the fact that these inductive studies led to different conclusions about what constitutes 'enough' morphology to trigger V-raising, but of much more importance is the way each author deals with a rather glaring historical fact: written evidence of V-raising continues to appear in texts for well over 100 years beyond the disappearance of the morphological trigger for V-raising (see Warner 1997).

In recognizing the chronology problem, Roberts proposes that children continued to hypothesize an abstract morphological marker which would have driven V-raising in the emerging grammar. Notice, however, that this view leaves us with a non-trivial question: if children in the Early Modern period hypothesized the abstract morphology necessary for V-raising, why couldn't subsequent generations of children continue to propose it? Presumably the trigger for the abstract morphology disappeared, but this is simply restating the problem.

Rohrbacher approaches the chronological disparity with the opposite impulse; he concludes that as soon as the morphological trigger for V-raising had disappeared, V-raising was no longer a productive mechanism for English speakers. Any evidence for the continued use of V-raising was purely archaic in nature, i.e. V-raising was used by writers out of familiarity with the older forms of the language, but given that the morphological trigger for V-raising was gone, so too was V-raising.

With this analysis, Rohrbacher proposes to account for the increased use of periphrastic *do*: given the loss of V-raising, speakers' grammars were forced to adopt the periphrastic form to encode Tense. However, it turns out that patterns during the early use of *do* actually point to a problem in the analysis; during the relevant period, *do* was used more frequently in formal texts, and V-raising was used more frequently in informal texts (see Ellegard 1953). This pattern contradicts what we would expect to find given Rohrbacher's scenario, i.e. if V-raising were no longer a productive mechanism in a writer's grammar, then we would expect *do* to be used more frequently in personal texts and V-raising to be used more in formal texts.

The empirical problems (for both morphology-based accounts) are underscored by a theoretical issue, namely motivating the assumption that the acquisition of syntactic mechanisms is triggered directly—and solely—by the richness of morphological paradigms. While it is true that there is some cross-linguistic evidence to suggest a correlation between richness of morphology and word order facts, i.e. languages with rich nominal inflections apparently allow 'freer' word order, there are also counter-examples, e.g. rampant scrambling in (essentially) morphology-less Chinese. Without a principled account of why/how syntactic mechanisms are triggered by rich morphological paradigms, the morphology-based approaches simply restate the problem.

A different sort of analysis, proposed by Watanabe, faces no less significant empirical or theoretical problems. Watanabe, adopting Chomsky's (1995)

Minimalist Program, assumes that the presence of periphrastic *do* in the linguistic input caused the loss of V-raising because *do* allowed V-raising to Procrastinate and thus provided more economical derivations. There are two empirical problems for this view. The first comes from observations made by van Kampen (1997) regarding the acquisition of Dutch: children acquiring Dutch over-generalize their early use of *doen* only to abandon it and converge on the adult forms with V-raising. While Watanabe's proposal might account for the early over-generalization of *doen*, the following question arises: if Dutch children can acquire V-raising, thereby (apparently) overriding the economy afforded by *doen*, why couldn't children acquiring English during the early Modern period do the same thing?

The other empirical problem is much broader and comes from the cross-linguistic observation that synthetic forms, when available in a language, generally block the use of a periphrastic form (see Poser 1992). This generalization is exactly the opposite of what we would expect if Watanabe's approach to language change were correct; given Watanabe's proposal concerning *do*, we would expect the general pattern to be that the economy provided by periphrasis (*vis a vis* overt movement) would typically block the use of the synthetic form because the synthetic form—being an instance of overt head movement—would be less economical than the periphrastic form. Thus, Watanabe's analysis of *do* flies in the face of the cross-linguistic pattern concerning periphrasis versus synthesis.

The theoretical problem is that Watanabe must allow the economy of derivations with *do* to be compared to derivations with V-raising, a comparison not allowed in the strict lexicalist system proposed by Chomsky. While it is true that the strict lexicalist hypothesis could be modified (as Arnold 1996 proposes), Watanabe doesn't address the issue, and thus the analysis encounters the following question: if the nature of the system is such that the relative economy of *do* cannot be compared to the relative economy of V-raising, and if children are receiving more input with V-raising than input with *do*, then why/how would children ever stop acquiring V-raising? Answering that question by saying that the linguistic input changed such that *do* was used more frequently than V-raising simply restates the problem.

With all that said, the important point is that the nature of the (attempted) explanation is very direct: 1.) the acquisition device is built to prefer economical derivations, 2.) *do* provided more economical derivations, 3.) V-raising was lost. This particular type of direct explanation, like the morphology-based approaches, encounters its own set of empirical and theoretical problems, but there is another set of problems—a sort of empirically based theoretical problem—which all three analyses face, namely that there is a wide range of historical facts related to the spread of *do* and the loss of V-raising which the above proposals are categorically unable to address. From a consideration of the full range of facts, discussed below, the need to entertain the possibility of a more indirect cause for the loss of V-raising becomes clear.

3. Towards a less direct cause of change

As a matter of illustrating the interconnectedness of the historical facts which motivate a unified analysis of various changes which took place during Middle English, the following list presents the chronological details addressed in Arnold's account of the loss of V-raising:

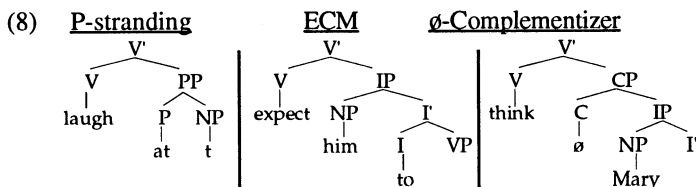
- (1) transitive verbs which had selected Dative complements in Old English (e.g. *help*) start showing novel passive forms (e.g. *The men were helped*)

in the 13th century; however, indirect objects (also marked Dative in Old English) did not appear as the subject of passive for 100-150 years after the appearance of the novel direct object passives (see Denison 1993);

- (2) the eventual innovation of indirect object passives (IOPs) ((1) above) aligned with the statistically significant use of periphrastic *do* : both first appeared with clear regularity at the end of the 14th century, and both remained relatively rare until the end of the 15th century (see Ellegard 1953 on *do*, and Denison 1993 on IOPs);
- (3) the disappearance of quasi-double object constructions (e.g. *Mary gave to John a book*) occurs in the same century in which verb movement had clearly given way to the widespread use of *do*, namely the latter half of the 16th century (see Visser 1963-1973 on double object constructions and Ellegard 1953 on *do*);
- (4) the innovation of complex prepositional passives (e.g. *John was taken advantage of*) paralleled the delayed innovation of IOPs (see Denison 1993);
- (5) deletion of *that* in *that*-trace contexts hits 100% in the 16th century (see Bergh and Seppanen 1992);
- (6) deletion of *that* in ECM and control structures (e.g. *Mary convinced him that to go*, *Mary expected that him to go*) approaches the modern standard in the 16th century (see Visser 1963-1973);
- (7) from 1400-1700, the relative frequency of periphrastic *do* was higher with transitive verbs than with intransitives; with respect to different sentence types, the relative frequency of *do* was highest with negative questions, then affirmative questions, then negative declaratives and was lowest in affirmative declaratives (see Ellegard 1953).

Arnold proposes an incorporation analysis of P-stranding, ECM, and *that*-deletion which allows the various diachronic developments outlined in (1-7) to be seen as the various consequences of a single underlying development rather than accidental similarities. Moreover, and crucially, neither the morphology-based approaches nor the economy approach to the loss of verb movement are capable of providing any insight about the chronological similarities outlined in (1-7).

The basic idea is that the spread of the novel incorporation constructions led to an increased use of *do*: *do* allowed the V head to remain in VP, thus allowing for a shorter movement when incorporating the relevant head into the verb.



For the various constructions in (8), movement of the verb out of VP would require LF incorporation of either P, infinitival-*to*, or the null complementizer to cross the VP projection. However, as illustrated in (9)—using a hypothetical P-stranding structure—use of *do* in (9b) allows incorporation to occur within the VP:

(9) a. neg. Q with V-to-I-to-C:

-P must cross three X^{\max} for incorporation into V

[CP [V_i+I]_j [IP t_j [NegP [VP t_i [PP P [NP (e)]]]]]]

b. neg. Q with *do*:

-P crosses no X^{\max} for incorporation into V

[CP do_j [IP t_j [NegP [VP V [PP P [NP (e)]]]]]]

Given that the derivation with *do* requires shorter movement in order to converge, it blocks the derivation with verb movement. Under this analysis, we can understand why *do* spread through the language—and we have an explanation for the patterns found by Ellegard, in the following way. First, since a transitive verb is more likely than an intransitive to have a complement containing an element which will incorporate into it, *do* was used more frequently with transitives than with intransitives. Second, the distinctions in the relative frequency of *do* in different sentence types follow from the overall degree of complexity in the different structures; (10a-d) represent structures with overt verb movement and a stranded P which must incorporate into the raised V head:

(10) a. Neg. Q:

-two instances of form chain, incorporation crosses three X^{\max}

[CP [V_i+I]_j [IP t_j [NegP [VP t_i [PP P [NP]]]]]]

b. Aff. Q:

-two instances of form chain, incorporation crosses two X^{\max}

[CP [V_i+I]_j [IP t_j [VP t_i [PP P [NP]]]]]]

c. Neg. Decl.:

-one instance of form chain, incorporation crosses two X^{\max}

[IP V_i [NegP [VP t_i [PP P [NP]]]]]]

d. Aff. Decl.:

-one instance of form chain, incorporation crosses three X^{\max}

[IP V_i [VP t_i [PP P [NP]]]]]]

In sum, the incorporation analysis provides an explanation for the spread of *do* during late Middle English and thereby provides an explanation for the decrease in the frequency of syntactic triggers for V-raising, i.e. the increased use of *do* meant fewer instances of subject/verb inversion or instances of a tensed verb preceding negation. Thus, the syntactic trigger for V-raising disappeared as a consequence of the spread of unrelated constructions, e.g. P-stranding, ECM, and *that*-deletion. Under this view, the 'direct' cause for the loss of V-raising was simply the loss of the syntactic trigger, but rather than simply restating the problem, the incorporation analysis allows us to understand that the loss of the syntactic trigger was itself a consequence of the spread of the novel incorporation mechanism. Thus, the

'indirect' cause for the loss of V-raising was the spread of P-stranding, ECM, and *that*-deletion.

4. Extending the analysis and finding new evidence

The proposal outlined above was originally developed to address the spread of *do*; however, as the majority of the texts in the Penn-Helsinki Corpus predates the widespread use of *do*, the specific details of the proposal must be slightly modified in order to test the theory.

To that end, recall that the operative force in the analysis is the economy provided by *do* in constructions with P-stranding, ECM, or *that*-deletion: *do* allowed the verb to remain in VP and thus allowed for shorter LF incorporation of the relevant functional head. If the proposal is right, we expect the following situation to emerge: in those cases where a writer could select between two grammatical forms, e.g. P-stranding versus pied-piping, the presence of a V head in VP would have increased the likelihood of the use of P-stranding; likewise, in clauses with clear evidence that the verb had raised out of VP, P-stranding would have been disfavored.

Of course, the problem is that in many clauses with a simple tensed verb, there is no definitive way to establish whether the particular sentence corresponded to an instance of V-raising or affix-hopping. In other words, *Mary likes John* could correspond to either its (parsed/analyzed) Middle English form, as in (11a), or its modern form, as in (11b):

- (11) a. [Mary [likes_i [t_i [John]]]]
 b. [Mary [(pres) [likes [John]]]]

The necessary approach, therefore, is to focus the research on clauses which provide clear evidence of the location of the V head. For the purposes of automatic searches in the corpus, there are two syntactic details which I take to provide clear evidence of either V-raising or V *in situ*: 1.) following Pollock (1989), the location of the Neg head relative to a finite V, or 2.) the presence of a non-finite V. (The second possibility, i.e. non-finite V, occurs either when tense is carried on a modal/auxiliary or the clause is infinitival.)

4.1 Some details about the Corpus

The Penn-Helsinki Parsed Corpus of Middle English is a collection of ASCII files which contain annotated sentences from a variety of Middle English texts. Each sentence from a given manuscript constitutes a separate token, and the annotation scheme provides low-level syntactic parsing as well as part of speech tags and locations of traces/elisions, e.g. s = subject, vt = tensed verb, a = auxiliary, p = preposition, %- = trace/empty. (12) provides a sample token:

- (12) (([f Al men 1[L [c-1 +tat] %s-1 r % [at wyll] [v her] [p of +te sege of Jerusalem], L]1] [I her] %[s +ge]% [at may] [v her] [p of gret meraculs 2[L [c-2 +tat] %d-2 r % [s almytty God] [vt wro+gt] 2.1[UP [a to] [v schow] [d his goodnys UP]2.1 L]2 and of gret vengans 3[L [c-3 +tat] %d-3 r % [s he] [vt toke] [p for syn] . L]3])(SIEGE,70.1)

Also, Anthony Kroch and Ann Taylor, the authors of the parsed corpus, provide a program which divides all of the sentences in a file into all of the separate

clauses which make up the sentences. (13) illustrates the result when the token in (12) is divided into separate clauses; the research reported here was conducted on files in which each clause is a separate token, as in (13).

(13) ([f Al men %L%] [l her] %[s +ge]% [at may] [v her] [p of gret meraculs %L% and of gret vengans %L%])(SIEGE,70.1)

(1[L [c-1 +tat] %s-1 r % [at wyll] [v her] [p of +te sege of Jerusalem] , L]1)(SIEGE,70.1)

(2[L [c-2 +tat] %d-2 r % [s almytty God] [vt wro+gt] %UP% L]2)(SIEGE,70.1)

(3[L [c-3 +tat] %d-3 r % [s he] [vt toke] [p for syn] . L]3)(SIEGE,70.1)

(2.1[UP [a to] [v schow] [d his goodnys] UP]2.1)(SIEGE,70.1)

Additionally, the files in the corpus are divided chronologically into four periods: M1 (being the earliest) through M4. This distinction is particularly helpful for testing the predictions of the incorporation analysis for the following reason. Given the proposal that the novel incorporation constructions (P-stranding, ECM, and *that*-deletion) spread during Middle English, it is insightful to compare the patterns found in the earliest texts to those in the latest. Throughout the discussion, I will make clear which subsets of files were searched for given patterns.

A final note concerns the particular searches which were conducted. First, in order to research the correlation of *that*-deletion with verb movement, the most certain diagnostic was the position of the tensed verb relative to Neg; as noted above, when the tensed verb precedes Neg, I assume that the sentence exhibits V-raising. Second, for P-stranding, the diagnostic used was the presence of a non-finite V; I assume that a non-finite V occurs in the VP and therefore provides a very local incorporation site for the stranded P. Finally, due to the coding conventions in the corpus, there is no automatic way to distinguish ECM from control structures; thus, testing the predictions *vis a vis* ECM is left for future work.

4.2 *That*-deletion

Before turning to the specific correlation between *that*-deletion and verb movement (as indicated by Neg), it is worth considering the basic pattern of *that*-deletion in the earliest texts compared to the latest. Of the eleven files in the earliest period, only six have an example of deleted *that*, and the overall relative frequency of *that*-deletion is 12% (17/138). By contrast, all 14 files with M4 designation contain at least one example of deleted *that*, and the overall frequency is 39% (247/641). These general numbers provide clear evidence that the option of deleting the sentential complementizer *that* spread during Middle English.

Turning now to the more specific prediction, the search required 'rebuilding' of clauses such that the status of the complementizer (overt versus null) could be correlated with the position of the governing verb. In other words, given that the relevant detail for *that*-deletion is the position of the verb which governs the complementizer, and given that the sentences in each file had been divided into separate clauses, it was necessary to concatenate a *that*-clause with the clause it was originally embedded in. Once the clauses were appropriately rebuilt, the search

could then determine the correlation between certain verb movement—as indicated by a tensed verb preceding Neg—and deleted *that*.

The findings are quite striking: of the 27 tokens in which Neg intervenes between the tensed verb and the complementizer position, there are only three cases of *that*-deletion—and all three are examples of the same construction using a form of *would* such as *I would not her falseness be known* . Even if we make nothing of the fact that all three tokens are examples of the same construction, the rate of *that*-deletion when Neg intervenes between the verb and the complementizer is only 10% (3/27), well below the average of 30% (366/1228) for the periods (M3 and M4) in which the tokens are found.

However, further investigation of the *would* construction suggests that such examples of *that*-deletion are themselves special cases requiring further study, for the following reason. A search for all relevant *would* constructions reveals that *that*-deletion in such constructions is 58% (18/31), far above the 30% average for the M3 and M4 periods. Based on the higher than average rate of *that*-deletion in *would* constructions, I set them aside for further study. Having set aside the tokens with *would*, the final result offers strong support for the theory being tested: there are no instances of *that*-deletion when Neg intervenes between the governing verb and the complementizer position. This fact follows directly from an analysis in which the possibility of deleting *that* is restricted in those cases when the verb has raised out of VP.

4.3 P-stranding

For P-stranding, the automatic searches become a bit more complex for two related reasons. First, P-stranding was quite productive in relative clauses in Old English; thus the early texts—when searched without regard to clause type—actually have a higher rate of P-stranding than of pied-piping, contrary to what we might expect given the standard view that the productivity of P-stranding was a Middle English innovation. Second, given the high rate of P-stranding in the earliest texts, a superficial glance at the corpus suggests that P-stranding was actually disappearing during Middle English, contrary to the standard view.

However, by separating relative clauses from non-relatives, it becomes clear that P-stranding in non-relative clauses increases during Middle English, in line with the standard view. The conclusion to be drawn from the preliminary work is that the searches should be targeted to non-relative clauses, for it is in the non-relatives that we can see the patterns heading towards the modern norms. With that in mind, we turn to the results of searching for P-stranding in non-relative clauses.

Recall the prediction: the presence of a V head in VP will increase the likelihood that P-stranding will occur. In other words, if the proposal for the incorporation analysis of P-stranding is right, then there should be a higher than average rate of non-finite verbs in those clauses which contain stranded-Ps. As Table 1 shows, the prediction obtains: the frequency of non-finite verbs in non-relative clauses is nearly three times higher than normal when the (non-relative) clause also contains a stranded-P.

	<u>all non-relative clausesnon-relatives with stranded-P</u>	
M1	24.59	65.00
M4	29.44	83.33

Table 1: Percentage non-finite V in non-relatives, with and without stranded-P

This same pattern obtains when the entire corpus is searched:

	<u>all non-relative clauses</u>	<u>non-relatives with stranded-P</u>
Corpus	27.56	71.01

Table 2: Percentage non-finite V in non-relatives, with and without stranded-P

Tables 1 and 2 illustrate the accuracy of the prediction when clause type is controlled for, i.e. setting aside relative clauses allows the pattern to stand out. Another similar research control arises if we consider that P-stranding due to A-movement was not at all possible in Old English, and thus another way to test the incorporation analysis would be to target so-called pseudo-passives, i.e. passive constructions in which A-movement creates a stranded-P.

Again, we expect that non-finite verbs should occur with a higher than normal rate in clauses with pseudo-passives. As before, the results are very compelling: the relevant verb is non-finite in all 35 tokens in which an overt subject is co-indexed with the trace of a stranded-P. Additionally, when the search is broadened to allow for empty subjects, i.e. A'-movement of the subject of a passive, only three of the additional nine tokens contain finite verbs. Thus, even without attempting to explain away the apparent counter-examples, the frequency of non-finites in pseudo-passives is 93%, well above the normal frequency for non-finites of 27.56%.

5. Conclusion

To close, there are two issues which are important to keep in mind. First, the findings are very robust: with respect to *that*-deletion, recall that there were no instances of deleted *that* when Neg intervened between the finite verb and the complementizer position; with respect to P-stranding, the frequency of non-finite verbs was three times the normal rate when the clause contained a stranded-P, and for the most clear-cut examples of pseudo-passive, every example of P-stranding co-occurred with a non-finite verb. The robustness of the findings strongly supports the incorporation analysis.

Second, these robust findings were found precisely because the incorporation analysis predicts a correlation between these novel Middle English constructions and the potential for verb movement. By contrast, in addition to the fact that the morphology-based accounts of the loss of verb movement are unable to address the original diachronic details outlined in (1-7) above, these new findings concerning the correlation of verb movement with other constructions are completely mysterious under the morphological analyses. Likewise, building an economy bias into the acquisition device in order to explain the spread of *do* allows for no explanation of the patterns presented here.

On a broader level, the findings provide evidence for a type of indirect cause/effect relationship we should be aware of when investigating the process of language change, for while it is true that there is a certain attractive certainty to the type of explanations which emerge from positing direct links between e.g. the acquisition of morphology and the acquisition of syntax, there is little reason to imagine that either the human language faculty or the acquisition device should be limited to such directly correlated systems. In other words, though the direct answers might appeal to our theoretical aesthetic, the data suggest a far more subtle analysis.

References

- Arnold, M. 1995. Case, Periphrastic *do*, and the Loss of Verb Movement in English. PhD dissertation, University of Maryland.
- Arnold, M. 1996. The History of Periphrastic *do*: An Argument for Lexical Insertion at Spell-Out, in the Proceedings of the 26th Annual Meeting of the North East Linguistic Society. GLSA, University of Massachusetts: Amherst.
- Arnold, M. 1997. A Unified Analysis of Preposition Stranding, ECM, and *that*-deletion, in the Proceedings of the 15th Annual Meeting of the West Coast Conference on Formal Linguistics. CLSI: Stanford.
- Bergh, G. and A. Seppanen. 1992. Subject Extraction in English: The Use of the *That*-Complementizer, in Fernandez, Fuster, and Calvo, (eds.), English Historical Linguistics 1992. Philadelphia: Benjamins.
- Chomsky, N. 1981. Lectures on Government and Binding. Dordrecht: Foris.
- Chomsky, N. 1995. Minimalist Program. Cambridge: MIT Press.
- Denison, D. 1993. English Historical Syntax: Verbal Constructions. New York: Longman.
- Ellegard, A. 1953. The Auxiliary *do*: The Establishment and Regulation of its Use in English, Behre, F. (ed.), Gothenburg Studies in English. Stockholm: Almqvist and Wiksell.
- Kampen, J. van. 1997. First Steps in Wh-movement. Wageningen: Ponsen and Looijen.
- Kemenade, A. van. 1987. Syntactic Case and Morphological Case in the History of English. Dordrecht: Foris.
- Lightfoot, D. 1979. Principles of Diachronic Syntax. New York: Cambridge University Press.
- Lightfoot, D. 1981. History of Noun Phrase Movement, in Baker, C. and J. McCarthy (eds.), Logical Problem of Language Acquisition. Cambridge: MIT Press.
- Lightfoot, D. 1991. How to Set Parameters: Arguments from Language Change. Cambridge: MIT Press.
- Pollock, J.-Y. 1989. Verb Movement, UG, and the Structure of IP, Linguistic Inquiry, 20: 365-424.
- Poser, W. 1992. Extending Morphological Blocking into Syntax in Sag, I. and A. Szabolcsi (eds.) Lexical Matters, 111-130. Stanford: CLSI 24.
- Roberts, I. 1993. Verbs and Diachronic Syntax: A Comparative History of English and French. Boston: Kluwer.
- Rohrbacher, B. 1994. The Germanic VO Languages and the Full Paradigm: A Theory of V-to-I Raising. PhD dissertation, University of Massachusetts, Amherst.
- Visser, T. 1963-1973. An Historical Syntax of the English Language. Leiden: E.J. Brill.
- Warner, A. 1997. The Structure of Parametric Change, and V-movement in the History of English, in Kemenade, A. van and N. Vincent (eds.) Parameters of Morphosyntactic Change. Cambridge University Press.
- Watanabe, A. 1993. Agr-Based Case Theory and Its Interaction with the A-Bar System. PhD dissertation, MIT.

Why Children Omit Function Morphemes: Metric vs. Syntactic Structure

Misha Becker
UCLA

1. Introduction

It is an interesting and noticeable fact that in the early stages of language production, children omit certain elements from their speech. Upon closer inspection, it becomes clear that children regularly omit certain function morphemes, as opposed to content words. Function morphemes (including determiners, pronouns, auxiliary verbs, and inflection morphemes) form a natural class from both a phonological and a syntactic perspective: phonologically speaking, in many languages (including English), function morphemes tend to be unstressed and monosyllabic; syntactically speaking, function morphemes are the heads of functional projections, which appear notoriously late in children's speech (Radford 1988). Thus, we can ask whether children omit functional elements for phonological or syntactic reasons.

If children omit function morphemes for reasons related to metrical stress, we should find similarities between unstressed function morphemes and unstressed lexical syllables, and, depending on our assumptions about metric structure, we might find asymmetries in the omission rates of subject vs. object determiners.¹ On the other hand, if children omit function morphemes for syntactic reasons, according to some accounts (e.g. Hoekstra, Hyams & Becker 1997, Clahsen et al. 1996) we should find correlations between omissions in the subject and the predicate, following the structural contingency known as spec-head agreement. In fact we find both of these patterns, which suggests that both metrical and syntactic processes constrain children's output forms, but in different ways.

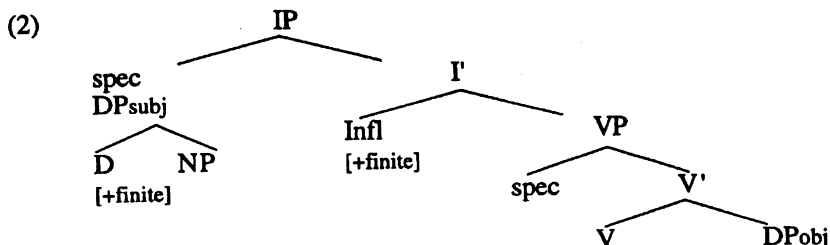
1.1 Overview of predictions

Many researchers have noticed the following pattern of syllable omission in young children's speech: while children show an overall preference for preserving strong (stressed) syllables, they also tend to preserve weak syllables in trochaic feet (S-w) more than weak syllables that are either in an iambic foot or unfooted. This fact has led a number of linguists to suggest that children's output forms are constrained by a so-called "trochaic template", such that weak syllables in trochaic feet are favored over all other weak syllables (Allen & Hawkins 1980, Wijnen, Krikhaar & den Os 1994, Gerken et al. 1990, Gerken 1991, 1994a, 1994b, 1996, i.a.).

To illustrate the predicted pattern, the determiner in (1a) should be omitted more frequently than the pronoun in the same sentence, because the subject determiner, unlike the object pronoun, fails to form a trochaic foot with anything.² Comparing (1b) and (1c), the object determiner in (1c) should be omitted more frequently than the one in (1b), because the syllabic inflectional morpheme on the verb in (1c) (namely *-es*) forms a trochaic foot with the verb root, forcing the determiner to form an iambic foot with the object noun (*duck*). (Stress is indicated by all caps.)

- (1)
- a. [The BOY] [HUGGED him]
 - b. [BERT] [PUSHED the] [DUCK]
 - c. [BERT] [PUSHes] [the DUCK]

A syntactic account of children's speech also predicts a certain pattern of function morpheme omissions. Within recent generative theories, the subject is represented in the specifier position of the projection whose head is the Infl (inflection) node, as shown in (2).



Based on the idea that a specifier and its head should "agree" in features (known as spec-head agreement), we would expect the subject and the verb to agree in their specification (or unspecification) of, for example, finiteness features. A proposal by Hoekstra, Hyams & Becker (1997) suggests, specifically, that the omission of a subject pronoun (i.e. a null subject) or a missing determiner (a bare N subject) is indicative of a lack of "nominal finiteness" in the subject, and should thus correspond to a lack of finiteness in the verb (as evidenced through a lack of verbal inflectional morphemes, such as 3sg -s). In brief, the expected correlation hinges on finiteness features in the verb (shown in (3)): if the verb is finite, the subject should be fully specified; if the verb is nonfinite, we might expect either a specified or an unspecified subject. For a detailed discussion of "nominal finiteness" see Hoekstra & Hyams (1996).

- (3)
- a. finite V -> full DP subject, *null subject/bare N
 - b. nonfinite V -> full DP subject, null subject/bare N³

2. Object determiner and pronoun omissions

Let us look at how the metrical and syntactic approaches square with children's omissions of object pronouns and determiners.

2.1 Gerken's data

In children's imitative reproductions of sentences of the sort in (1), Gerken (1991, 1994, 1996) found that children in fact omitted an object pronoun or determiner less often if it occurred in a trochaic foot. In example (4) I give some of the results from a study by Gerken (1991, average age 2;3), where the percentages below the unstressed words indicate the rate of omission:

- (4) a. [she KISSED] [the DOG]
 39% 28%
 b. [the DOG] [KISSED her]
 39% 0%
 c. [PETE] [KISSED the] [DOG]
 11%

As we see in (4), the object pronoun in (4b) is never omitted, and the object determiner in (4c) is omitted only 11% of the time; both of these words occur in trochaic feet. In contrast, the object determiner in (4a) is omitted much more frequently: at 28% of the time.

More specifically, Gerken (1996) shows that children (age 2;1-2;3) display a contrast in omission frequencies of the object determiner in the near minimal pair given in (5):

- (5) a. he [KICKS the] PIG
 16%
 b. he [CATCHes] [the PIG]
 29% 48%

As predicted by Gerken's metrical hypothesis, the children in her experiment omitted the object determiner in pattern (5b) more frequently (48% of the time) than the one in pattern (5a) (only 16%). Note that there is no syntactic difference between the two verbs, so no syntactic account would predict a difference in the omission rates between the determiners in the two conditions.

2.2 Progressive verbs

Note that monosyllabic verbs with progressive aspect (e.g. *running*) constitute a trochaic foot by themselves.

- (6) [RUNning]
 S w

Thus, given the facts outlined above for the omission of object determiners and pronouns following finite verbs, we should expect that object determiners & pronouns following progressives should be omitted as frequently as in (5b), as opposed to (5a). The expected omissions in this condition, according to the metrical hypothesis, are given in (7), where the morphemes with high expected rates of omission are underlined:

- (7) a. [The BOY] is [HUGging] [the DUCK].
 b. [The BOY] is [HUGging] him.

I searched 12 files of spontaneous (natural) speech from the Adam corpus (Brown 1973, from the CHILDES database, MacWhinney & Snow 1985), in which Adam's age was 2;3 to 3;7 (years;months). I found a total of only 18 examples of sentences in which Adam produced a progressive verb with an unambiguous singular count noun object (e.g. *eating (an) apple*). Of those 18 sentences, 13 (or 72%) contained an overt object determiner or pronoun, and only 5 (28%) lacked an object determiner or pronoun. I show the results in Table 1, comparing this result

with his omissions of object determiners or pronouns following plain, monosyllabic verbs. I give example sentences for each category.

Table 1. Adam's omission of object determiner/pronoun by verb type (age 2;3-3;7)

object type	progressives: n (%)	plain Vs: n (%)
overt det/prn	13 (72%) dey are holding <u>it</u> Mommy making <u>a</u> road	79 (85%) Adam do <u>it</u> dat fits <u>my</u> train
null det/prn	5 (28%) dis man is getting __ doggie getting __ meat	14 (15%) baby have __ tooth Adam see __ sun

$\chi^2 < 1$

Table 1 shows that Adam omits object determiners and pronouns after progressive verb forms at a higher rate (28%) than after plain, monosyllabic verbs (15%) (e.g. *eat(s) (an) apple*), as would be predicted by the metrical account.

2.3 A finiteness effect

Although the predictions of the metrical hypothesis for object omission are confirmed by these data, the metrical approach does not give us the whole story. In particular, there is a finiteness effect on Adam's omissions of object determiners and pronouns. That is, while 14 (25%) of Adam's sentences containing nonfinite plain verbs lacked an object determiner or pronoun, none of his finite utterances did. The breakdown by finiteness is given in Table 2.

Table 2: Adam's object det/prn omissions for finite vs. nonfinite verbs (age 2;3-3;7)⁴

object type	finite V	nonfinite V
overt det/prn	38 (100%) he takes <u>it</u>	41 (75%) kitty eat <u>a</u> apple all up
null det/prn	0 (0%)	14 (25%) have __ mouth

$\chi^2 = 9.48, p \leq .005$

The same pattern of distribution with respect to finiteness occurs with Adam's progressive verbs, as shown in Table 3.⁵

Table 3: Adam's object det/prn omissions for finite vs. nonfinite progressives (age 2;3-3;7)⁶

object type	finite progressive (is)	nonfin. progressive (0-is)
overt det/prn	4 (100%) dey <u>are</u> holding <u>it</u>	9 (64%) Mommy __ making <u>a</u> road
null det/prn	0 (0%)	5 (36%) Adam __ taking __ diamond

$\chi^2 < 1$

Another child, Nina (Suppes 1973), shows a similar pattern of object determiner and pronoun omission to that of Adam. Searching 20 files, in which Nina ranged in age from 2;0 to 2;6, Nina omitted an object determiner or pronoun

after a plain, monosyllabic verb about 10% of the time, but about 20% of the time following progressive verbs.⁷ I give the results in Table 4:

Table 4: Nina's omissions of object determiner & pronouns by verb type

	progressives	plain verbs
overt object det/prn	140 (80%)	1089 (90%)
null object det/prn	37 (20%)	115 (10%)

$$\chi^2 = 19.16, p \leq .001$$

Nina shows a slight finiteness effect, but not a significant one.

Nina's and Adam's data came from spontaneous, natural speech samples, while Gerken's results come from experiments involving elicited speech (in particular, imitative speech), making comparisons difficult. In order to better compare these sets of data, I conducted a small pilot study, adopting Gerken's methods. I measured omission rates for object determiners and pronouns following progressive verbs. The study involved 8 native English-speaking children between ages 2;0 and 2;6. Each child was tested on a total of 36 different sentences, broken down into two sets with two randomizations of each set. Each sentence was recorded onto a tape, spoken with normal intonation and stress, by a native speaker of English who was blind to the task.

I played the tape for each child individually, stopping the tape after each sentence. I then asked the child to repeat exactly what he or she had heard on the tape (a sample was given beforehand to ensure that the child understood the task). The sentence was replayed up to 3 times if the child gave no response, and if there was still no response after 3 repetitions, the next sentence was given. All sessions were recorded on another tape, and after each session, these tapes were checked against the results recorded during the actual experiment.

The predicates of the test sentences and the results are given in Table 5. I compared productions of utterances containing the verbs *hug* and *push*: These verbs differ in their metric structure in present tense (*hugs* vs. *pushes*), they are both monosyllabic in past tense (*hugged*, *pushed*), and they are both disyllabic in progressive aspect (*hugging*, *pushing*). Although the total number of omissions I recorded was quite small, the trend followed the pattern that Gerken reports. Object determiners and pronouns following progressives were omitted at roughly the same rate as the object determiners and pronouns following plain verbs with syllabic inflection, and more frequently than object determiners and pronouns after finite verbs with nonsyllabic inflection.

Table 5: object det/prn omissions following di- vs. monosyllabic verbs (age 2;0-2;6)

[V+es/ing] + the/him	n	%	[V+the/him]	n	%
[PUSHes] the	2	10%	[PUSHED the]	0	0%
[PUSHes] him	1	5%	[PUSHED him]	1	5%
[PUSHing] the	2	10%	[HUGGED the]	1	5%
[PUSHing] him	1	5%	[HUGGED him]	0	0%
[HUGging] the	1	5%	[HUGS the]	0	0%
[HUGging] him	1	5%	[HUGS him]	0	0%

As shown in Table 5, object determiners and pronouns following disyllabic verbs were omitted slightly more frequently than those following monosyllabic verbs (e.g. *pushes the* vs. *pushed the*). There was one exception involving an object pronoun: *pushes him* vs. *pushed him*.

As for a connection between verb finiteness and object determiner or pronoun omission, we find a similar correlation with finiteness as was found in Adam's natural speech data. That is, the children in my pilot study produced sentences with null object determiners and pronouns both with finite and nonfinite verbs (i.e., they produced a nonfinite verb if they omitted the verbal inflectional morpheme in the utterance they repeated). However, a much higher percentage of finite utterances contained an overt object determiner or pronoun (97%) than nonfinite utterances (86%). I give the relevant figures in Table 6.

Table 6: finiteness and object det/prn omission: data from pilot study (2;0-2;6)

	finite verbs	nonfinite verbs
overt object det/prn	145 (97%)	38 (86%)
null object det/prn	4 (3%)	6 (14%)

$$\chi^2 = 6.21, p \leq .05$$

To summarize this section, Gerken's experimental data suggest that children preserve weak syllables at a higher rate if they occur in trochaic feet, and the results of my small pilot study appear to support her claims. However both the spontaneous speech data I studied and my experimental data suggest a possible finiteness effect. Thus a syntactic correlation between verb finiteness and omission of object dets/prns cannot be ruled out.

3. Subject determiner and pronoun omissions

Let us turn now to the predictions regarding omission patterns of function morphemes in the subject. Both the metrical approach and the syntactic approach described above predict that subject determiners and pronouns should be omitted, but the precise predictions each theory makes are different. The metrical view makes the straightforward prediction that subject determiners and pronouns should be omitted relatively frequently, since they will not form a trochaic foot with a stressed syllable, as we saw earlier in (cf. 1a, 2a-b). Gerken (1991, 1994) provides experimental support for this hypothesis. Gerken 1991 showed that subject determiner and pronoun omissions ranged from 25-39%. In contrast, stressed syllables in the subject were omitted less frequently: 3-25% of the time. In (8) below, I give example sentences from her 1991 experiment; percentages indicate rates of omission (average age 2;3).

- (8) a. He kissed her.
35% 0%
- b. He kissed John.
33% 0%
- c. He kissed the dog.
39% 28% 0%
- d. The boy kissed her.
39% 3% 0%

- e. The boy kissed John.
 25% 25% 0%
- f. The boy kissed the dog.
 28% 11% 14% 0%

The syntactic account proposed in Hoekstra, Hyams & Becker (1997) makes a more specific prediction. Namely, at a stage at which children have a tendency to omit subjects and verb inflection, we expect to find a contingency between omissions of the subject or subject determiner, and finiteness in the verb. That is, when the verb is finite, the subject should be fully specified (an overt pronoun or determiner); if the verb is nonfinite, the subject may be specified or unspecified (either null or lacking a determiner). Concretely, we should see examples of (9a, b & c) in children's speech, but not (9d).

- (9) a. The boy dances
 b. \emptyset boy dance/ \emptyset dance
 c. The boy dance
 d. * \emptyset boy dances

As table 7 shows with data from Nina (from the CHILDES database), this prediction is borne out. While Nina's finite verbs almost always take a fully specified subject (i.e. one with an overt determiner), her nonfinite verbs are more evenly distributed with respect to occurrence with overt and null subject determiners.

Table 7: Nina's production of subject determiners and verb finiteness (2;4-2;10)

	Finite verbs	Nonfinite verbs
overt determiner	34 (92%)	12 (57%)
null determiner	3 (8%)	9 (43%)

$$\chi^2 = 7.85, p \leq .01$$

As we can see in Table 7, there is a strong correlation between overt determiners and finite verbs. Adam's data show a slightly different pattern, but one that is still consistent with the syntactic hypothesis. As shown in Table 8, 93% of Adam's finite verbs take an overt subject determiner, but almost none of his nonfinite verbs do.

Table 8. Adam's production of subject determiners and verb finiteness⁸ (2;3-3;7)

	Finite Verbs	Nonfinite Verbs
Overt determiner	53 (93%)	2 (5%)
Null determiner	4 (7%)	39 (95%)

$$\chi^2 = 71.64, p \leq 0.001$$

As for subject pronouns, 70% of Adam's utterances with a pronominal subject were finite (30% were nonfinite), and of Adam's nonfinite utterances, only 29% of them had an overt pronominal subject. (That is, Adam produced many utterances of the type *He is sleeping/he sleeps* and very few utterances of the type *He sleeping/he sleep*.) This fact about Adam's speech shows that the same nonrandom cooccurrence pattern between subject determiners and verb finiteness also holds

true for pronominal subjects. The metrical hypothesis does not predict such a correlation.

3.1 Subject-finiteness correlation from pilot data

The results of my pilot study also confirm the predictions of the syntactic account (i.e. there is a subject-verb correlation for finiteness). The data are given in Table 9.

Table 9: subject-verb finiteness correlation from pilot study (2;0-2;6)

	finite verb (pushed, pushes, hugged, hugs, is)	nonfinite verb (push, hug, null-is)
overt subject / subj det (Bert, the boy, he)	135 (85%)	20 (53%)
null subject / subj det (0, boy)	25 (15%)	18 (47%)

$$\chi^2 = 17.17, p \leq 0.0001$$

Within finite clauses, there is a strong tendency also to have an overt subject pronoun or determiner (85%). However when children failed to produce a finite verb, there was roughly an even number of productions containing as lacking a subject or subject det (53% and 47%, respectively). Note that this is the same pattern we found in Nina's speech (cf. Table 7), and is the pattern predicted by the syntactic account of Hoekstra, Hyams & Becker (1997).

4. Universality

A further potential advantage of the syntactic account is its universality. We find the same pattern of correlated finiteness between the subject and verb in other child languages such as German, Dutch and French. Some child data from German are given in Table 10.

Table 10: Subject determiners and finiteness in German children (avg. age 2;6)⁹:

	finite verb	nonfinite verb
overt subject det.	9 (90%)	2 (15%)
null subject det.	1 (10%)	11 (85%)

$$\chi^2 = 9.798, p \leq 0.005$$

Furthermore, evidence from child Italian (Schaeffer 1996) suggests that children inflect past participles to agree with the object if their object clitic is overt. When the object clitic is null, they do not produce agreement inflection on the participle. Whether the participle shows agreement or not, it has the same metric structure. The sentences in (10a-b) show the adult pattern of agreement. (11a) shows children's pattern with overt clitics, (11b) shows their pattern with null clitics, and (11c) shows that they do not inflect with null clitics. (The examples in (10-11) are schematic examples of the pattern with transitive verbs.)

- (10) a. le ho viste
 3pl.obj.cl. have-1sg. seen-3pl.
 "I have seen them."
 b. la ho vista
 3sg.fem.obj.cl. have-1sg. seen-3sg.fem.
 "I have seen her."
- (11) a. le ho viste (same as 10a)
 b. ho visto *pro* (meaning = (11a))
 have-1sg. seen-1sg.
 c. *ho viste/vista

Although these data do not run directly counter to Gerken's hypothesis, since she does not claim that all grammatical omissions are necessarily due to metric conditions (Gerken 1991), they show that there is a syntactic effect on object omissions that is not also a metric effect. That is, it is not the case that the omission of object clitics in child Italian is correlated with a metric effect on other words in the utterance, but rather with syntactic agreement inflection that does not bear on the metric structure of the sentence.

5. Data collection

Finally, let me point out some problems and advantages of using the two methods of data collection I used here. The main difficulty of using spontaneous speech, especially when looking for object omissions, is that in many cases it is difficult to tell whether an object was in fact omitted, i.e. whether it is obligatory. Particularly in the case of progressive verbs, many normally transitive verbs (e.g. *pushing*) can be used without an object in progressive aspect, taking a meaning like "push habitually". In this respect, using imitative speech is helpful since it is completely clear when a given morpheme is omitted.

On the other hand, it seems quite plausible to me that cadence and metrical structure play a much more important role in imitative than spontaneous speech. If this is true, then imitation tasks may well tell us about how metrical structure constrains a certain type of child speech, but not necessarily children's natural speech.

6. Conclusion

The main point of this paper was to explore the possibility that children's pattern of function morpheme omissions might be explained entirely by either a metrical or a syntactic story. There is strong support for both approaches, i.e. children preserve function morphemes that occur in trochaic feet, and they preserve determiners and pronouns in the subject and object when their verb is finite. Nevertheless, we have seen that there are cases like (5), which are explainable by a metrical story but not by a syntactic one, and the finiteness patterns in Tables 2 and 3, and 6-10, which are explainable by a syntactic story, but not by a metrical one. Thus, as the metrical and syntactic theories stand, neither one can fully cover the range of omission patterns we find in children's speech. In the future, we might find a way of refining and/or expanding one of the two approaches such that it captures the full range of phenomena, or we might discover a new approach that takes both metrical and syntactic phenomena into account.

Notes:

¹This view is supported in Gerken 1991, 1994a, 1994b, 1996, Wijnen, Krikhaar & den Os, 1994, Fikkert 1994, among others.

²Brackets indicate foot boundaries; feet are formed according to the principles outlined in Gerken (1996).

³There is no direct prediction for omissions in the object, given finiteness features in the verb.

⁴Finite verbs that took a syllabic inflectional morpheme (/es/) were excluded from this data set, since the metrical hypothesis predicts that object determiners/pronouns following them should omitted more frequently. There were not enough productions of verbs taking or requiring syllabic inflection to do a comparison strictly between them and verbs not requiring syllabic inflection.

⁵A progressive verb is considered finite if the auxiliary verb *is* is overt; it is considered nonfinite if the auxiliary is omitted. This analysis is supported by considerations discussed in Hoekstra & Hyams, forthcoming.

⁶Please note that while the object omissions after progressive verbs are all predicted by the metrical approach, regardless of finiteness, the correlation between verb finiteness in plain verbs and object omissions cannot be a result of an accidental overlap of metrical & syntactic predictions. In fact, the metrical hypothesis predicts the sequence [push the/him] to have a higher rate of preservation for the object determiner or pronoun than its finite counterpart; however as table 2 shows, the nonfinite verbs had a lower rate of preserved object determiners and pronouns.

⁷I am very thankful to John Grinstead (UCLA) for supplying me with the relevant utterances from the Nina corpus.

⁸Table 2 is taken from Hoekstra, Hyams & Becker (1997).

⁹Table 10 is reproduced from Hoekstra, Hyams & Becker (1997).

Bibliography

- Allen, George & Sarah Hawkins. 1980. Phonological rhythm: Definition and development. *Child Phonology*, vol. 1: Production. ed. by Yeni-Komshian, Kavanagh & Ferguson. New York: Academic Press. 227-256.
- Brown, Roger. 1973. *A first language*. Cambridge, MA: Harvard University Press.
- Clahsen, Harald, Sonja Eisenbeiss & Martina Penke. 1996. *Lexical Learning in Early Syntactic Development. Generative Perspectives on Language Acquisition*. ed. by Harald Clahsen. Philadelphia: John Benjamins.

- Fikkert, Paula. 1994. On the acquisition of prosodic structure. Ph.D. dissertation. Dordrecht: HIL.
- Gerken, LouAnn, Barbara Landau and Robert Remez. 1990. Function morphemes in young children's speech perception and production. *Developmental Psychology* 27. 204-216.
- Gerken, LouAnn. 1991. The metrical basis of children's subjectless sentences. *Journal of Memory and Language*, vol. 30. 431-451.
- Gerken, LouAnn. 1994a. A metrical template account of children's weak syllable omissions. *Journal of Child Language*, vol. 21. 565-584.
- Gerken, LouAnn. 1994b. Young children's representation of prosodic structure: Evidence from English-speakers' weak syllable omissions. *Journal of Memory and Language*, vol. 33. 19-38.
- Gerken, LouAnn. 1996. Prosodic Structure in Young Children's Language Production. *Language* 72. 683-712.
- Hoekstra, Teun & Nina Hyams. 1996. The syntax and interpretation of dropped categories in child language: a unified account. *Proceedings of WCCFL XIV*. Stanford: CSLI.
- Hoekstra, Teun & Nina Hyams (forthcoming) Agreement and the finiteness of V2: Evidence from child language. *Proceedings of the 22nd annual BUCLD*, Brookline: Cascadilla Press.
- Hoekstra, Teun, Nina Hyams & Misha Becker. 1997. The Underspecification of Number and the Licensing of Root Infinitives. *Proceedings of the 21st annual BUCLD*. Brookline: Cascadilla Press.
- MacWhinney, Brian & Catherine Snow. 1985. The Child Language Data Exchange System. *Journal of Child Language*, vol. 12. 271-296.
- Radford, Andrew. 1988. Small children's small clauses. *Transactions of the Philological Society*, vol. 86. 1-43.
- Schaeffer, Jeannette. 1996. Direct object scrambling in Dutch and Italian child language. UCLA Ph.D. dissertation.
- Suppes, Patrick. 1973. The semantics of children's languages. *American Psychologist*, vol. 88. 103-114.
- Wijnen, Frank, Evelien Krikhaar & Els den Os. 1994. The (non)realization of unstressed elements in children's utterances: A rhythmic constraint? *Journal of Child Language*, vol. 21. 59-83.

Person Reference and References to People in Brazilian Sign Language¹

Norine Berenz
University of the Witwatersrand

1. Introduction: person reference

The systematic examination of sign language forms functionally equivalent to spoken language personal pronouns got underway in the first decade after William Stokoe's article "Sign Language Structure" appeared in 1960. However, analyses proposed over the next two decades did not include the semantic level. (See Wilbur 1979 for a review of the early analyses. Recent analyses include Aarons et al 1992, 1994; Kegl 1990; Liddell 1994, 1995, 1996; Lillo-Martin & Klima 1990.) It was not until the late 1980s that the question was raised as to whether sign languages actually encode grammatical person.

Ahlgren (1990) challenged the linguistic universality of personal pronouns as proposed by Benveniste ([1966] 1971) and Lyons (1977), claiming that Swedish Sign Language has only demonstrative pronouns. Meier (1990) claimed that only the first/nonfirst distinction is encoded in American Sign Language, an analysis that has been widely accepted for ASL (e.g. Liddell 1994, Padden 1990). and extended to other sign languages — among these, Danish (Engberg-Pedersen 1993), Norwegian (Greftegreff 1995), and Argentine (Massone 1993). If these analyses hold, then sign languages are importantly different from spoken languages in not encoding both conversational participants: the sender and the recipient.

2. Sender only

Arguing for the grammaticization of the conversational role of sender in the form of a first person pronoun, Meier claims that points to center chest do not denote the signer as an individual but only refer to the signer in the role of sender, a meaning consistent with that proposed for spoken languages by Bühler ([1934] 1982) and others. Evidence for this claim is the fact that such points can be understood in particular contexts to indicate other individuals in the role of sender. This property of the ASL sign had been noted earlier by Kegl and others, but no one had examined it in the light of the semantics of first person deixis.

However, with respect to the grammaticization of the role of recipient, Meier concluded that there is no regular formational distinction between the forms used to indicate second and third person referential objects, and here he falls back on the Lillo-Martin and Klima (1990) "personless" analysis of a single pronoun with associated referential indices.

3. Conversational participants

Drawing on research on Brazilian Sign Language, carried out over several extended periods of fieldwork in which I video-recorded naturally-occurring conversations in a variety of settings, and my reanalysis of published ASL sources, I claim that at least ASL and Brazilian Sign Language encode the participant/nonparticipant distinction as well as the first/nonfirst distinction —

that is, they have grammaticized personal pronouns for both conversational roles: sender and recipient or addressee.

3.1 Signer's body as part of the phonological form

Although my analysis of first person pronouns differs from Meier's in important ways, in this paper I will concentrate on the issue of greatest disagreement: the grammaticization of the conversational role of recipient.

The first point is that analysis of the addressee form in different sign languages has been complicated by an assumption that a parallel relationship should exist between the sender's physical body as part of the description of the form of the first person pronoun and the addressee's physical body as part of the description of the form of the second person pronoun. A number of researchers (e.g. Engberg-Pedersen 1993) have noted that the signer can't point to an individual in the role of addressee to reference another individual in the role of addressee, as the signer can point to self to reference another individual in the role of sender. This observation has been put forth as evidence against a grammaticized second person pronoun.

The error here is to see the signer indicating self as sender in a kind of type-token relationship. Greftegref (1992) proposes an analysis of Norwegian Sign Language forms that reanalyzes what in the Stokoe paradigm is a particular place of articulation as, instead, a token indicated by a pointing hand. Among the examples she gives are NOSE, THINK, and I — points to the nose, forehead, and chest, respectively. An important feature of these articulations Greftegref fails to note is that the signer can refer to her nose, someone else's nose (even a nonhuman one) or noses in general only by pointing to her own nose; a point to anyone else's nose indicates that specific nose and not, I would add, by means of the nominal sign NOSE.

That is, it is only the signer's body which is the source of the phonological elements of the sign, not the bodies of others present in the conversational setting. Although Meier does not make this error in his analysis of the first person pronoun, relying instead on the configuration of observable elements of the articulation itself, he makes a related error as part of his argument against a second person pronoun in concluding that the ASL sign cannot be described "independent of the location of the real or hypothetical addressee" (1990:188). As much for the second person as for the first, it is the articulatory array presented by the signer in the role of sender — and that alone — which is relevant at the level of form; the physical location (and social identity) of the individual in the role of recipient only becomes relevant at the level of interpretation.

3.2 The role of gaze and the midline

What then is the articulatory array presented by the signer which is the form of the second person pronoun? I take as a starting point a passage from Appendix A of the *Dictionary of American Sign Language on Linguistic Principles*, which says:

First and second persons in signing are the opposite and interchangeable ends of an imaginary but well-defined line of sight. A third person or a fourth — even a fifth if needed — is designated by pointing at an angle to that line. (1976:281)

Here the use of the word person differs in an important way from the first mention to the second. The first use is clearly a reference to grammatical categories familiar from spoken languages but the second use shades into reference to individual entities. No discussion posits set-internal contrasts in form or meaning which would justify distinctions among third, fourth and fifth persons.

The DASL passage, however, reveals an intuition about the differing linguistic status of first and second person, on the one hand, and third, fourth and fifth person, on the other: first and second persons “are” but third, fourth, or fifth person “is designated by”. That is, in contrast to putative other “persons,” the first and second persons exist by virtue of the act of address itself.

We see in the word “interchangeable” a reflection of the shifting deictic center Benveniste, Fillmore, and Lyons have described, as the conversation moves between participants, an interchange in which the nonparticipant (that is, the nonfirst/nonsecond “person”) has no part to play. The nonparticipant is located outside the conversational interchange, at an angle to it. On the level of form, the line of sight and the midline² are crucial to a description of personal pronouns.

3.2.1 Examples

Considering and rejecting the possible elements in the form of a second person pronoun, Meier uses examples which are flawed in a number of ways: (a) the transcription is not sufficiently detailed to demonstrate the claim he is making and (b) the performance of the relevant sign is influenced by exigencies peculiar to the situation in which it is performed.

One of Meier’s examples contrasts two utterances, each of which has a conjoined pronominal subject. Four signs are glossed as INDEX, as if they are articulatorily equivalent, distinguished only by subscripts for participant roles and referential indices. In particular, gaze is transcribed as if it were of equal duration in all performances, yet I have noted that Brazilian signers typically mark third person reference with a much briefer gaze, a mere glancing. In actual practice, where more than a single potential addressee is present and attending to a conversation, it can be a delicate matter to make a third person reference to one of them — in part, I would suggest, because the sender’s glance is necessarily cursory.

LSB, example 1 (role of gaze):

Evidence from my data, recorded in a classroom in Rio de Janeiro, is a situation where a Brazilian Sign Language instructor apologizes to the student who will be the referential object of a demonstration of the third person pronoun and assures the student that he is “just playing around.” The form the instructor then produces shows a very brief glance to the conversational nonparticipant.

LSB, example 2 (role of the midline):

Moreover, the relationship of the hand and arm vis à vis the signer’s torso would differ in a predictable way between the performances of INDEX: to indicate the addressee, the hand and arm would be positioned near the midline of the signer’s body; to indicate the nonparticipant, the hand and arm would be at a

distinct angle to the midline. In another classroom in Rio where the students are arrayed in a wide semi-circle in front of the instructor, for a demonstration of the third person pronoun his hand and arm are positioned in almost the same plane as his torso, in contrast to the almost perpendicular positioning of the second person pronoun.

LSB, example 3 (interplay between gaze and midline in multi-party conversation):

A third example is typical of many of the conversations I observed during fieldwork. It is a three-party interaction. At first JC and NP are engaged in conversation about a Brazilian Sign Language class they've just finished co-teaching and about the next day's classes. GR, who is waiting for NP, stands off to the side several feet away. When the topic of the conversation becomes more general, GR waves his hand towards NP to get his attention, and both NP and JC turn towards GR. As they do, JC and NP step back from each other, making the positions of the (now) three interactants approximately equidistant.

In the exchange that follows, as each takes a turn, his head and gaze are directed towards one or the other of his two interlocutors so that head and gaze orientations shift between the two interlocutors' positions. That is, although both interlocutors fix gaze on the signer, the signer can only fix gaze on one person at a time. The chest orientation, however, is disjunct from the head and gaze orientations, remaining oriented at an angle about midway between the positions of the two interlocutors. At the end of the exchange, NP and JC turn towards each other, and the chest orientation is brought in line with head and gaze.

Fillmore (1975) proposes *audience* as a person deictic category, by which he means someone present in the setting not being addressed at a given moment in a conversation. Levinson (1983) proposes *bystander* as a person deictic category, by which he means someone not currently the addressee who is both present and actively attending to the conversation. Goffman (1976), Sacks, Schegloff, and Jefferson (1974), and Clark and Carlson (1982) examine conversational moves attributable to the presence in the setting of others not directly addressed. It is a question whether the particular positional elements I report here should be analyzed at the level of grammar or communicative practice or both. For the present, the important points are that gaze behavior is more complex than Meier's transcription and analysis reveal and that gaze plays a role at more than one level of the system. A closer look at the data may reveal the formational elements criterial to the distinction between second and third person pronouns.

Meier says of interactional settings such as these (i.e. involving signer, addressee, and audience) that "the set of pointing gestures we might identify as second person largely, if not completely, overlaps the set we would identify as third person." This does not hold if we are accounting for the orientation of handshape with respect to gaze and to the midline of the signer's body. I observed instances of second or third person reference wherein it was unclear who or what was the intended referential object — that is, reference did not go through — but no instances of a confusion between the two grammatical categories.

3.3 Corroboration from native signers

Signers' intuitions also support the claim that gaze and midline are crucial determiners of second and third person pronominal forms. I was fortunate to get on videotape two native deaf signers discussing pronouns and agreeing on the criterial role of these elements. In another recording session, a clearly exasperated sign language instructor chides a student by signing "I was not addressing you," and admonishes the student to recognize as relevant to the roles of sender and recipient in signed discourse a corridor of communication, which he indicates by two flat-hands, called B-hands, upright and parallel a few inches apart, moving outward from center chest. The second person pronoun is performed within the bounds of this corridor of communication.

3.4 Corroboration from citation forms and narrative

There are two final lines of evidence for the existence of patterned regularities sufficient to distinguish second and third person pronouns. First, signers unhesitatingly provide forms which show a conjunction of gaze, chest, and hand/arm orientation away along the midline of the torso in elicitations of a second person pronoun and equally consistently provide a disjunction of the chest and hand/arm orientation in elicitations of the third person pronoun. The citation form of the third person pronoun is especially interesting because it invariably orients to the signer's ipsilateral side (same side as active hand/arm), when a point to the contralateral side would seem to be articulatorily easier. The explanation may be found in the need to avoid the midline in order to keep the second and third person forms maximally distinct.

Second, for reported conversations using role shift, which rely on a typification of the conversational setting, the default position for a hypothetical addressee is directly opposite the signer in the role of reported sender. Pointing signs away along the midline of the body are interpreted as second person pronouns. This is the unmarked form. If a sign intended to have a second person rhetorical force differs from this, it will be understood to be reproducing the positions, relevant to each other, of the signer and addressee as they were at the time of the conversation being reported. Interpretation relies on context provided by the reported conversation. These marked forms, then, simultaneously encode second person and proximity relations between sender and recipient. Evidence for this claim is the fact that an interlocutor could challenge the position off the signer's midline but not the conjunction of formational elements with the midline — that is, the point away along the midline of the body neither asserts nor presupposes that the addressee was in fact so positioned. Contrary to what Meier and others have said, the form does not rely on the addressee's location to make second person reference.

4. From practice to grammar: references to people

I have argued here for the grammaticization of the conversational roles of sender and recipient in the form of first and second person pronouns. The course that process took is open to speculation. Other pro-forms have origins which are more clearly attributable to the exigencies of everyday conversational practice. Significant interactional factors which give rise to these forms are privacy needs and politeness considerations.

4.1 Privacy needs: shielded third person pronoun and the spatial anaphor

The need for privacy has given rise to a number of communicative practices, two of which I will mention here: the shielded third person pronoun and the spatial anaphor.

The *shielded third person pronoun* serves to hide the act of reference from nonsigners. It derives from a conversational practice functionally similar to whispering in spoken languages, whereby one hand hides the other hand while it signs. Two facts are evidence that the form is grammaticized. First, the tip of the index hand, called the G-hand, contacts the palm of the B-hand — other non-contact signs do not become contact — and the “shield” persists only for the duration of the single sign and does not persevere over a string of signs, as is commonly the case in signed “whispering”. Second, Brazilian signers readily volunteer the form; they do not offer shielded alternates of other signs.

As the form is well-known in the Brazilian Deaf community, the shielded third person pronoun cannot be used to ensure privacy when the possible “overhearers” are signers. In this case the *spatial anaphor* may be used. This form takes the spread hand, called the 5-hand, as the base on which to locate referents using the other hand in the G-handshape. I have been told that the form originated in discourse about soccer, where it represented the positions of players on the field. In Rio, the base hand is prone and contact is on the extended fingers; in Recife in the Northeast of the country, the base hand is supine and contact is on the palm at the juncture with the fingers. (In Porto Alegre, in the South, the form seems not to be used.) Although the position of the sender is relevant to the use of the form, there are no set-internal contrasts of sender, addressee, and nonparticipant; the form is third person by default, like demonstratives.

Both the shielded third person pronoun and the spatial anaphor are grammaticized forms whose origins in practice are still recoverable, as is the motive: to hide the visible.

4.2 Politeness considerations: formal personal pronouns and the classifier

The claim has been made (Ferreira Brito 1995) that pointing is not rude in the Brazilian Deaf community, but this is not true if the variety of alternate referential devices is any measure. Among these, the simplest of cases is the substitution of the B-hand for the G-hand. In my data, B-hand pronouns occur more frequently in situations marked by asymmetrical social status or unfamiliarity among interactants. On this basis, I categorize these as formal pronouns. (The interplay between politeness and the recipient’s attentional state, although relevant to pronoun choice, will not be discussed here.)

Another substitute motivated by politeness considerations takes the 4-hand as a base and the G-hand as the active hand. The 4-hand also serves in predications about long, upright objects. While this form can be used to indicate a collectivity of upright objects, including the nonhuman, in that use the signer cannot single out an individual within the collectivity. Since reference to human beings is the only condition under which a single object can be individuated within the collectivity, a notion of person is relevant to the use of the form, at least to the extent that the grammatical notion of person derives from the ordinary

sense of the word. Because the form is only used to make reference to a person at some remove from the conversational participants, it never happens that either sender or recipient is associated with one of the upright objects. There are no set-internal person contrasts.

In sum, Brazilian Sign Language data provide support for the claim that personal pronouns are universal in natural languages. The pronominal forms whose use rests on privacy needs and politeness considerations show that grammaticization can be an outcome of the exigencies of everyday conversation in the gestural-visual modality.

Notes

¹ The research on which this paper is based was supported by the Wenner-Gren Foundation for Anthropological Research (#5454), the National Science Foundation (#DBS-9214764), and Tinker Foundation Travel Grants. The writing of the paper was supported by a Wenner-Gren grant (#6151).

² Brentari (1998) makes reference to the midline in her analysis of the phonological process in ASL called "weak drop."

Bibliography

- Aarons, Debbie, Ben Bahan, Judy Kegl, and Carol Neidle. 1992. Tense and agreement in ASL. Poster presented at the Theoretical Issues in Sign Language Research Conference. San Diego, CA. July.
- , —, —, and —. 1994. Subjects and agreement in American Sign Language. In Ahlgren, Bergman, and Brennan, 13-28.
- Ahlgren, Inger. 1990. Deictic pronouns in Swedish and Swedish Sign Language. In Fischer and Siple, 167-174.
- , Brita Bergman, and Mary Brennan. (eds.) 1994. Perspectives on sign language structure: papers from the Fifth International Symposium on Sign Language Research. Durham, UK: ISLA.
- Benveniste, Emile. [1966] 1971. Problems in General Linguistics. Coral Gables, FL: University of Miami Press.
- Brentari, Diane. 1998. An optimality theoretic account of 'weak-drop' in two-handed signs in ASL. Paper presented at the annual meeting of the Linguistic Society of America, New York.
- Bühler, Karl. [1934] 1982. The deictic field of language and deictic words. Speech, Place, and Action ed. by R.J. Jarvella and W. Klein, 9-30. New York: John Wiley and Sons, Ltd.
- Clark, Herb H. and Thomas B. Carlson. 1982. Hearer and speech acts. *Language*, vol. 58, no. 2:332-373.
- Engberg-Pedersen, Elisabeth. 1993. Space in Danish Sign Language: the semantics and morphosyntax of the use of space in a visual language. Hamburg: Signum-Verl.
- Ferreira Brito, Lucinda. 1995. Por uma gramática de Língua de Sinais. Rio de Janeiro: Edições Tempo Brasileira.
- Fillmore, Charles J. 1975. Santa Cruz lectures. Deixis I & II. Ms.

- Susan D. Fischer and Patricia Siple. (eds.) 1990. *Theoretical Issues in Sign Language Research*, vol. 1. Chicago: University of Chicago Press.
- Goffman, Erving. 1976. Replies and responses. *Language in Society* 5:257-313.
- Greftegreff, Irene. 1995. SLLING (Sign Language Linguist List), November.
- . 1992. Orientation in indexical sign in Norwegian Sign Language. *Nordic Journal of Linguistics*, vol. 15, no. 2: 159-182.
- Kegl, Judy A. 1990. Predicate argument structure and verb class organization in the ASL lexicon. In Lucas, 149-175.
- Levinson, Stephen C. 1983. *Pragmatics*. New York: Cambridge University Press.
- Liddell, Scott. 1994. Tokens and surrogates. In Ahlgren, Bergman, and Brennan, 105-120.
- . 1995. Real, surrogate, and token space: grammatical consequences in ASL. *Language, Gesture, and Space* ed. by Karen Emmorey and Judy Reilly, 19-42. Chicago: University of Chicago Press.
- . 1996. Spatial representation in discourse: comparing spoken and signed languages. *Lingua* 98: 145-167.
- Lillo-Martin, Diane, and Edward S. Klima. 1990. Pointing out differences: ASL pronouns in syntactic theory. In Fischer and Siple, 191-210.
- Lucas, Ceil. 1990. *Sign language research: theoretical issues*. Washington, DC: Gallaudet University Press.
- Lyons, John. 1977. *Semantics*, vol. 2. Cambridge: Cambridge University Press.
- Massone, Maria Ignacia. 1993. *Lengua de Señas Argentina: diccionario bilingüe*. Buenos Aires: Editora Sopena Argentina S.A.
- Meier, Richard. 1990. Person deixis in American Sign Language. In Fischer and Siple, 175-190.
- Padden, Carol A. 1990. The relation between space and grammar in ASL verb morphology. In Lucas, 118-132.
- Sacks, Harvey, Emmanuel Schlegoff, and Gail Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50, no. 4: 696-735.
- Stokoe, William C. 1960. *Sign Language Structure*. Buffalo, NY: University of Buffalo.
- , Dorothy C. Casterline, and Carl G. Croneberg. 1965. Revised 1976. *A dictionary of American Sign Language on linguistic principles*. Silver Spring, MD: Linstok Press.
- Wilbur, Ronnie B. 1979. *American Sign Language and sign systems*. Baltimore, MD: University Park Press.

Language and Responsibility: Perceptions of NATO Expansion

Piotr Cap
University of California at Berkeley

1. Introduction

During its early-July 1997 summit in Madrid, the North Atlantic Alliance formally invited Poland, the Czech Republic, and Hungary to membership negotiations. Soon after that, on July 10, President Clinton, whose substantial support for NATO expansion had conclusively influenced the decision, went for a visit to Poland, in order to officially declare the beginning of a new era in the post-communist system of world security. The key event in the agenda was speeches to be delivered by Clinton and Aleksander Kwasniewski, the current Polish President, to the people of Warsaw gathered at the Castle Square. The occasion was perceived as a momentous occasion Poland, and as such had even turned the people's eyes away from the disastrous flood which was raging over southwestern regions of the country at the time.

The speeches, of which the first was given by Kwasniewski and second by Clinton, received a great deal of attention from the American media. A few public TV channels broadcast them live; also, most information and press agencies reported on them in considerable detail. In general, statements of both presidents were perceived as responsible and solid performances, displaying readiness to equally share both privileges and obligations resulting from the emergence of the new security structure. This predominant feeling of the apparently fair distribution of advantages and duties was quite representatively expressed in the USIA coverage of President Kwasniewski's words:

Warsaw - - Polish President Aleksander Kwasniewski welcomed President Clinton to Warsaw July 10, saying that for Poland, "an invitation to NATO is an accomplishment, but also a challenge. We are well aware of tasks ahead of us and their costs. Security does not come free. It is not offered as a gift. We stand ready to assume obligations resulting from NATO membership."

Noting that "everything going on these days has a symbolic dimension," Kwasniewski said "that the era we live in is symbolized not by walls dividing people but by bridges linking them. Not by hostility, but by cooperation. Not by a balance of fear but by common security."

Poland, he said, "will contribute to NATO its accomplishments of the recent years. We are a country of stable democracy. Independent courts, freedom of speech, and independent local government are now taken for granted in Poland. The civil control over the army testifies to its democratic standard. Our economic growth is one of the highest in Europe.... Polish reforms have passed a critical point. They are now irreversible."

Kwasniewski told the Polish people, "you are the greatest heroes of these transformations. It is to be remembered, however, how important for us was the support of our foreign friends, especially the USA."

To the Americans, he extended his thanks, saying that they are aware that "peace, security and prosperity of ourselves, our children and grandchildren" are priceless. (Washington File 97071009. GWE, USIS, USIA.)

Judging by length and a fair level of generalization, the main building blocks of the USIA report seem to be first paragraph and third paragraph, which respectively address issues of Poland's prospective benefits and contributions entailed by NATO expansion. At a closer look in fact, transition from the former to the latter frame is traceable as early as toward the end of the opening paragraph. All in all, the first paragraph and the third paragraph in combination appear to set up an aura of mutual cooperation, as well as a fully active membership on the part of Poland. The concluding paragraphs develop the idea of partnership and legitimized equality by bridging the people of the USA and Poland in their efforts dedicated to *peace, security and prosperity*.

Sadly enough, whether the USIA coverage was actually **intended** to fit the expectations of American taxpayers or not (this factor will be dealt with later in this study), it has nonetheless failed to convey what seems the predominant mood of President Kwasniewski's performance, that is, an implicit shift of anticipated responsibilities toward the American side. Furthermore, it ignored numerous markers of unwillingness to participate in the build-up of new security structures. Later we shall see that the USIA report on President Clinton's address turns out to be a comparably leveled misinterpretation ("disinterpretation"?). The coverage of Clinton's performance has been in turn a neutralization of a harsh imposition of obligations pertaining to NATO membership. The USIA attempts to position both speeches somewhere in between the non-impositive extreme and the forcefully impositive extreme, made in circumstances where one speaker (Kwasniewski) was clearly sticking to the former and the other speaker (Clinton) to the latter, seems either at least a linguistic and political blunder, or at most a purposeful act of manipulation. This paper has been devoted to a pragmatic analysis of both.

2. Kwasniewski: "Here we are, claiming a deserved seat in NATO's limousine..."

Let me sound unoriginal and provide a metaphor to illustrate and summarize President Kwasniewski's points, as it seems to capture the four all-important messages his words might contain. First, there is an aura of admiration of the American partner in the speech, the latter powerful but yet capable of ensuring comfort. Second, the address sees NATO as an infallible mechanism, which obviously can accommodate new members but does not necessarily depend on them for proper functioning. Third, President Kwasniewski seems to have either

inherited from his predecessors or simply learned from Polish history books the rhetoric of national uniqueness and natural feel of "belonging" to fate-dictating forces in the world, even though apparently he himself would rather dissociate membership as such from active participation (which is the moment when his Messianic heritage ends). Fourth, throughout the speech, the spirit of mere "being" clearly outweighs that of "doing" - as the idea of NATO membership is never to be found followed by an enactment of self-directed responsibilities and obligations.

Since this paper is not devoted to componential analyses of political metaphors (especially invented ones), I want to abandon the sub-title metaphor at this point, in hope that the reader will find no difficulty whatsoever in correlating automotive and political domains into the four hypotheses I am putting forward. An important thing to remember, however, is that just as the lexical items of the provided metaphor generate meaning in sequential structure dictated by rules of syntax, President Kwasniewski's implications work hand-in-hand toward achieving the superordinate goal, whose linguistic enactment comes via piling up ideological and pragmatic frames, rather than via free distribution of pre-set messages.

2. 1. America: powerful and protective

President Kwasniewski's address opens with a warm welcome:

(1) We welcome you with pleasure to Polish soil, Mr. President. We welcome the leader of the superpower that shapes global policy. We welcome the man who is doing so much for world order, security and peace. We welcome a close friend of Poland.

Apparently a pragmatically neutral celebrity, the welcome seems to carry some important illocutionary load. First of all, by applying a considerable dose of definiteness/uniqueness when characterizing the guest, Kwasniewski positions the US side clearly superior to the rest of the world. The US's past and present actions are being applauded and expected to continue (note the use of the progressive aspect in *doing*). Their outcomes are perceived as naturally beneficial to the world and as such evade questioning (presupposition *so much*). In general, strategies of positive politeness (cf. Leech 1983; Brown and Levinson 1987) adopted by the Polish President in the opening paragraph of the speech seek to maximize approbation of the US partner, in order to encourage further activity on his side. Left to Poland is in turn the spectator's role.

The admiration frame coupled with the assumption of passiveness on the part of Poland is to be found later in the address as well, especially when President Kwasniewski talks about American promises:

(2) On 22 October 1996, you said in Detroit, Mr. President, that the first group of prospective NATO members would be invited to preliminary talks in the early

summer of 1997. And you have been as good as your word. Thank you, Mr. President!

That the passive voice is used in the first of the two sentences could perhaps be attributed to a rhetorical routine, if not for another implication of passiveness in the other sentence. Surprisingly enough, by engaging into personal characterization of the partner (*as good as your word* - let me drop phraseological considerations here), Kwasniewski in fact further builds up an aura of superiority around the American side. This time however, its fate-dictating powerfulness is presented not so much in terms of ruling capacity, but rather in terms of protectiveness. It is the image of the proverbial Uncle Sam that appears to be invoked in the passage, the vision of a country which not only "knows better" and, consequently, has its say in the matter on grounds of internal strength as such, but also **kindly** uses its potential for the sake of others.

An important remark to be made is that President Kwasniewski's acknowledgment of American superiority, however rhetorically inconvenient in front of the home crowd at the Castle Square, is going to be balanced by oncoming enactments of historically marked presence of Poland in the international arena. This, I believe, proves coherence of presentation of the four pragmatic frames within the speech (cf. Van Dijk 1977). Still, when expounding on the past, Kwasniewski will in turn be reluctant enough not to address issues ranging back in time beyond the Declaration of Independence. Doing so, not only might he undermine his partner's moral right to rule and protect, but also damage the aura of any partnership at all.

2. 2. Fitting in

There are two major discrepancies between the USIA coverage of President Kwasniewski's address and what seems to constitute his actual message. One of them relates to Poland's perception of the Alliance as such, and the other to Poland's perception of how the Alliance might possibly benefit from the accommodation of new members. Since I consider it convenient for the reader to follow the analysis of the speaker's words in accordance with chronology of the address, the latter question will be dealt with in (2.4.). Right now I wish to cast some light on the manner President Kwasniewski handles his outline of world security under NATO command. The following examples are found approximately one-third of the way through the entire speech:

(3) Everything going on these days has a historic dimension. The end of the Yalta order is approaching. The area of stability and security is expanding. The peaceful order is being consolidated. Poland has not been in a better situation for centuries.

(4) The era that we live in is symbolized not by walls dividing people but by bridges linking them. Not by hostility, but by cooperation. Not by a balance of

fear but by common security. America won the Cold War so that the fruit of this victory could serve free nations. America has shown the world that to be a superpower means to participate, to care about, and build the future.

(5) *It [Poland's membership in NATO] is a natural and logical choice...*

Perhaps the most striking pragmatic quality of these passages is that they are densely packed with direct assertions, which clearly outweigh other illocutionary acts. The pragmatic and rhetorical power of assertion has been thoroughly investigated in recent years. What has been found is that although the primary function of assertive acts is to assert a state of affairs only, many such acts trigger additional inferences on the part of the addressee. In somewhat more precise terms, assertions seem to be capable of enacting credibility, on which the speaker can formulate or impose other acts, not necessarily of an assertive nature alone. In actual fact, it is the very domain of political language that appears particularly characteristic of a number of **directives** clustered around respective assertions (cf. Wilson 1990; Cap 1997). And so, that assertions can be forceful is something to bear in mind when analyzing President Kwasniewski's speech.

How does he acquire his credibility, then? Simply by reiterating messages which stay in line with the pre-existing beliefs most of his audience might share (cf. Festinger 1957). *The end of the Yalta order* is a fact realized, amongst other political events, by the crumbling of the Soviet empire. President Kwasniewski's mention of the *American victory in the Cold War* is another reiteration performed on a similar basis. That *the fruit of this victory* is going to *serve free nations* is in turn an enactment of anti-isolationism, an idea which, having originated with Franklin Delano Roosevelt, has been imposed upon the American minds ever since - to the effect that it gradually became one of the major face-creating assets which have been battled for in presidential campaigns since World War II (cf. Windt 1994). It has therefore been relatively safe for Kwasniewski to tackle such issues, in hope that his audience would find him credible enough to accept later points, such as the claim for Poland's membership in NATO on grounds of historical dogmata and natural order in (5). Interestingly, this claim would most probably have been less appealing if not for the apparent deagentization of assertions in (3) and (4). There are virtually no agentive forces being assigned any "responsibility" for *the end to the Yalta order, stability, security, peaceful order, bridges, or cooperation*, even if we all know, recognize and perhaps appreciate NATO's role in shaping both post-war and post-1989 security structures. However obvious the latter might be, and further given that President Kwasniewski in (4) finally acknowledges American contribution (probably for the reasons explained in 2. 1.), his earlier, deagentized claims do not cease to work toward the enactment of naturalness of historical change which **itself** ultimately allowed for Poland's integration into the Alliance. There seem to be no better circumstances for assuming passiveness, which I believe President Kwasniewski

partly does toward the end of (3), saying *Poland HAS NOT BEEN in a better SITUATION for centuries* [capitalization mine]. Hand-in-hand with this assumption comes an implicit **imposition** of not only an inherent right to join NATO, but also a **directive** at "whoever" is going to be responsible for world security now to further deal with what "naturally" emerged as a self-construed reality (after all, powerful and kind-hearted leaders need neither be replaced nor substantially assisted, do they?)

2. 3. In search of compensation...

To say that the speech of President Kwasniewski pays merely an adequate tribute to past accomplishments of the Polish people over centuries would certainly be an understatement, considering that approximately 50% of verbal forms used in the address take a past ending. About half-way through the speech, Kwasniewski gives a chronological account of four distinct periods in Polish history. Let me be particularly considerate to the reader here and recognize the limitations of space, choose single sentences for exemplification:

(6) *It was here that the first European constitution was passed in 1791.*

(7) *Later, in 1939, there came a moment when we were attacked from the west and the east at the same time.*

(8) *Poland and the whole of Central Europe - through no fault of their own - were left excluded from all post-war reconstruction projects, integration, and cooperation in the field of security.*

(9) *"Solidarity", the protests of the 1980s, the "round-table talks" marked the beginning of a great transformation.*

If you look at the thematic structuring of examples (6)-(9) (note that selection follows the chronology of President Kwasniewski's speech), you will probably find that sentences (6) and (9) somehow tend to "embrace" the remaining two. This is due to the fact that, in contrast to (7) and (8), they both deal with remarkable achievements, which indeed have been a contribution to the European civilization. On the contrary, sentences (7) and (8) set up an aura of isolation, during which period no such contribution was made. Yet, for the reason that sentences (7) and (8) are framed by (6) and (9), they might be perceived as thematically-marked and therefore non-representative in context (cf. Halliday and Hasan 1976). What follows is that the addressee will consequently consider the inside of the text in terms of an exception to its complex thematic frame. In actuality then, President Kwasniewski enacts an image of continuity of Polish presence on the international arena, only temporarily disrupted by external factors. An emerging implicatum is that, since Polish membership in the international community is a stability-related (stability-ensuring, in fact) norm and a drive

toward stability has been only self-evident in recent years, the *peaceful order* in Europe ought to be restored to such standards that would automatically recognize Poland's deserved place in the community of ancient democracies.

The implications do not stop here, though. Actually, it appears that sentence (8) thematically mirrors a great deal of the present-day discourse on NATO expansion. Let us imagine a fragment of a cover story saying, (10) *Countries of Central Europe have been integrated into the newly emerging security structure*, which as such is about to become an instance of reality, rather than mere hypothesis. Now, let us compare it to how the actual sentence (8) is structured. What we find is that there holds an analogy in the subject reference, another analogy in the verbal voice, and still another in adverbial phrase reference. That sentences (8) and (10) lack agents is a follow-up analogy being formally true, yet misleading in terms of the generated implicatures. There is good reason to believe that, for historical and political reasons, one might consider the American side an underlying co-agent in the case of the former sentence and a full agent in the case of the latter sentence. Consequently, if we assume that the USA used to have grounds to compromise with the USSR on the seclusion of the Central Europe in exchange for the maintenance of the wobbly balance of power back in 1945, we are naturally prone to conclude that nowadays, with the Soviet threat most probably gone, there is nothing that might prevent Uncle Sam from fostering European integration. Mindful of the size of tragedy and humiliation that followed the 1945 compromise, the addressee is likely to have such implicature turn into an obligation - forceful enough to remove any NATO-related responsibilities from the shoulders of those who apparently deserve compensation for years of suffering.

2. 4. Fitting in (continued)

It is quite remarkable that, as long as they are interpreted in isolation from the rest of the speech, the first paragraph and the third paragraph of the USIA report do provide a reasonably balanced account of costs and benefits relative to the expansion of NATO - the reason being the fairly general character of the third paragraph. Although it seems natural for political audiences to expect the idea of "what we are" to be followed up by "how we (will) act" (cf. Cap 1997), such topical and pragmatic shift is virtually absent from President Kwasniewski's address, certainly not for the limitations of space. In fact, before the speech ends there is enough of it to acknowledge a skillful use of the defeasibility of an action implicature, which is inferable from the USIA quotations if and only if the whole speech has been analyzed. The key problem is that, although the enumeration of *accomplishments* indeed cancels the assumption of *obligations*, no earlier than upon a complete analysis is this cancellation traceable. With only a portion of the actual text provided, the reader is likely to perceive the third paragraph of the USIA coverage in terms of an introductory generalization of what he believes

must have followed it. On the other hand, when approaching the end of the full-size text, he will consider the implicature canceled back at the very first moment it was followed by an enactment of **presence** instead of **action**.

2. 5. Summary

As it has been noted earlier in the paper, the assumption of passiveness on the Polish side is an underlying implicatum behind all four of the analyzed frames. Negotiating a safe "passenger seat" in the "NATO limousine", President Kwasniewski settles on his major rhetorical asset, a continual adherence to a multidimensional status quo. On the one hand, he would be pressing for membership on account of Polish heritage and an established contribution to the making and cultural growth of the Old Continent. Doing so, he would remind his audience of the first constitution ever passed in Europe, bring up images of World War II, or even presuppose Poland's place in NATO ((11) *Not for a moment will Poland in NATO forget the aspirations and hopes of Romania, Slovenia, ...* , he eventually says). One might perhaps consider this strategy negotiation of a civilizational status quo. On the other hand, however, President Kwasniewski would encourage the US side to continue the mission which has been enjoying an almost worldwide applause over the past five decades, and there is hardly any more spectacular way to cling to a status quo in security leadership than with the welcome opening of the speech. Nevertheless, it should be remembered that neither of these strategies is capable of enacting political engagement, one being merely a glorious recollection of the latter and the other surrendering it up to the leader.

3. Clinton: "Do your part, unless..."

President Clinton was to speak second at the Castle Square. Following is the USIA coverage of his address:

Warsaw -- "Three years ago this week," President Clinton told the citizens of Warsaw July 10, "I came to this great city and made this pledge: Nothing about you without you. Nic o was bez was."

"Now Poland is joining NATO," he said. "Poland is taking its place in the community of democracies. Never again will your fate be decided by others. Never again will the birthright of freedom be denied you. Poland is coming home."

In remarks to Polish citizens gathered in Castle Square in downtown Warsaw, Clinton had this message for the American people and their representatives on Capitol Hill: "To the citizens of my own country I say, this land where I speak has known the worst wars of the 20th century. By expanding NATO we will help to prevent another war involving Poland, another war in Europe, another war that also claims the lives of Americans."

One week ago was the 4th of July, America's Independence Day, the President said. "More than 200 years ago, you sent your sons to help to secure our future. America has never forgotten. Now, together we will work to secure the future of an undivided Europe for your freedom and ours."

"That is the promise that brings us together today," Clinton said. "That is the promise that will keep us together in a new Europe for a new century. That is our promise to all the young people here today and to generations yet to come: Security, for 100 years. Sto lat. Democracy for 100 years. Freedom for 100 years." (Washington File 97071006. USIS, USIA.)

When compared to the account of President Kwasniewski's speech, the USIA report on President Clinton's address reveals a striking proportion of similarities. In fact, both coverages hardly differ in structure; nor do they elucidate any different thematic/pragmatic content of the speeches. First, both reports bring up an apparently equal share of mutual costs and benefits following NATO expansion, their structural parallel being the first and, predominantly, the third paragraph of Kwasniewski's coverage vis a vis exclusively the third paragraph of Clinton's coverage. Second, both elicit an image of a fair partnership built on historical grounds. The structural correspondence holds even stronger here: the layout of the reports has the fourth paragraphs follow virtually an identical thematic line, except that Clinton tactfully reverts Kwasniewski's tribute to the American support. Third, among the most important messages of both speeches will be what the third paragraphs of the respective coverages deal with, the parallel appearing in the matter of function, rather than content. We have namely seen thus far how misleading the third paragraph of the USIA report on President Kwasniewski's address might be, especially when it comes to the issue of political involvement. Not to jump ahead of the actual analysis, let me only suggest now that the interpretation of the third paragraph of the USIA report on President Clinton's speech will again be an exercise in spotting intricacies of powerful implicatures, as well as sociopsychological underliers of political talk. To close this outline of analogies I should repeat what I said in the introduction: both coverages fail to account properly for the pragmatic load of the speeches as such.

3. 1. "..., unless you surrender to a foreign rule"?!

There is not a single word in the USIA coverage of President Clinton's speech that would either state or imply Poland's responsibilities entailed by the expansion of NATO, contrary to tens of such markers in the whole address. There is not a single suggestion of actual costs following the membership, one that would do justice to numerous implications of a new share in financing the new security structure in Europe. Finally, there is nothing in the report that would account for what might possibly happen should Poland prove unable to meet NATO standards. Instead (!), the USIA report offers a lengthy reassurance right in the

third paragraph, by itself longest in the coverage. Let us then look at a **necessarily** longer fragment of President Clinton's speech at the Castle Square:

(12) *NATO is doing its part - taking in new members, taking on new missions, working with new partners. (...)*

(13) *Now, as your President has said, you must continue to do your part. Poland, Hungary, and the Czech Republic will now become full members of our Alliance, with the full responsibilities of membership - the responsibility to nurture and strengthen and defend your democracies, because, as we in America know, after more than 200 years the struggle for democracy is never over, it must be fought every day.*

(14) *The responsibility to continue the remarkable transformation of your economies, because, having known poverty, you know the true value of the prosperity you have only begun to achieve. The responsibility to reach out to all your neighbors, to the East as well as the West, including the people of Russia. You must continue to build in tolerance what others destroyed in hate.*

(15) *The responsibility to meet NATO's high military standards and to help to bear its cost, because true security requires strength and readiness. We know you are ready to share the burdens of defending freedom because you know the price of losing freedom.*

In actuality President Clinton's address, insofar as it deals with practical implications of Poland's membership in NATO, demonstrates a pragmatic continuum. It sets up a context for a directive, performs the directive both on- and off-record (cf. Brown and Levinson 1987; Thomas 1995), supplements an entailment, and ultimately backs up the whole sequence with an implicit warning.

Namely, in (12) Clinton asserts a substantial contribution on the part of the Alliance, whereby he implicates an absence of such contribution from the prospective members. This built an implicature, self-defeasible by definition, undergoes a half-cancellation by the *continue* part in (13), only to have provided the addressee with sufficient background to take what follows in (13) and (14) as a matter of course. Consequently, the indirect orders that characterize paragraph (13) and paragraph (14) are practically bound to fall within the addressee's latitude of acceptance, rather than rejection. Furthermore, the internalization of obligations imposed by the *you must* repetitions is another factor that pre-alleviates an attitudinal dissonance, which the audience is unlikely to experience anyhow, having been exposed to an indisputable assertion beforehand.

The directives issued throughout (13) to (15) seem to fall into two categories. On the one hand these are indirect commands of a relatively small degree of implicitness, as in *you must continue* to or *the responsibility to* parts. President Clinton performs them baldly on-record, as they have already been neutralized by the implicature of passiveness in (12). On the other hand, paragraphs (14) and (15)

contain implicit warnings; respectively, of the return to *poverty*, and the restoration of Poland's political dependence (*losing freedom*). The warnings will, however, slide off-record in Clinton's speech, since he has made no attempt to pre-neutralize them the way his orders have been pre-neutralized. And so, President Clinton's warnings in (14) and (15) reveal a perplexing mixture of indirectness, hinting via lexical juxtapositions ("stick and carrot" tactics), incompleteness of nominalizations (the agent canceled, it is left unclear who might re-attempt a political takeover of Poland), and semantically neutralizing metaphors (note the use of the Clausewitz's concept of political *price*, *gains*, and *losses*).

Somewhat analogously to the presence of a dependency link between the assertion in (12) and all the subsequent directives which I have just pointed to, President Clinton's performance of the latter ones entails **more** directives. Positioned as an integral part of the overall package of obligations resulting from the prospective membership is namely a responsibility to assume a regular share in financing the new NATO. A potentially face-threatening act, Clinton's imposition of such responsibility over Poland comes once again within a neutralizing frame. The principal components of this frame are the assertion following the directive (*because true security...*), the verb preceding it (*help*), and in fact paragraphs (12) to (14), which have been building up a truly ideological setting for the consideration of the down-to-earth implications of NATO expansion.

Having expounded on the US perception of the new Alliance, President Clinton finally visualizes a situation which might be the result of either a blatant rejection of his ideas by the countries of Central Europe, or, more probably, their inability to meet the enforced standards. The implicit warning in (15) seems quite reminiscent of the "tiger metaphor" coined by President Kennedy in his 1961 inaugural (cf. Windt 1994), yet it is by far better constructed and placed in the speech. Not only does it provide a forceful closing to the most important section of the address, but it does so with a presupposition of partnership and mutual understanding, one that adds to the common ground prescribed by the occasion.

4. Conclusion

Imagine the USIA gives a truly comprehensive account of Clinton's points. Imagine the reader discerns more of strenuous persuasion than of courtship and reassurance. There is obviously no room for persuasion and argument in a political address unless there is an underlying problem to overcome. There is no need for a linear buildup of argument unless resistance is presupposed, either. Thus, the expansion of NATO posing a problem, is it worthwhile to support an increasingly bigger spendings it incurs? For an average American taxpayer, is the vague benefit going to pay for the cost?

Imagine the USIA takes an effort to interpret Kwasniewski's words or simply provides longer quotations. Is it worthwhile to invest in a country, however strategically located, whose head would talk of the past rather than future, in an address that has been expected to delineate actual policies?

As we have seen, both coverages were far from inciting such questions. A government-sponsored nonprofit agency, the USIA would rather reassure the addressee his money buys American security. That it same buys loyalty of the media he will probably never know.

References

- Brown, P. & S. Levinson (1987) *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.
- Cap, P. (1997) "On the Pragmatic Organization of JFK's Inaugural Speech" in B. Lewandowska-Tomaszczyk ed. *New Trends in Language Studies*. Lodz: Wydawnictwo Uniwersytetu Lodzkiego.
- Festinger, L. (1957) *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press.
- Halliday, M. A. K. & R. Hasan (1976) *Cohesion in English*. London: Longman.
- Lakoff, G. (1990) *Metaphor and War*. Distributed via WWW.
- Leech, G. (1983) *Principles of Pragmatics*. London: Longman.
- Levinson, S. (1983) *Pragmatics*. Cambridge: Cambridge University Press.
- Thomas, J. (1995) *Meaning in Interaction: an Introduction to Pragmatics*. London: Longman.
- Van Dijk, T. A. (1977) *Text and Context*. London: Longman.
- Wilson, J. (1990) *Politically Speaking: the Pragmatic Analysis of Political Language*. Oxford: Basil Blackwell.
- Windt, T. (1994) *Presidential Rhetoric*. Dubuque, IO: Kendall & Hunt.

The Tough Construction is neither Extraction nor Control

Yoon-Suk Chung
UC Berkeley

1. INTRODUCTION. The English *tough* construction (TC), illustrated in (1), is a structure in which one of a small class of adjectives like *tough*, *easy*, etc. (= *tough* predicates) has a subject which is understood as the so-called missing argument of an infinitival complement.

- (1) a. John is easy to please.
- b. Mary would be difficult for anyone to talk to.

The TC has been identified in generative grammar either as a species of extraction, or as a species of control. In this paper, I will argue that the TC is neither extraction nor control, but a third type of construction that shares one salient grammatical property with these other two constructions.

First, I will review two accounts of the TC, the extraction account and the control account, and point out some problems with these accounts. I will then suggest a new construction grammar-based account of the TC.

2. TWO ACCOUNTS. The essential problem presented by the TC is how to account for the missing argument embedded under a *tough* predicate, the so-called *tough* gap.¹

For the past four decades, this question has been answered in terms of already existing analyses of English syntactic phenomena. That is, the TC has sometimes been treated as a kind of extraction construction which otherwise include things like *wh*-questions, and topicalization constructions, etc. as illustrated in (2).

- (2) a. Who do you think John admired?
- b. Mary, I think John admired.

Alternatively, it has been treated as a kind of control construction, which otherwise includes *equi* and raising phenomena, as in (3).

- (3) a. John tried to leave.
- b. John seems to have left.

2.1. EXTRACTION ACCOUNT. The Extraction account basically says that the *tough* gap is created when an argument is extracted or displaced from an otherwise realized embedded argument position.

The extraction account started with Chomsky (1977), and has become the dominant view. For example, in GB, Chomsky (1986) suggests a D-structure (4a) for sentences like (1a):

- (4) a. John is easy [_{CP} [_{IP} PRO to please OP]
- b. John_i is easy [_{CP} OP_i [_{IP} PRO to please t_i]

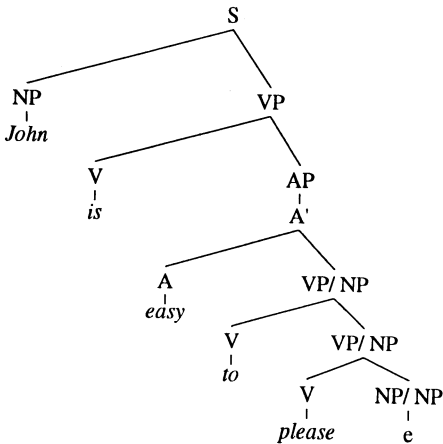
In (4b), a trace is created in the missing argument position as a result of movement of a null operator (OP) to the clause-initial, non-argument position of the embedded

infinitival phrase. This operator is later coindexed with the matrix subject of the TC (= the *tough* subject).

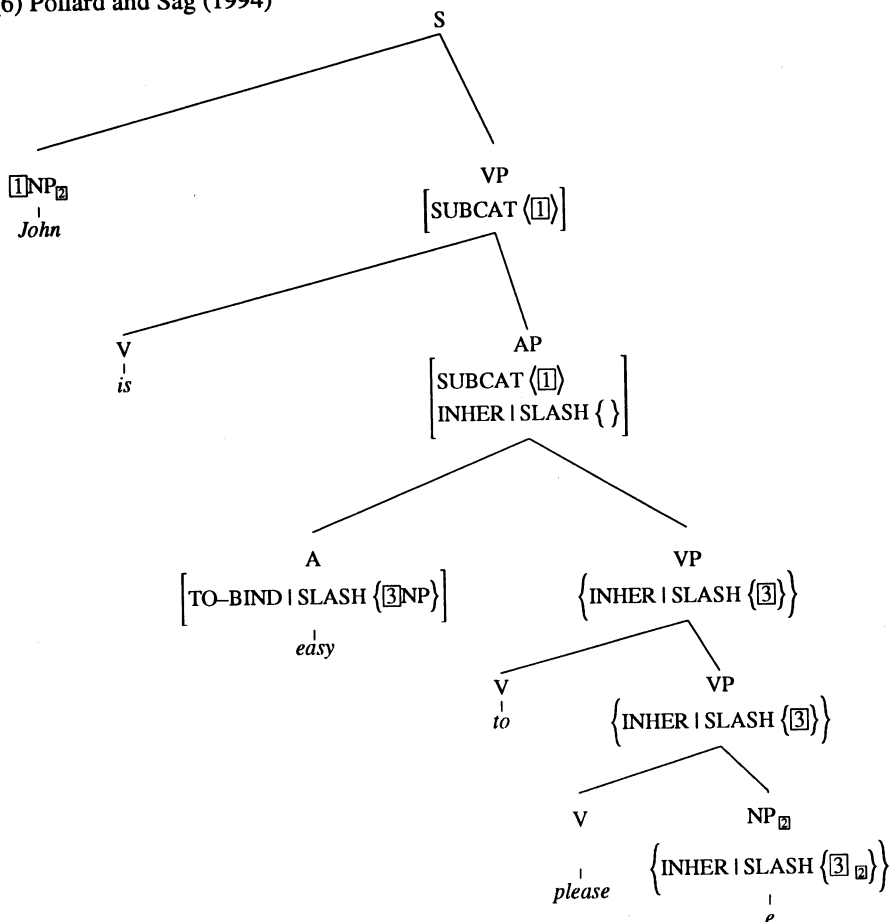
Basically, the same idea has been implemented in monostratal frameworks. In standard GPSG or HPSG such as Gazdar et al. (1985), and Pollard and Sag (1994), the *tough* gap arises as the value of the gap feature SLASH propagates from the gap position via extraction path to the position where it is cashed out, or receives its value.

The trees shown in (5) and (6) should make this clear. In (5) of the Gazdar et al. (1985) treatment, the SLASH value is cashed out by the *tough* subject at the node of a matrix sentence. In (6) from Pollard and Sag (1994), the SLASH value is cashed out by the *tough* predicate at the node of an AP. The lexical entry *easy* ensures the coindexation of the trace with the *tough* subject in the lexicon.

(5) Gazdar et al. (1985)



(6) Pollard and Sag (1994)



2.2. ARGUMENTS FOR THE EXTRACTION ACCOUNT. There are three main arguments for the Extraction account. First, like extraction, seen in (7a), the TC may involve long distance dependencies. See (7b).

- (7) a. Which gadget did you persuade people to buy?
b. That gadget would be difficult to persuade people to buy.

In (7), the dependencies between the missing argument positions and their antecedents occur across more than one clause boundary.²

Second, like extraction, the TC is subject to some island constraints.³ Consider the examples in (8)-(9).

- (8) a. *Who do you imagine the likelihood of entertaining?⁴
 b. *John would be difficult to imagine the likelihood of entertaining.
- (9) a. *Who do you imagine a person who entertains?
 b. *John would be difficult to imagine a person who entertains.

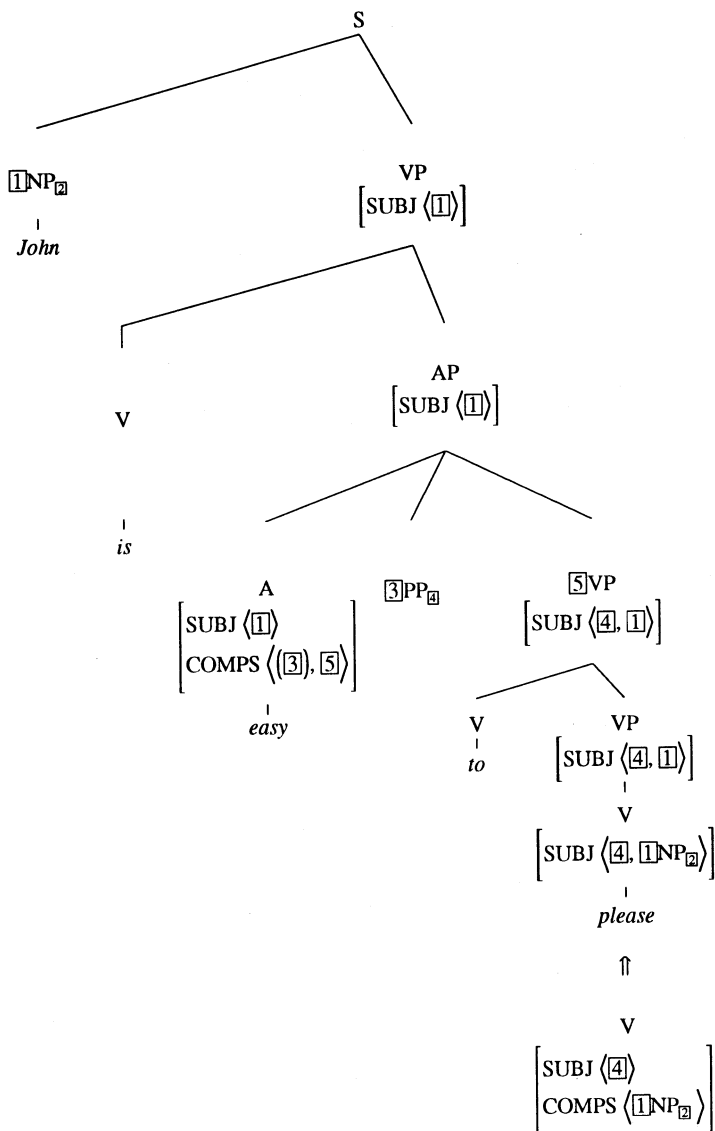
(8b), like (8a), is ungrammatical since it violates the Complex NP Constraint. The same is true with (9).

Finally, like extraction in (10a), the TC allows parasitic gaps, like the one in (10b).

- (10) a. Which paper do you think you filed without reading?
 b. This paper is easy to file without reading.

2.3. CONTROL ACCOUNT. The Control account, on the other hand, says that the *tough* gap is a 'controlled' argument in the way that the understood subject in an equi or raising construction is controlled by an argument of a higher predicate. Control-type accounts include the *tough* deletion account, which is essentially the same as the Equi NP deletion account, and a movement analysis which parallels the traditional treatment of raising. Control accounts also include some less familiar analyses in monostratal frameworks like Grover (1995). (11) is a representation for a structure (1a) on this account.

(11) Grover (1995)



(12) Grover's lexical rule

$$\left[\text{SYNSEM} \mid \text{LOC} \left[\text{CAT} \left[\begin{array}{l} \text{SUBJ} \langle \dots \rangle \\ \text{COMPS} \langle \dots, \boxed{\text{NP}}, \dots \rangle \end{array} \right] \right] \right] \\ \Rightarrow \left[\text{SYNSEM} \mid \text{LOC} \left[\text{CAT} \left[\begin{array}{l} \text{SUBJ} \langle \dots, \boxed{\phantom{\text{NP}}} \rangle \\ \text{COMPS} \langle \dots \rangle \end{array} \right] \right] \right]$$

For Grover's HPSG analysis, the *tough* gap is created when the missing argument is promoted, by a lexical rule (12), from the COMPS list to the SUBJ list of the embedded verb's subcategorization structure, seen at the bottom of the diagram (11). The promoted subject argument, $\boxed{\text{NP}}$ in (11), is controlled by the *tough* subject in accordance with general principles that govern the relationship between the subject of the controlled VP and its controlling subject.

2.4. ARGUMENTS FOR THE CONTROL ACCOUNT. There are two arguments for the Control account. First, the 'controller' of the *tough* gap occurs in an argument position, i.e. specifier of IP, not a non-argument position, i.e. specifier of CP. So, if your theory forces you to choose between control and extraction accounts, this fact makes the TC look more like a control phenomenon.

Second, the TC is only licensed by specific lexical predicates such as *tough*, *easy*, etc., just as control constructions are only licensed by certain predicates.

- (13) a. This door is impossible/*possible to open.
b. John tried/*screamed to open the door. (*agentive reading)

Just as a predicate like *screamed* in (13b) does not participate in the equi construction, a predicate like *possible* in (13a) may not participate in the TC.⁵

2.5. PROBLEMS WITH THE EXTRACTION AND THE CONTROL ACCOUNTS. Proponents of the extraction account emphasize the extraction properties, and de-emphasize the control properties, dismissing them as irrelevant or trivial. Proponents of the control account, on the other hand, play down the extraction properties, and emphasize the control properties of the TC. Thus, each account is unable to account for some portion of the data.

Furthermore, there are other properties of the TC which one or the other or both accounts cannot explain. First, unlike extraction, the TC is not always an island for extraction. Multiple gaps are sometimes permitted in the TC, like (14a). This property is problematic for an extraction account since canonical extraction constructions *are* islands to extraction, as seen in (14b).

- (14) a. Which violin is this sonata easy to play _ on _?⁶
b. * Which violin did you ask which sonata Mary played _ on _?

Second, (15) shows that the TC does not allow nominative gaps.

- (15) a. *John is difficult for me to believe _ went to Chicago.
b. *Mary is tough to believe _ likes John.

This property is problematic for an extraction account since, as seen in (16), embedded subject gaps *are* possible with canonical extraction constructions.

- (16) a. Who do you believe _ likes John?
b. Mary, I think _ likes John.

Third, as seen in (17a), the case marking in the TC is different between the gap and its antecedent: the antecedent is nominative, but the gap is in the accusative position. This property is problematic for an extraction account since in extraction constructions the case assigned in the missing argument position is carried over to the antecedent as in (17b).

- (17) a. She _{NOM} is easy for me to please _ _{ACC}
b. Whom _{ACC} do you think you pleased _ _{ACC}?

Fourth, unlike control, the 'controllee' of the TC must be a non-subject of the non-finite verb phrase, as shown by (15), repeated here as (18).

- (18) a. *John is difficult for me to believe _ went to Chicago.
b. *Mary is tough to believe likes _ John.

This is in contrast with control where the controllee is typically a subject.

Finally, there are problems for both the extraction and the control accounts. Depending on the speaker, a *tough* gap may or may not be associated with the *tough* subject across a finite clause boundary. For example, while Hukari and Levine (1991), and Grover (1995) judge the examples in (19) to be ungrammatical, Bresnan (1971), and Pollard and Sag (1994) accept them.

- (19) a. %John is difficult for me to believe Mary loves.
b. %Mary is easy for me to think John cheated.

The extraction approach cannot account for why extraction from finite clauses is bad for some speakers since extraction is perfectly good with vanilla extraction constructions, like the ones in (20).

- (20) a. Who do you want me to believe Mary loves?
b. John, I really want you to believe Mary loves.

The control approach, on the other hand, cannot account for why extraction from finite clauses is good for the other speakers since the controllee of control predicates typically occurs within a non-finite VP complement, and in (19), the VP complement is finite.

Despite these apparent problems, the relegation of the TC to either extraction or control is inevitable in most major syntactic frameworks, since they lack any

means to capture the common ground shared by extraction and control. In HPSG (Pollard and Sag 1994, Grover 1995), control is handled in the lexicon, whereas extraction is treated in the syntax, with a strict line drawn between the two components. In GB (Chomsky 1977, 1981, 1986, Cinque 1990), both control and extraction are handled in the syntax, but by different modules. The occurrence and interpretation of the understood subject of a non-finite VP complement, control, is basically regulated by the PRO theorem or NP-movement, and extraction is regulated by subadjacency constraints, together with some surface filters. Since control and extraction are regulated by independent modules, a unified account of a phenomenon which exhibits both of the properties is not feasible.

3. NEW PROPOSAL

3.1. SIMILARITIES AND DIFFERENCES BETWEEN EXTRACTION, CONTROL, AND TOUGH CONSTRUCTION. Table 1 represents the common ground between the TC and extraction, on the one hand, and control on the other hand. From this point, I will adopt Construction Grammar terminology for extraction and control. We call extraction the Left Isolate Construction, and we call control the Coinstantiation Construction.

	Left Isolate	Coinstantiation	Tough
Non-local argument construal	+	+	+
Long distance dependency	+	-	+
Single valence structure	-	+	+

Table 1. Relationship between the Left Isolate, Coinstantiation and Tough Construction

The first row of Table 1 shows that all three constructions involve non-local argument construal. In other words, they involve specification of an argument's semantics by an argument that is not realized syntactically within a local structure. So, in the resulting construct, some argument is construed as instantiating a valence requirement of a predicate that occurs in a syntactic domain not governed by that predicate. The second row of Table 1 indicates that the TC shares with the Left Isolate Construction the property, absent in the Coinstantiation Construction, that the dependency may be a long-distance one. The third row indicates that the TC shares with the Coinstantiation Construction the property, absent in the Left Isolate Construction, that the entire construction occurs within a single valence structure. That is, in the TC, there is no left-isolate sister as a distinct syntactic position, as there is in the Left Isolate Construction.

3.2. THE THREE CONSTRUCTIONS. Now I will demonstrate how the similarities and differences between the three constructions are captured in a construction grammar approach.

Before discussing each of these constructions, two preliminary notes are in order. First, the descriptions of the constructions we will look at employ feature structures consisting of attributes and their appropriate values. Second, the descriptions of the constructions will be fleshed out just enough to illustrate the similarities and the differences between them.

3.2.1. THE LEFT ISOLATE CONSTRUCTION. The Left Isolate Construction (= extraction) is given in Figure 1.

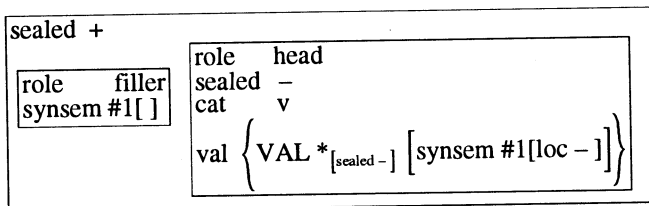


Figure 1. Left Isolate Construction

Figure 1 says that the Left Isolate Construction consists of two daughters. It is a verbal structure since the right head daughter is verbal, as indicated by 'role head', 'cat, v'. This is achieved by an independent Head Principle in Construction Grammar. Note that since the right daughter does not mention any information about specific verbal forms, it can be either finite or non-finite.

If we move to the valence structure at the bottom of the right daughter, we see some notation in the valence set. For the moment, ignore the operator Kleene-starred VAL with the attached subscript feature structure [sealed -]. The remaining valence value, represented as 'synsem #1[]', when read in conjunction with the same synsem value in the left daughter, has the effect of saying that the left daughter fills or satisfies a valence requirement of a predicator that occurs in the right daughter. Note that what is unified is syntactic and semantic information, so that the case marking effects we saw in (17) are maintained between an 'extracted' element and the corresponding valence requirement. The 'loc -' value in the synsem of the right daughter represents the fact that the valence requirement is not locally instantiated, that is, it is extracted. Note also that no restriction is imposed on the extracted element so that the extracted element can satisfy any grammatical function in the valence of the right daughter.

Recall that there might in principle be an unrestricted number of other intervening predicators between the left-isolated valence element, and the predicator whose valence it satisfies. The operator VAL* indicates this.

The formal expression VAL* [synsem #1] has the effect of saying that there is a valence element, which unifies with the left-isolated daughter, which is arbitrarily multiply embedded in the right daughter.

The feature structure 'sealed +' in the external structure of the Left Isolate Construction captures the fact that the Left Isolate Construction is an island, that nothing can be extracted or left-isolated from it, as we saw in (8)-(9). This feature structure is in contrast with the one 'sealed -' in the right daughter since any predicator in the right daughter may permit its argument to be left-isolated.

3.2.2. THE COINSTANTIATION CONSTRUCTION. A simplified version of the Coinstantiation Construction (= control) is given in Figure 2.

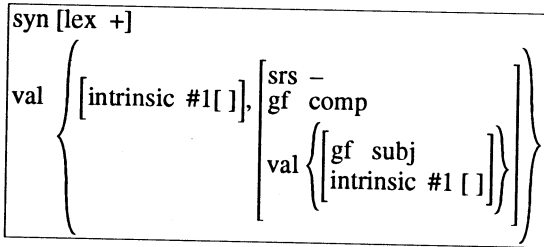


Figure 2. Coinstantiation Construction

The top line of the Coinstantiation Construction encodes the fact that the construction is licensed by a specific lexical item.

The lexical predictor heads a valence structure that has two valence elements: the first is an argument, and the second is a controlled subjectless non-finite verbal complement, as indicated by 'srs -' and 'gf comp'.

The unification index #1[] in the controlled complement, when read in conjunction with the same intrinsic value of the first valence element, indicates that the subject requirement of the complement has its intrinsic features — non-relational syntactic and semantic information — supplied by another valence element of the predictor. The fact that the controlling element is not realized as a separate constituent corresponds to the fact on other accounts that it will be satisfied by a constituent that appears in an argument, as opposed to a non-argument, position.

3.2.3. THE TOUGH CONSTRUCTION. A simplified version of the TC is given in Figure 3.

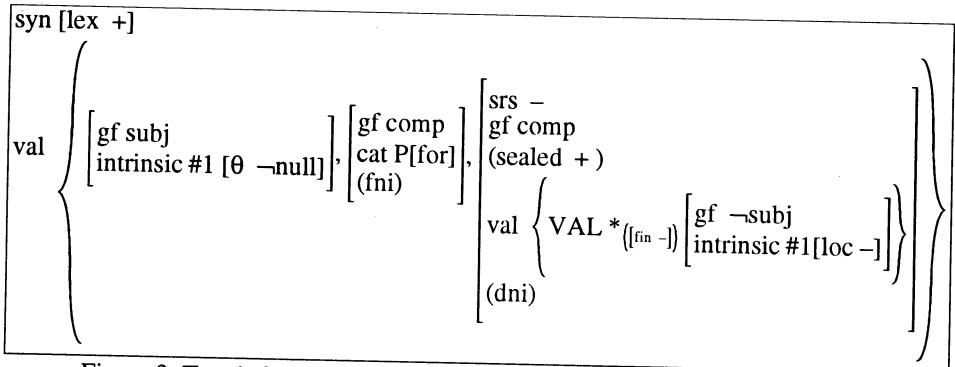


Figure 3. Tough Construction (a simplified version)

The top line says that, like Coinstantiation, the construction is licensed by a specific lexical item. The valence structure says that the predicate has three valence elements.

The first element encodes the *tough* subject. As indicated in footnote 2, the *tough* subject may be equi, but not raising, objects. This restriction is marked 'θ -null' in the *tough* subject intrinsic valence.

The second and third valence elements represent the two oblique arguments, indicated as 'gf comp' that *tough* predicates take. The second element is a prepositional phrase headed by the preposition *for*. This is the 'for anyone' constituent that we saw in (1b). However, the syntactic realization of this prepositional phrase is not mandatory. When the prepositional phrase is not manifested, the missing constituent is interpreted as people in general or as the speaker or the addressee or some third party whose perspective is being employed in a discourse: free null instantiation (fni).

The third valence element is a subjectless infinitival complement, similar to the controlled complement in the Coinstantiation Construction. The constituent is not always an island, unlike the Left Isolate Construction, as we saw in (14a) and footnote 6. This restriction is indicated by '(sealed +)'. The constituent may also be absent. When it is absent, the address may be able to construe its content from the the context, and is marked as '(dni)' (definite null instantiation) at the bottom of the valence structure.

The third valence element has its own valence requirement. It resembles the valence requirement of the Left Isolate Construction's right daughter in that it allows for multiple embedding. It differs from the Left Isolate Construction in that only intrinsic information is shared with another valence element, namely, the *tough* subject. It differs from the Coinstantiation Construction in that the grammatical function of the unified element must be a non-subject within the complement so that the case markings between the *tough* subject and the *tough* gap must be different. As we saw in (19), the coindexation of the *tough* subject with the *tough* gap may or may not, depending on the dialects, be made across the a finite clause boundary that an intervening predicator hosts. This fact is captured by marking '(fin -)' on a valence structure that any intervening predicator has.

4. CONCLUSION. In this paper, we hope to have shown that the TC has its own syntactic properties that cannot belong to either the extraction construction or the control construction.

On the Construction Grammar account, the TC is subsumed neither under the Left Isolate Construction nor under the Coinstantiation Construction. The TC instead shares particular properties of both constructions, and satisfies its own well-formedness conditions. Such an account is not available to other approaches due to their general organizational principles.

If the analysis proposed is correct, it casts doubt on the basic assumption of the distinction between lexicon and syntax in GB or HPSG. The case study of the TC shows that such a widely accepted view cannot be maintained, and points to the desirability of a more flexible design of grammar such as that provided by Construction Grammar.

NOTES

1. This issue should not be confused with another independent issue of equi vs. raising which concerns the semantic relation between the *tough* predicate and the subject. See also Grover (1995).

2. Jacobson (1992) notes, however, that it is sometimes hard to get a deeply embedded gap in some *tough* sentences:

- (i) ?John is hard for me to imagine Mary wanting to meet.

But such difficulties might be explained by the semantic oddness of the to-infinitive phrase, i.e., 'I imagine Mary wants to meet X'. More generally, Zwicky (1987) independently notes that the anomaly seems to be related to the fact that object-to-subject raised NPs, unlike equi NPs, usually do not participate in the TC. Contrast (i) and (ii) with (iii):

- (ii) ?John is hard to expect to marry Leslie.
- (iii) John is hard to persuade to marry Leslie.

I will not look into the details of the issue, since it does not have any bearing on the extraction vs. control issue. The observed differences between raising objects and equi objects will, however, be incorporated in our analysis in section 3.

3. Or the subadjacency condition since it subsumes many, though not all, of the island constraints.

4. Grover (1995) notes, however, that some NPs including picture nouns behave somewhat differently:

- (i) *Sandy is hard to sell some pictures of.
- (ii) Who did you sell some pictures of?
- (iii) *France would be impossible to meet the king of.
- (iv) Which country did George meet the king of?

Since the examples involved share the same syntactic structure, the unexpected difference in grammaticality shown in (i, ii) and (iii, iv) might be due to some semantic factors, for which we offer no account here.

5. See Akatsuka (1979) for reasons why a predicate like *possible* cannot function as a *tough* predicate.

6. Thus there is a well-known difference in grammaticality judgements between (14a) and (i):

- (i) *Which sonata is this violin easy to play_on_?

We do not, however, offer any account of the difference here.

ACKNOWLEDGEMENTS

I am grateful to Chuck Fillmore, Andreas Kathol, especially Paul Kay, and David Peterson for their helpful contributions. Usual disclaimers apply.

REFERENCES

- AKATSUKA, NORIKO. 1979. Why *Tough*-Movement is impossible with *possible*. Chicago Linguistic Society 15.1-8.
- BRESNAN, JOAN. 1971. Sentence Stress and Syntactic Transformations. Language 47.257-281.
- CHOMSKY, NOAM. 1977. On *wh*-movement. Formal syntax, ed. by P. Culicover, T. Wasow, and A. Akmajian, 71-132. New York: Academic Press.
- CHOMSKY, NOAM. 1981. Lectures on Government and Binding. Dordrecht, Holland: Foris.

- CHOMSKY, NOAM. 1986. *Knowledge of Language: Its Nature, Origin and Use*. New York: Praeger.
- CHOMSKY, NOAM. 1995. A minimalist program for linguistic theory. *The minimalist program* 167-218. Cambridge, MA: MIT Press.
- CINQUE, GUGLIELMO. 1990. *Types of A'-Dependencies*. Cambridge, MA: MIT Press.
- FILLMORE, CHARLES, and PAUL KAY. To appear. *Construction Grammar*. Stanford, CA: CSLI.
- GAZDAR, GERALD; EWAN KLEIN; GEFREY. K. PULLUM; and IVAN SAG. 1985. *Generalized Phrase Structure Grammar*. Cambridge, MA: Harvard University Press.
- GROVER, CLAIRE. 1995. *Rethinking Some Empty Categories: Missing Objects and Parasitic Gaps in HPSG*. Edinburgh: University of Edinburgh dissertation.
- HUKARI, THOMAS, and ROBERT LEVINE. 1991. On the Disunity of Unbounded Dependency Constructions. *Natural Language & Linguistic Theory* 9.97-114.
- JACOBSON, PAULINE. 1992. The Lexical Entailment Theory of Control and the *Tough*-Construction. *Lexical Matters*, ed. by I. Sag and A. Szabolcsi, 269-299. Stanford: CSLI.
- KAY, PAUL, and CHARLES FILLMORE. 1998. Grammatical constructions and linguistic generalizations: the *What's X doing Y?* construction. *Language* 74.
- POLLARD, CARL, and IVAN SAG. 1994. *Head-Driven Phrase Structure Grammar*. Stanford, CA: CSLI.
- ZWICKY, ARNOLD. 1987. Slashes in the passive. *Linguistics* 25.639-669.

REDUCED RELATIVES: LEXICAL CONSTRAINT-BASED ANALYSIS

Hana Filip
University of Rochester

1 Introduction

One of the central controversies in sentence processing concerns the role of structural factors. The touchstone for the two main competing schools of thought, structure-based and constraint-based, are garden-path sentences, such as *The horse raced past the barn fell*, discussed in sentence processing literature since at least Bever (1970). I will here present a recent structure-based approach to garden-path sentences of Stevenson and Merlo (1997), which is couched within Government and Binding Theory, and contrast it with an alternative account that presupposes constraint-based approaches to natural language description in psycholinguistics and linguistics (HPSG and Construction Grammar). The goal is to show that thematic properties, which characterize the two fuzzy cluster concepts Proto-Agent and Proto-Patient (Dowty, 1988, 1991), can account for a great number of processing differences between sentences with reduced relative clauses based on unergative verbs, on the one hand, and on unaccusative verbs, on the other hand. One advantage of this novel way of looking at the garden-path phenomenon is that it allows us to understand the influence of the main predicate in a sentence on the magnitude of the garden-path effect. This type of data has so far gone unnoticed, to my knowledge, and it is problematic for *purely* structure-based accounts.

2 Data and the Main Question: Sentences with Reduced Relative Clauses

It is well-known that sentences with reduced relative clauses vary from being hard or nearly impossible to interpret to being very easy, at least on a first pass. For example, (1a) is unacceptable or difficult to process, while (1b) is far easier:

- (1) a. #The horse RACED past the barn fell.
b. The butter MELTED in the pan was fresh.

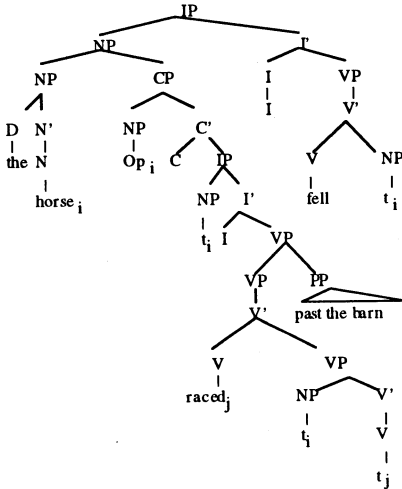
Many researchers agree that this difference is correlated with the type of verb used in the reduced relative clause. As far as their inherent or basic lexical class is concerned, *raced* is unergative and *melted* unaccusative, a distinction introduced by Perlmutter (1978), and also noticed by (Hall, 1965). In standard Government and Binding approaches (Hale and Keyser, 1993, for example), unergative verbs are syntactically characterized as having an external argument, but no direct internal argument, while unaccusative verbs have no external argument, and a direct (non-clausal, non-PP) internal argument. According to semantic characterizations given by Van Valin (1990) and Dowty (1991), for example, unergative verbs tend to entail agentivity in their single argument and to be aspectually atelic. Unaccusative verbs take a patient-like argument and are mostly telic. Assuming that the difference between (1a) and (1b) is correlated with the unaccusative-unergative distinction, the following questions arise: 'What is the nature of the lexical knowledge that differentiates between unergative and unaccusative verbs?'

3 Stevenson and Merlo (1997)

According to Stevenson and Merlo, it is the differences in structural configurations at the lexical level that set unaccusative verbs apart from unergative ones, and that ultimately result in the differences in the processing difficulty of sentences with

reduced relative clauses. They presuppose Hale and Keyser's (1993) lexical syntax and also apply some ideas of Dowty's (1979) much earlier lexical decomposition approach that is grounded in generative semantics. On Stevenson and Merlo's view, the difficult sentence (1a), for example, is assigned the syntactic structure (2):

(2) Stevenson and Merlo (1997:382)



The parser, a symbolic/connectionist hybrid, developed by Stevenson (1994a,b), cannot activate the structure (2) needed for a grammatical analysis of sentences like (1a), "because of its limited ability to project empty nodes and to bind them in the structure" (Stevenson and Merlo, 1997:397). Hence, their model predicts that sentences with reduced relatives headed by passive participles derived from unergatives are judged "all mostly or completely unacceptable" (p.355). In particular, manner of motion verbs "lead to a severe garden path in the RR construction" (p.353). There are "sharp distinctions between unergative RR clauses and RR clauses with other verbs" (p.396). Some of their examples (p.353) are given in (3):

- (3) a. The students advanced to the next grade had to study very hard.
- b. The clipper sailed to Portugal carried a crew of eight.
- c. The troops marched across the fields all day resented the general.
- d. The model planet rotated on the metal axis fell off the stand.
- e. The dog walked in the park was having a good time.
- f. The ship glided past the harbor guards was laden with treasure.

In contrast, "unaccusative RRs are all completely acceptable or only slightly degraded" (p.355). Stevenson and Merlo's (1997:353) examples are given in (4):

- (4) a. The witch melted in the Wizard of Oz was played by a famous actress.
- b. The genes mutated in the experiment were used in a vaccine.
- c. The oil poured across the road made driving treacherous.
- d. The picture rotated 90 degrees was easy to print.

Although Stevenson and Merlo (1997) acknowledge that other factors are important in processing, such as lexical frequencies, thematic fit, context, world knowledge, etc., syntactic constraints are claimed to override all the other factors in cases of conflict and syntactic constraints alone can cause failure to interpret a sentence (see p.392).

However, a close look at the relevant data reveals a different picture from that proposed by Stevenson and Merlo. Sentences with reduced relatives based on unergative verbs, including manner of motion verbs, manifest a considerable degree of variability in acceptability, and, in fact, perfectly acceptable sentences of this type are easy to find. Consider examples in (5):

- (5) a. The victims rushed to the emergency room died upon their arrival.
- b. The cart rolled down the ramp overturned when it hit a bump.
- c. The dog walked in the park was wearing a choke collar.
- d. The diplomats jetted to Iraq were unable to diffuse the crisis.
- e. The convict moved into an isolation cell became depressed.
- f. The soldiers marched across the fields were ambushed by the enemy.

There are also sentences with reduced relatives headed by participles derived from unaccusative verbs that are more difficult to interpret than some sentences with reduced relatives based on unergatives. Examples are given in (6):

- (6) a. The theatre darkened for the movie frightened some preschoolers.
- b. The plaster hardened in the oven cracked with loud popping sounds.
- c. The bubble burst in the hallway made the principle jump.
- d. The paper yellowed in the sun was wrinkled.

What has so far gone unnoticed is that both types of sentences exhibit similar gradient effects in acceptability that are crucially influenced by the lexical semantics of the main verb in a matrix clause. To put it in the simplest terms, the fewer agent-like properties and the more patient-like properties the main verb assigns to its subject, the easier the whole sentence with a reduced relative clause is judged. This idea will be discussed in detail in section 4, but let me illustrate it here with a few examples. In (7a) the subject of *complained*, *the patients*, is a volitional agent in the denoted event, and we see that the whole sentence is less acceptable than (7b) with *died* as the main verb, whose subject undergoes a change of state. A similar contrast can be found in (8):

- (7) a. The patients rushed to the emergency room *#complained to the nurse*.
- b. The patients rushed to the emergency room *died*.
- (8) a. The Great Dane walked in the park *#tugged at the leash*.
- b. The Great Dane walked in the park *wore a choke collar*.

Similar to reduced relatives with passive participles derived from unaccusative verbs, such as *darkened* in (9), we see that the use of *frightened* as opposed to *smelled* in the matrix clause is correlated with a difference in the acceptability of the whole sentence. The reason is that *frightened*, but not *smelled*, presents the subject *the theatre* as the cause of the change of the psychological state in the referent of the direct object *some preschoolers*. Other similar examples are given in (10):

- (9) a. The theatre darkened for the movie #frightened some preschoolers.
 b. The theatre darkened for the movie smelled like popcorn.
- (10) a. The genes mutated in the experiment #attacked their host.
 b. The genes mutated in the experiment were used in a new vaccine.

To summarize, there is no sharp contrast between sentences that contain unergative-based reduced relatives and those that contain unaccusative-based ones. Second, and more importantly, both types of sentences are similar in exhibiting clear gradient effects with respect to acceptability judgments and parsing difficulty, which are influenced by the lexical semantics of the main verb in a matrix clause. It must be emphasized that sentences with reduced relatives based on unergative verbs manifest far greater variability and are on average significantly harder to interpret than unaccusative-based ones.

Most importantly, different degrees of acceptability observed in (5) - (10) resist an explanation in purely structure-based terms, including those couched in lexical syntax of Stevenson and Merlo (1997). Recall that they predict that *all* sentences with reduced relatives headed by inherently unergative verbs are predicted to pose 'sharp difficulty' (p.392) for an interpreter, and they cannot be assigned a grammatical analysis by the parser. In order to account for unaccusative-based reduced relatives that are *not* easy to interpret, such as those in (11), Stevenson and Merlo resort to semantic factors, in addition to structural ones, to argue that they are unergative. The reason is, according to them, that verbs like *caramelize*, *solidify*, and *yellow* entail 'internal causation' (see Levin and Rappaport Hovav, 1995:210-11) in their semantic description, a feature that distinguishes unergative verbs from unaccusative ones, the latter being 'externally caused' (see *ibid.*). Only unaccusative verbs, but not unergative ones, sanction the expression of a causal 'external' agent or force as their subject-NP, when they are used transitively. Stevenson and Merlo illustrate this point with examples in (12).

- (11) a. #The candy caramelised in an hour burned.
 b. #The wax solidified into abstract shapes melted.
 c. #The paper yellowed in the sun shrank.
- (12) a. The sun yellowed the paper. Stevenson and Merlo, 1997:365
 b. #The chain-smoker yellowed the papers.
 c. #The sculptor solidified the wax.
 d. The sculptor hardened the wax.

By this test, *yellow* in (12b) and *solidify* in (12c) are unergative and *harden* in (12d) is unaccusative. Moreover, (12b) is less acceptable than (12a), because its subject referent may be intentionally involved in the denoted event, while in (12a) it cannot, the denoted change of state is "indirectly brought about by some natural force" (p. 365). Yet at the same time, Stevenson and Merlo observe (p. 357) that agentive manner of motion verbs, which are unergative, used transitively *require* their subject argument to be an Agent: cp. **The downpour marched the soldiers to the tents* vs. *The commander marched the soldiers to the tents*. (This observation is based on Cruse, 1972; Jackendoff, 1972; Levin and Rappaport Hovav, 1995.) This inconsistency clearly indicates that a test based on the possibility of the overt expression of an agent/causal force cannot be the right diagnostic for deciding the membership of verbs in the unaccusative and unergative class. The main source of confusion are here the correlations 'non-agentivity - external causation - possibility

of an overt expression of an external agent/force', on the one hand, and 'agentivity - internal causation - prohibition against an overt expression of an external agent/force', on the other hand. Moreover, what is lacking is a precise characterization of the notions 'internal and external causation', introduced by Levin and Rappaport Hovav (1995).

The fact that Stevenson and Merlo do resort to rather subtle semantic criteria in order to account for difficult cases is instructive, because it shows that purely structure-based accounts are insufficient and that semantic explanations are necessary in addition to structure-based ones. Indeed, one may ask to what extent, if at all, syntactic factors are necessary in addition to semantic ones in order to account for the garden-path phenomenon. If we focus on the differential semantics of the verbs in the material discussed here, we can begin to understand the overlapping distribution of sentences with reduced relatives as well as the great deal of variability with respect to how good or bad they are judged to be. A particularly fruitful way of capturing the relevant semantic entailments of verbs is in terms of Dowty's theory of thematic Proto-Roles and argument selection. In the next section I will introduce Dowty's theory and its application to the interpretation of sentences with reduced relative clauses.

4 Thematic Properties of Verbs and Language Comprehension

4.1 Dowty (1988, 1991)

Dowty proposes that the only thematic roles are two cluster concepts, Proto-Agent and Proto-Patient, each characterized by a set of verbal entailments, given in (13) (see Dowty, 1991:572):

- (13) Contributing properties for the Agent Proto-Role:
- a. volitional involvement in the event or state
 - b. sentience (and/or perception)
 - c. causing an event or change of state in another participant
 - d. movement (relative to the position of another participant)
 - (e. referent exists independent of action of verb).

- Contributing properties for the Patient Proto-Role:
- a. undergoes change of state
 - b. incremental theme
 - c. causally affected by another participant
 - d. stationary relative to movement of another participant
 - (e. does not exist independently of the event, or not at all).

Proto-Agent and Proto-Patient properties are 'higher-order generalizations about meanings'. "[A]n argument of a verb may bear either of the two proto-roles (or both) to varying degrees, according to the number of entailments of each kind the verb gives it" (Dowty, 1991:547). The Argument Selection Principle (p.576) determines the direct association of clusters of Proto-Agent and Proto-Patient properties with grammatical relations in a many-to-one fashion:

- (14) Argument Selection Principle
- In predicates with grammatical subject and object, the argument for which the predicate entails the greatest number of Proto-Agent properties will be lexicalized as the subject of the predicate; the argument having the greatest number of Proto-Patient properties will be lexicalized as the direct object.

4.2 Compatibility between Subjects in Sentences with Reduced Relative Clauses

I suggest that a decisive factor (though not the only one) for the acceptability of a sentence with a reduced relative clause is the constellation of Proto-Patient and Proto-Agent properties assigned by the main verb in a matrix clause and the passive participle in a relative clause to their respective subjects. This idea is formulated in (15):

(15) Hypothesis

The acceptability of sentences with reduced relative clauses, headed by passive participles derived from unergative and unaccusative verbs, increases when the passive participle and the main verb of a matrix clause assign their subject-NPs more Proto-Patient, and fewer Proto-Agent, properties.

The intuition behind (15) is that sentences are easier to interpret when there is an internal coherence among the interpretations of their constituents. One way this coherence can be achieved is in terms of compatible or even identical assignments of thematic properties to different NP arguments that are associated with one and the same participant in the domain of discourse. In sentences with a reduced relative clause the internal coherence depends in part on the thematic compatibility between the subject NP in the matrix clause and the PRO subject of the reduced relative. Let us take (1a) #*The horse raced past the barn fell*. At the point when *raced* is processed, the preferred syntactic-semantic pattern is that of the main clause with a highly agentive subject-NP. However, when *fell* is processed, *raced* must be understood instead as a passive participle. Passive participles in general presuppose the existence of corresponding active transitive verbs whose subjects correspond to active direct objects (see Sag and Wasow, 1997:164, for example; but passive subjects do not always correspond to active direct objects, see Postal, 1986; Zwicky, 1987; and others). Let us now look at the assignment of thematic properties by the verb *raced* in its intransitive and transitive (causative) use:

(16) Distribution of Proto-Agent (PA) and Proto-Patient (PP) properties for *RACED*

a. Vi: unergative

The horse RACED past the barn

|
PA
(+ volition)
+ sentience
+ movement

b. Vt: lexical causative

The rider RACED the horse past the barn

PA	PA & PP	
+ volition	(+volition)	+causally affected
+ sentience	+ sentience	
+ causing	+ movement	
event/change		

The subject *the horse* of the intransitive *raced* corresponds to the object of the transitive *raced*. They share three Proto-Agent properties, two of which are also the Proto-Agent properties assigned to the subject *the rider* of the transitive *raced*. Although *the rider* and *the horse* as arguments of the transitive *raced* are close in Proto-Agent properties, they differ in so far as the former is the 'causer of the denoted event', a Proto-Agent property, while the latter is entailed to have the

Proto-Patient property of being 'causally affected by another participant'. This motivates the lexicalization of *the rider* and *the horse* as the subject and the direct object, respectively, as predicted by Dowty's Argument Selection Principle (see (14)).

Typical unergative verbs used transitively do not fit the semantics of prototypical transitives. Intuitively, the latter can be understood in terms of a 'billiard ball model', as Langacker (1986) calls it, which involves two participants that interact in an asymmetric and unidirectional way, whereby one of them is directly affected by some action (possibly involving movement, contact, effect, and the like) instigated or caused by the other participant. Now, the direct object of an inherently unergative verb used transitively corresponds to the intransitive subject and has a thematic make up of a "good" Agent (having a few Proto-Patient and several Proto-Agent properties), rather than of a "good" Patient. That is, we may understand a sentence like (16b) as being semantically decomposable into two causally related clauses: (i) *The rider did something to the horse* and (ii) *The horse raced*. (See also Fillmore, 1971:46-7, for a similar example.) The horse is a Patient-like participant from the point of view of the first clause and an Agent-like participant from the point of view of the second. Having to reconcile these two different perspectives is directly related to the awkwardness often associated with the transitive use of unergative verbs. This also carries over to passive participles derived from inherently unergative verbs, because prototypical passive participles require a high number of Proto-Patient properties in their subject arguments. These observations correctly predict that reduced relatives with passive participles derived from inherently unergative verbs are hard to interpret if the transitive and/or passive use of unergatives is judged hard. This is illustrated by examples in (17) - (19):

- (17) a. #John waltzed the debutante across the dance floor.
 b. The debutante was waltzed across the dance floor.
 c. #The debutante waltzed across the floor wore a beautiful dress.
- (18) a. John glided the puck across the ice.
 b. #The puck was glided across the ice.
 c. #The puck glided across the ice slipped through the goalie's mitt.
- (19) a. #The trainer danced the bears.
 b. #The bears were danced around the ring by their trainer.
 c. #The bears danced around the ring were amusing.

Of course, not all transitive and passive uses of inherently unergative verbs are odd. For example, *John walked his dog* and *Fido was walked by John tonight* are perfectly natural. Apart from the thematic compatibility discussed here, other factors, such as certain expectations related to the occurrence of highly conventionalized combinations of words and general world knowledge, may come into play and override the semantic mismatch described above.

To return to our lead example, the PRO subject of the passive participle in (20) has the same thematic properties as the corresponding active object in (16b), it is not a "good" Patient, and hence it does not fit the prototypical semantics of passives. The analysis of *raced* as a passive participle is then further made difficult by the main verb *fell*, because it assigns the Proto-Agent property 'movement' to its subject *the horse*. If, on the other hand, the main verb of a matrix clause assigns

Proto-Patient, rather than Proto-Agent, property (or properties) to its subject, the magnitude of the garden path effect is diminished. This is illustrated in (21) with the main verb *died*, which is somewhat easier to interpret than (20).

- (20) The horse_i [< PRO_i > *RACED past the barn*] *fell*
- | | | |
|---------------|--|-----------------------------------|
|
PA |
PA & PP |
PP |
| + movement | (+ volition) + causally affected
+ sentence
+ movement | + movement
(+ change of state) |

- (21) The horse_i [< PRO_i > *RACED past the barn*] *died*
- | | | |
|-----------------------------|--|-----------------------------|
|
PP |
PA & PP |
PP |
| + undergoes change of state | (+ volition) + causally affected
+ sentence
+ movement | + undergoes change of state |

Although the difference between (20) and (21) may appear subtle, the subject of *died* is clearly a “better” Patient than the subject of *fell*, as it undergoes a permanent change of state.

Let us now look at sentences with reduced relatives headed by passive participles derived from unaccusative verbs. As (22) and (23) show the subject of the unaccusative *melted*, the object of the corresponding active transitive *melted* and the PRO subject of the passive participle *melted* are all entailed to have the same three Proto-Patient properties: ‘change of state’, ‘Incremental Theme’ and ‘causally affected’.

(22) Distribution of Proto-Agent (PA) and Proto-Patient (PP) properties for *MELTED*:

- | | | |
|--|--|---|
| a. Vi: unaccusative | b. Vt: lexical causative | |
| <u>The butter MELTED in the pan</u> <u>The cook MELTED the butter in the pan</u> | | |
|
PP |
PA |
PP |
| + undergoes change of state
+ Incremental Theme
+ causally affected | + volition
+ sentence
+ causing event/change | + undergoes change of state
+ Incremental Theme
+ causally affected |

- (23) The butter_i [< PRO_i > *MELTED in the pan*] *was fresh*
- | |
|---|
|
PP |
| + undergoes change of state
+ Incremental Theme
+ causally affected |

Consequently, the reanalysis of the unaccusative *melted* as a passive participle in (23) involves no change in the thematic properties of the argument *the butter*. This

facilitates an easy recovery from a garden-path, provided the lexical semantics of the main verb in the matrix clause also entails a high number of Proto-Patient properties in its subject. For example, (23) with the main predicate *was fresh* is judged easier to process than (24) with the main verb *dripped*, as is predicted by the hypothesis in (15).

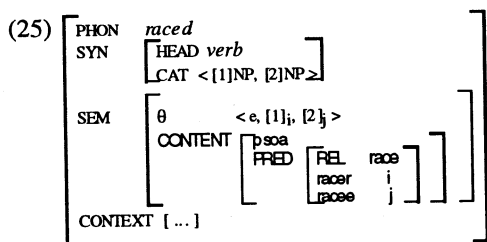
(24) #The butter melted on the stove dripped onto the kitchen floor.

To summarize, I showed that the unaccusative-unergative distinction that Stevenson and Merlo characterize as a syntactic distinction correlated with difficulty or ease of processing in reduced relative clauses can be re-cast as a distinction that concerns the assignment of thematic roles. One advantage of this novel way of looking at the garden-path phenomenon is that it allows us to understand a set of data that have never been systematically commented on before: namely, the influence of the main predicate in a sentence on the magnitude of the garden-path effect (see (7) - (10)). The analysis in terms of Dowty's thematic roles, encapsulated in (15), also makes the correct predictions here. The semantic account proposed here is understood as part of a constraint-based sentence processing system. As far as its linguistic assumptions are concerned, it presupposes constraint-based approaches to natural language description: HPSG and Construction Grammar. In what follows I will outline the main characteristics of a constraint-based grammar assumed here and some plans for the future research.

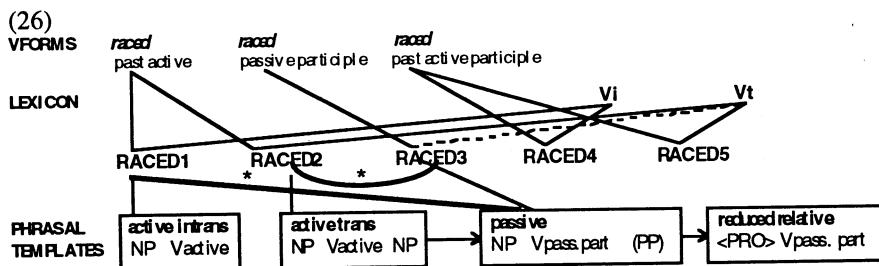
5 Lexical Entries and Constraint-Based Approach

The lexical constraint-based grammar presupposed here has all the main hallmarks of recent versions of HPSG (see Sag, 1997; Sag and Wasow, 1997, for example). Assumptions about lexical semantics of verbs and linguistic information directly associated with extra-linguistic context and general world knowledge are influenced by Fillmore's work and Construction Grammar (see Fillmore and Kay, in press). The overall architecture is monostratal, non-derivational and non-modular. It is characterized declaratively by specifying types of well-formed linguistic expressions (e.g., words, phrases, part of speech classes, argument structure classes, and traditional morphological classes, for example) and constraints on those types. All properties of linguistic expressions are represented as feature structures. The main explanatory mechanism is unification in the narrow sense of structure sharing of token-identical feature structures (cf. Pollard and Sag, 1994).

Constraint-based approaches in linguistics and psycholinguistics share two main assumptions: First, a sentence's interpretation requires satisfaction of multiple (possibly differentially weighted) constraints from various domains of linguistic and non-linguistic knowledge. Second, the integration of such diverse constraints is facilitated by the information contained in lexical entries. Verb-based syntactic and semantic patterns provide a guide for interpreting core aspects of the sentence's structure and meaning, whereby semantic constraints often have a privileged status. Let me, therefore, introduce the main features of lexical entries using a simplified lexical entry for the transitive active *raced* in (25):



(25) contains phonological, syntactic, semantic and pragmatic information, encoded as values of the feature attributes PHON, SYN, SEM and CONTEXT, respectively. The value of SYN encodes syntactic information required for constructing syntactic projections headed by *raced*. The linking between the syntactic (SYN) and semantic (SEM) structure in the lexicon is mediated via co-indexation of syntactic arguments and thematic argument slots, and motivated by Dowty's Argument Selection Principle (here given in (14)). Each argument slot in the thematic structure of a verb corresponds to a cluster of Proto-Agent and/or Proto-Patient properties. Thematic argument slots in turn are co-indexed with individuals in the predication feature structure PRED, which together with 'psoa' constitutes the value of CONTENT. The feature structure PRED captures the assumption that verbs semantically express relations between individuals. The attributes 'racer' and 'racee' include properties that we associate with the individuals 'i' and 'j' on the basis of knowing that the statement 'i raced j' is true. The attributes 'racer' and 'racee' correspond to 'frame-specific participants' in Fillmore (1986) or 'individual thematic roles' in Dowty (1989). In a given single-clause predication, further semantic restrictions on participants are imposed by the interpretation of noun phrases. For example, '[racer i]' will be constrained by the content of the NP filling the '[1]NP' place. The proper association of '[1]NP' with '[racer i]' is ensured through the co-indexation in the thematic structure '[1]_i'. PRED does not provide an exhaustive account of all that we know about the meaning of a given verb. For example, what role an individual plays in a given situation depends on a number of other factors, including world knowledge, which is encoded under 'psoa', a parametrized state of affairs. (For a related, though not identical, use of 'psoa' see Pollard and Sag, 1994; Sag and Wasow, 1997.) Apart from the lexicon, a constraint-based grammar also includes the syntactic level with phrasal templates. This is illustrated in a highly simplified diagram (26):



In general, types at each level of representation are cross-classified in multiple

inheritance hierarchies according to their shared information. The information shared by a given class of objects is associated with a general type and is automatically passed down from the general type to specific members of the class. For example, RACED2 (past active transitive) and RACED5 (past active participle transitive) inherit information from the generic lexical entry for transitive verbs, here located in the node *Vt*. Types directly subsumed under the same supertype represent mutually exclusive alternatives. For example, RACED2 (past active transitive) and RACED3 (passive participle) are mutually exclusive (here indicated by the thick starred lines). Mutually exclusive types often represent multiple interpretation alternatives and differ in frequency of occurrence in the language. For example, the active intransitive use of *raced* is more frequent than the active transitive one. Such frequency information is also encoded in the lexical entries of verbs.

Unification allows us to represent dependencies and connections within one particular level of representation and also among different levels. Feature structures representing compatible types are unified in a new coherent structure by linking them to a single feature structure: e.g., [VFORM PASTACTIVE] %o [SYN Vi]. One advantage of this system is that it allows us to capture the observation that different types of information that characterize the use of a given word are dependent on each other so that accessing one type of information during sentence processing results in accessing others compatible with it. For example, if the sequence *The horse raced* ... is understood as the main clause, the information associated with the verb *raced* will be a complex feature structure comprising the information that this verb shares with all active past tense verbs. If the same sequence is understood as the head noun modified by a reduced relative clause, *raced* will be associated with the information shared with all passive participles, and due to its passive argument structure it will also activate the information associated with the active transitive use of *raced*.

Selected References

- Bever, T. G. 1970. *The Cognitive Basis for Linguistic Structure*. New York: Wiley.
- Carlson, G. N. and M. K. Tanenhaus. 1988. "Thematic Roles and Language Comprehension". *Syntax and Semantics*, Volume 21, pp. 263-300.
- Cruse, D. A. 1972. "A Note on English Causatives." *Linguistic Inquiry* 3, 520-528.
- Dowty, D. R. 1979. *Word Meaning and Montague Grammar. The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ*. Dordrecht: Reidel.
- Dowty, D. R. 1988. "Thematic Proto-Roles, Subject Selection, and Lexical Semantic Defaults", ms. (Paper presented at the 1987 LSA Colloquium, San Francisco)
- Dowty, D. R. 1989. "On the Semantic Content of the Notion of 'Thematic Role'". Chierchia, G., Partee, B. and R. Turner (eds.), *Property Theory, Type Theory and Natural Language Semantics. Volume II: Semantic Issues*. Kluwer Academic Publishers. Dordrecht, Boston, London.
- Dowty, David. R. 1991. "Thematic Proto-Roles and Argument Selection". *Language* 67, 547-619.
- Fillmore, Ch. J. 1968. "The Case for Case". Bach, E. and R. Harms (eds.), *Universals in Linguistic Theory*. New York: Holt, Rinehart and Winston, Inc. pp. 1-88.
- Fillmore, Ch. J. 1971. "Some Problems for Case Grammar". *Twenty-Second Annual Round Table. Linguistics: Developments of the Sixties - Viewpoints for the Seventies*. Georgetown University. pp. 35-56.
- Fillmore, Ch. J. 1986. *On Grammatical Constructions*. Department of Linguistics, The University of California at Berkeley. ms.

- Fillmore, Ch. J. and P. Kay. *Construction Grammar*. Stanford: CSLI Publications. In press.
- Hale, K. and J. Keyser, 1993. "On argument structure and the lexical representation of syntactic relations". K. Hale & J. Keyser (eds.), *The view from building 20*, pp. 53-110. Cambridge, MA: MIT Press.
- Hall, B. 1965. Subject and Object in Modern English. Cambridge, MA: MIT Ph.D. Thesis.
- Jackendoff, R. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: The MIT Press.
- Langacker, R. 1986. "Transitivity, Case, and Grammatical Relations: A Cognitive Grammar Prospectus."
- Levin, B. and M. Rappaport Hovav. 1995. *Unaccusativity: At the Syntax-Lexical Semantics Interface*. Cambridge, Mass.: The MIT Press.
- MacDonald, M. C. (ed.) 1997. *Lexical Representations and Sentence Processing*. A special issue of *Language and Cognitive Processes*. Hove, Great Britain: Psychology Press Ltd.
- McRae, K., Spivey-Knowlton, M. J. and Tanenhaus, M. K. 1998. "Modeling the Influence of Thematic Fit (and Other Constraints) in On-Line Sentence Comprehension." *Journal of Memory and Language*.
- Perlmutter, D. M. 1978. "Impersonal Passives and the Unaccusative Hypothesis". *Berkeley Linguistics Society*. Berkeley 4, 15-189.
- Pollard, C. and I. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Stanford: CLSI/Chicago: The University of Chicago Press.
- Postal, P. M. 1986. *Studies of Passive Clauses*. Albany: State University of New York Press.
- Sag, I. 1997. "English Relative Clause Constructions". To appear in *Journal of Linguistics*.
- Sag, I. and T. Wasow. 1997. *Syntactic Theory: A Formal Introduction*, ms. Stanford University.
- Spivey-Knowlton, M. J. 1996. *Integration of Visual and Linguistic Information: Human Data and Model Simulations*. Ph.D. Thesis. University of Rochester, Rochester, NY.
- Spivey, M. J. and M. K. Tanenhaus. "Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency". *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
- Stevenson, S. 1994a. "Competition and Recency in a Hybrid Network Model of Syntactic Disambiguation." *Journal of Psycholinguistic Research* 23, 295-322.
- Stevenson, S. 1994b. *A Competitive Attachment Model for Resolving Syntactic Ambiguities in Natural Language Processing*. Ph.D. Thesis, Department of Computer Science, University of Maryland.
- Stevenson, S. and P. Merlo. 1997. "Lexical Structure and Parsing Complexity". Maryellen C. MacDonald (ed.) *Lexical Representations and Sentence Processing*. Psychology Press. pp. 349-399.
- Tabossi, P., Spivey-Knowlton, M.J., McRae, K. and M. K. Tanenhaus. 1994. "Semantic Effects on Syntactic Ambiguity Resolution: Evidence for a Constraint Resolution Process". Umiltà, C. and M. Moscovitch (eds.). *Attention and Performance XV*, pp. 589-615. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Tanenhaus, M. and G. Carlson. 1989. "Lexical Structure and Language Comprehension". W. Marslen-Wilson (ed.), *Lexical Representation and Processes*, pp. 529-561. Cambridge: MA: MIT Press.
- Tanenhaus, M. K. and J. C. Trueswell. 1995. "Sentence Comprehension". *Speech, Language, and Communication*. Academic Press.
- Van Valin, R. D. Jr. 1990. "Semantic Parameters of Split Intransitivity". *Language* 22, Number 2, 221-260.
- Zwicky, A. M. 1987. "Slashes in the Passive". *Linguistics* 25, 639-669.

The Discontinuous Intonation Contour: A Case for Rethinking Intonation Contour/Intonation Phrase Isomorphy

Phil Gaines
Montana State University

Imagine the following utterance occurring in a news broadcast following a local election:

'Although the rest of the city incumbents went down to defeat in yesterday's election, the mayor was re-elected handily.'

Here is the intonation contour for the last clause of the sentence:



The mayor was reelected handily.

Now, if the same clause included a medial non-restrictive relative clause, giving new information about the mayor, the intonational shape of the entire clause complex might very well be:



The mayor, whose husband is unemployed, was reelected handily.

Within the framework of Prosodic Phonology, this utterance would be accounted for as containing three intonational phrases, each associated with a separate intonation contour. Missing from such an account, however, is the observation that the shape of the contour of the first sentence is identical, with respect to gross tonal structure, to that of the combined intonational strings associated with the segments of the utterance which have been interrupted by the medial relative clause. I suggest that such an interrupted contour, hereafter called the Discontinuous Intonation Contour (DIC), cannot be adequately accounted for by the standard theory of prosodic phonology, which posits a strictly linear and adjacent series of intonational phrases isomorphic to intonation contours.

The system for describing the structure of intonation contours (hereafter ICs) that will be assumed in this study is that developed in Pierrehumbert 1980 and later extended in Beckman & Pierrehumbert 1986, Pierrehumbert & Beckman 1988, and Pierrehumbert and Hirschberg 1990, which posits two levels of tones, low and high, which are associated with three pitch loci: the pitch accent, phrase accent, and boundary tone. Pitch accents are tones associated with stressed syllables. The phrase accent 'controls the pitch in the region between the last accent and the boundary tone' (Hayes & Lahiri 1991:52). The boundary tone is that associated with the terminal boundary of the contour.

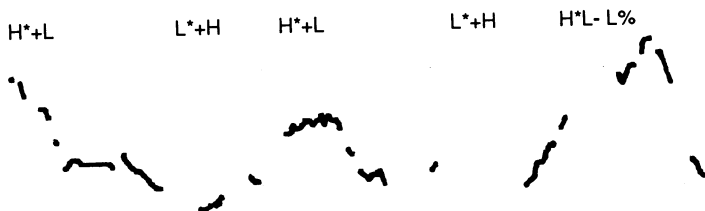
In keeping with the Pierrehumbert model of IC description, it will be maintained here that '[t]he well-formed [ICs] for an intonation phrase are comprised of one or more pitch accents followed by a phrase accent and then a boundary tone' (Pierrehumbert 1980:9). These three components are realized even in the shortest mono-syllabic utterance. Here, H* is the pitch accent, L- is the phrase accent, and L% is the boundary tone:

H*L-L%



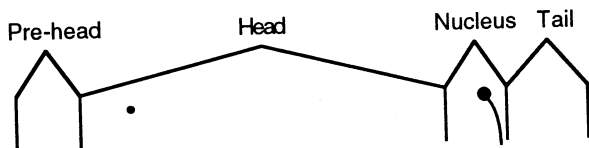
Q h !

Complex contours can combine multiple bi-tonal pitch accents:



[He] beat around the bush and ended up by s aying, 'Qh'

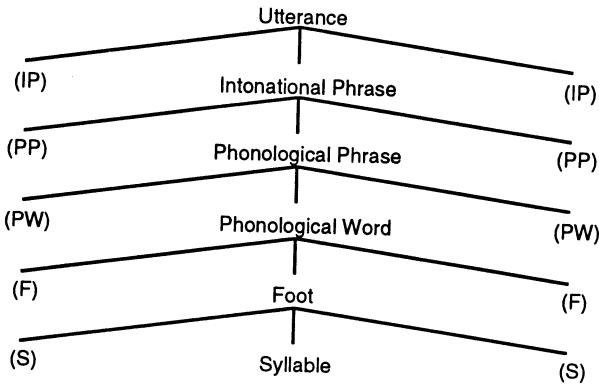
In addition to the obligatory pitch accent/phrase accent/boundary tone configuration, it will also be assumed that a well-formed IC contains a uniquely prominent pitch accent or--in the British terminological tradition--nucleus. Here, the unique prominence of the second pitch accent is acknowledged by identifying it as the nucleus of the IC.



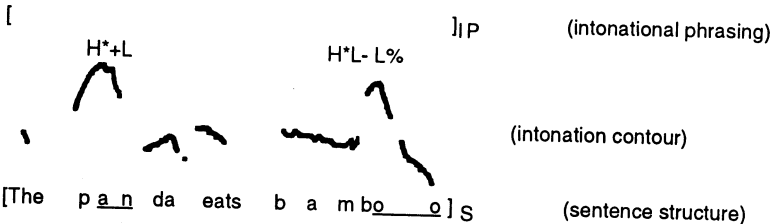
The mayor was reelected han dily

In summary, a well-formed IC consists, from left to right, of 1) one or more pitch accents, the last of which is nuclear, 2) one phrase accent, which provides a tonal transition between the nuclear accent and the end of the IC, and 3) one boundary tone, which establishes the tonal level at the terminus of the IC.

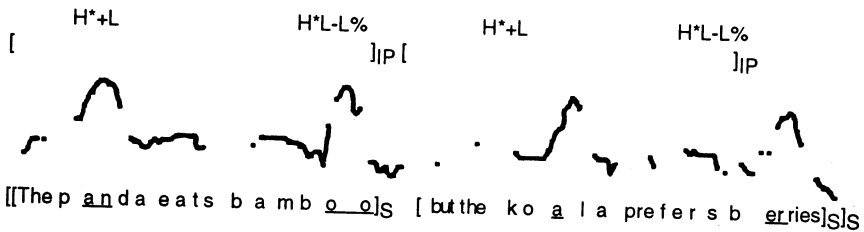
The IC plays an important role in the theory of prosodic phonology, originally formulated by Selkirk and expanded in later work by Selkirk (1980a, 1980b, 1981, 1984, 1986) and Nespor & Vogel (1982, 1986). Prosodic structure represents 'a suprasegmental, hierarchically arranged organization to the utterance, not a simple linear arrangement of segments and boundaries' (Selkirk 1978:111).



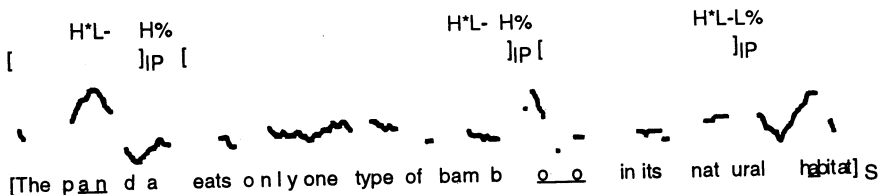
In the prosodic hierarchy, the intonational phrase (hereafter IP) is dominated by the utterance (U) and dominates the phonological phrase (PP). According to the theory, 'an intonational phrase is a unit of prosodic constituent structure with respect to which the characteristic intonational contours of a language are defined' (Selkirk 1984:197). This one-to-one mapping of ICs to IPs is a crucial component of the model. The corollary Strict Layer Hypothesis requires that each level of the prosodic hierarchy consist exclusively of constituents of that level's type and that constituents at any given level are exhaustively dominated by those of the next higher level. The following sentence contains a single clause, IP, and IC.



Multiple IPs are identified in sentences with multiple clauses,



or sentences in which single IPs have been restructured according to certain syntactic or performance conditions--such as the presence of clear pause:

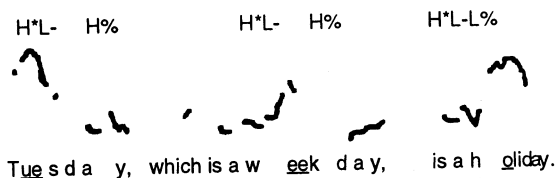


Recent work by Ladd (1986, 1992) and Gussenhoven (1988, 1990) represents significant movement away from the tenets of standard prosodic phonology at the level of the IP. The essential notion from which these analyses emerge is captured in Gussenhoven's claim that '[d]omains for intonational structure should be mapped onto prosodic constituents, but they cannot consistently be mapped onto any particular prosodic constituent' (Gussenhoven & Rietveldt 1992:89). Although it is not possible to explain or exemplify this work here, it should be noted that it presents convincing evidence bringing into question a necessary one-to-one relationship between ICs and IPs. However, none of their discussions specifically treats the phenomenon I have called the discontinuous intonation contour, an IC which, although rendered discontinuous by the presence of an interrupting string, is nevertheless describable as an individual unit which is not isomorphic to an IP.

The presence of certain syntactic structures in sentences mandate an obligatory IP, e.g. non-restrictive relative clauses, direct quote ascriptions, tag questions, vocatives, and preposed adverbials. When the structure in question is sentence-medial, the sentence, under the standard analysis, has three IPs--one formed by the sentence-medial structure and one each formed by the strings to the left and right of it:

[Tuesday,]IP [which is a weekday,]IP [is a holiday.]IP

This requirement is in keeping with the Strict Layer Hypothesis, which mandates that adjacent constituents at a specified level in the prosodic hierarchy be of the same type. Since an IP boundary exists to the left and right of the sentence-medial structure, the remaining strings on both sides of the structure must form their own IPs. Crucially, since IPs and ICs are isomorphic, the intonational material associated with each of these strings is identified as a separate IC, each with its own pitch accent, phrase accent, and boundary tone:



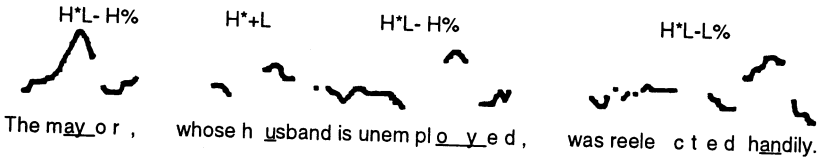
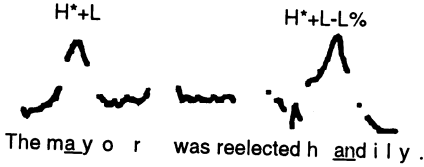
Since the ICs associated with each of the three IPs are easily identifiable as well-formed structures with respect to the pitch accent account of Pierrehumbert, there is no problem with a one-to-one IC/IP mapping in this case. However, the standard model does not account for certain phenomena of ICs seen in sentences

containing the kinds of sentence-medial syntactic structures in view here. Consider again the following two utterances:

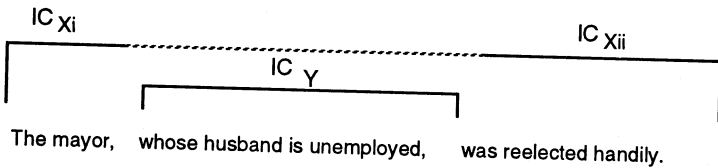
[The mayor was re-elected handily.]IP

[The mayor.]IP [whose husband is unemployed,]IP [was re-elected handily.]IP

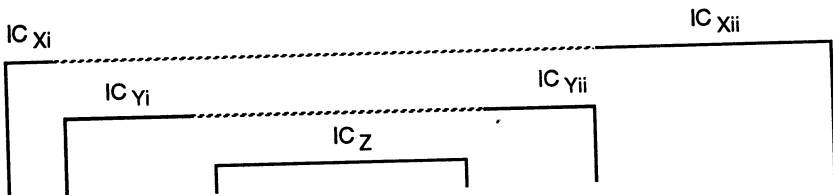
The first sentence contains only one obligatory IP, whereas the second contains three. Note the IC structure of the two sentences:



The standard formulation for this utterance does not capture the observation that the actual phonetic shape of the intonational contour of the combination of the material to the left and right of the medial IP is essentially identical to that of the first sentence without the medial structure. If one posits a sequentially discontinuous yet phonetically unitary IC to account for this phenomenon, the second sentence can be accounted for as having three IPs, yet only two ICs: one continuous, associated with the sentence-medial structure, and one discontinuous, associated with the combined material to the left and right of that structure:



ICY is embedded in the discontinuous IC_{Xi+ii} in a fashion parallel to the embedding of the relative clause in the root sentence. Such embedding could conceivably involve an unlimited number of levels. Here, there are five IPs yet only three DICs:



John, the real hero— and I mean that sincerely— of the trip, ended up with frostbite.
 []IP[]IP[]IP[]IP[]IP

Other sentence types further illustrate the phenomenon: a yes/no question...

^

Do you really think you can get away with this?

...with an adverbial,

Do you really think, as a rational human being, you can get away with this?

a wh-question...

Which poor player do you intend to penalize now?

...with an expletive,

Which poor player -- damn it! -- do you intend to penalize now?

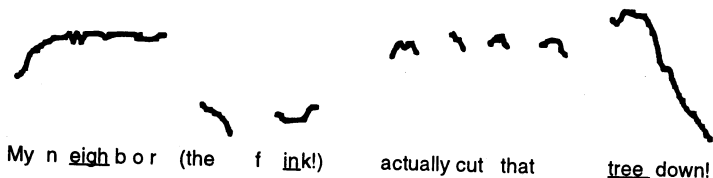
and an exclamatory...

What a groovy trip this vacation is!

...with an adjectival.



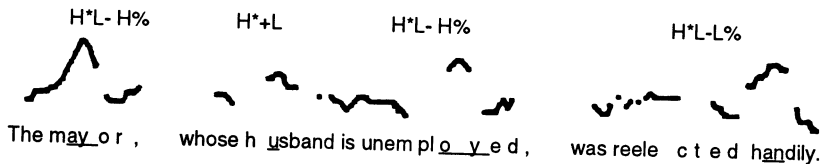
Vivid evidence for the stability of the DIC can be found in utterances with an especially wide pitch range. The resetting of tone after a sentence-medial obligatory IP in order to preserve the stability of the contour of the DIC can sometimes be fairly dramatic:



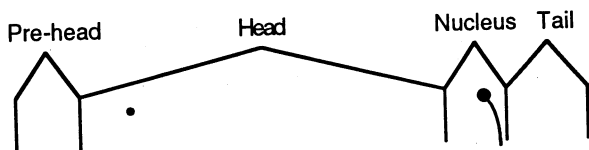
Incidentally, the fact that a given speaker might intone the two separate components of a DIC differently than the uninterrupted contour is not an important argument against the notion of a DIC. Choices made by speakers in intonation vary greatly depending on numerous conditions including attitude, information structure (such as given/new, focus, etc.), interactional context, and personal style.

On the standard account, commitment to a linear organization of IPs demands that the ICs associated with them be described as independent wholes, 'each with its own defining characteristics' (Selkirk 1978:130). This assumption, however, fails to account for important differences in the internal structure of the individual strings of tonal material to the left and right of a sentence-medial structure and of the complete DIC itself.

If each IC associated with an IP is a complete contour containing all requisite components, then each of the three ICs of an utterance containing a sentence-medial IP should be well-formed. In many cases, of course, the pitch accent algorithm can be applied to each of the tonal strings, producing three well-formed sequences. For example,

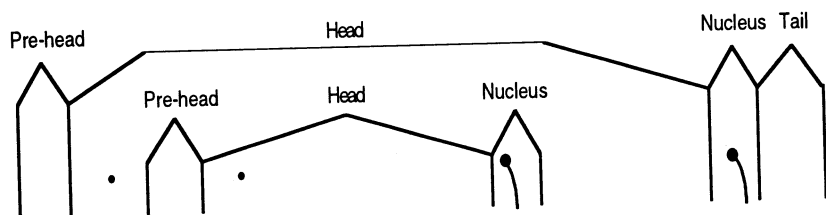


What is missing here, however, is any account of the nuclear accent in the DIC. A nucleus-based treatment of the IC of the utterance without the embedded structure shows a nuclear accent near the end.



The mayor was reelected han dily

According to Couper-Kuhlen, the nucleus is 'the last stressed syllable with any kind of noticeable pitch modulation' (Couper-Kuhlen 1986:81). A Beckman & Pierrehumbert corollary states that 'there is a rule forbidding any accent after the nuclear accent in an intonation contour' (Beckman and Pierrehumbert 1986:266). This accounts for why the accent on 'ma-' cannot be a nucleus. With the sentence-medial structure included, the description shows the DIC interrupted early in the head:



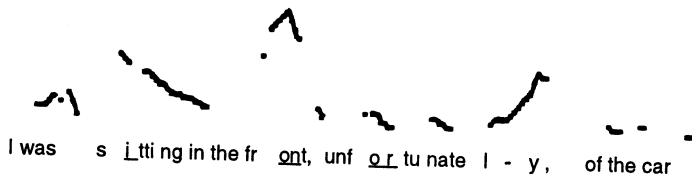
The mayor, whose husband is unemployed, was reelected han dily .

The resulting IC structure shows a nucleus in the second and third IPs but not in the first. It would be gratuitous to posit nucleus status to 'ma-' in the first utterance merely because an IP boundary is 'required'. Since by definition all ICs have a single nucleus, the presence of DICs reveals the importance of accounting for the fact that prosodic units which are supposedly equal in quantity (i.e., IPs) are associated with ICs that differ significantly in quantity--that is, one having a nucleus and one not.

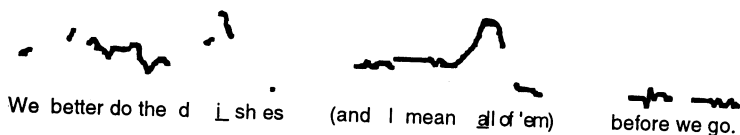
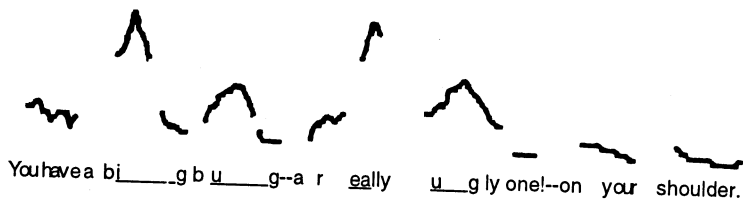
DIC segments reveal not only a difference in IC quantity as regards the presence or absence of a nucleus but even as regards the presence of a pitch accent itself. It is altogether possible for one segment of a DIC to have no pitch accent at all:



I was sitting in the fr ont of the car

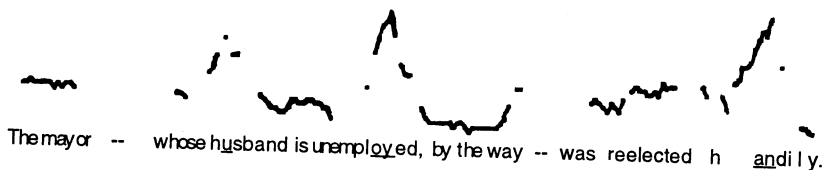


In the second sentence, the tail--consisting of all unaccented syllables following the nucleus--is separated from the nucleus. To identify the third string, in which no pitch accent is present, as a separate IC is highly suspect and runs counter to both the nucleus-based and pitch-accent based systems of IC description. Other examples can be shown in which segments of DICs contain no pitch accent. In these two examples, as in the previous one, the string containing no pitch accent is that following the medial structure:

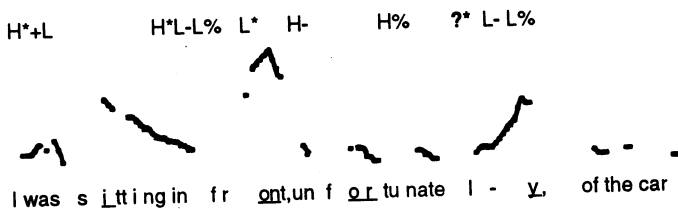


However, it is also possible for the relevant string to be the first in the sentence, as, for example in this case, in which no pitch accent is associated with 'ma-' due to the fact that 'mayor' is old information.

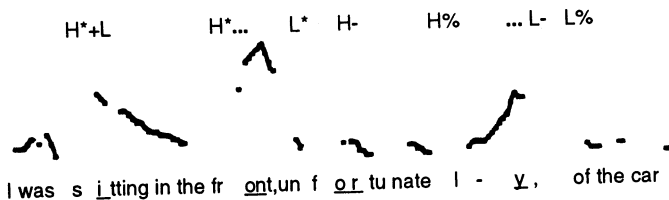
'How did the mayor with that rich salesman husband of hers do in the election yesterday?'



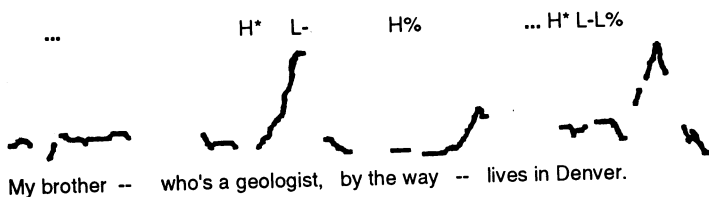
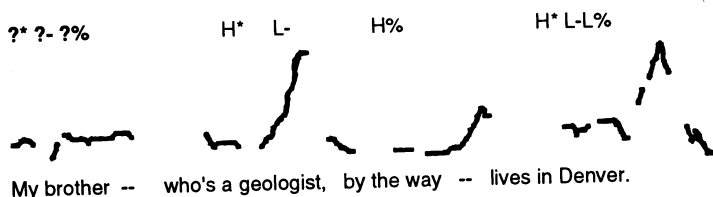
The existence of DICs calls for a change in the way ICs are represented. In this example, for instance, the positing of a phrase accent and a boundary tone located at the word 'front' suggests the termination of a complete IC, which is not the case.



Instead, a DIC, including the tonal material associated with 'of the car', is in evidence, and the only actual phrase accent and boundary tone are located at the very end of the sentence. The use of ellipses indicates the discontinuity and allows for a clear representation of all and only the actual tonal components of the interrupted contour:



Ellipses can also represent the complete absence of salient tonal material in a DIC segment, as in the following sentence in which 'my brother' is given:



Recursively embedded DICs could include double ellipses in their representations:

- Gussenhoven, C. 1988. Intonational phrasing and the prosodic hierarchy. *Phonologica* 1988, ed. by W. Dressler, H. Luschützky, O. Pfeiffer, and J. Rennison. London: Cambridge University Press.
- Gussenhoven, C. 1990. Tonal association domains and the prosodic hierarchy in English. *Studies in the pronunciation of English: a commemorative volume in honour of A. C. Gimson*, ed. by S. Ramsaran. London: Routledge.
- Gussenhoven, C., and A. Rietveld. 1992. Intonation contours, prosodic structure, and pre-boundary lengthening. *Journal of phonetics* 20.283-303.
- Hayes, B., and A. Lahiri. 1991. Bengali intonational phonology. *Natural language and linguistic theory* 9.
- Ladd, D. 1986. Intonational phrasing: the case for recursive prosodic structure. *Phonology* 3.311-340.
- Ladd, D. 1992. Compound prosodic domains. Paper distributed by Department of Linguistics, University of Edinburgh.
- Nespor, M., and I. Vogel. 1982. Prosodic domains of external sandhi rules. *The structure of phonological representations*, ed. by H. Vander Hulst and N. Smith. Dordrecht: Foris Publications.
- Nespor, M., and I. Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris Publications.
- Pierrehumbert, J. 1980. The phonology and phonetics of English intonation. PhD dissertation. Massachusetts Institute of Technology.
- Pierrehumbert, J., and M. Beckman. 1988. *Japanese tone structure*. Cambridge, MA: The MIT Press.
- Pierrehumbert, J., and J. Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. *Intentions in communication*, ed. by P. Cohen, J. Morgan, and M. Pollack. Cambridge, MA: MIT Press.
- Selkirk, E. 1978. On prosodic structure and its relation to syntactic structure. *Nordic prosody II*, ed. by T. Fretheim, 111-140. Trondheim: TAPIR.
- Selkirk, E. 1980a. Prosodic domains in phonology: Sanskrit revisited. *Juncture*, ed. by M. Aronoff and M.-L. Kean. Saratoga, CA: Anma Libri.
- Selkirk, E. 1980b. The role of prosodic categories in English word stress. *Linguistic inquiry* 11.563-605.
- Selkirk, E. 1981. On the nature of phonological representation. *The cognitive representation of speech*, ed. by J. Anderson, J. Laver, and T. Meyers. Amsterdam: North Holland.
- Selkirk, E. 1984. *Phonology and syntax: the relation between sound and structure*. Cambridge, MA: MIT Press.
- Selkirk, E. 1986. On derived domains in sentence phonology. *Phonology* 3.371-405.

The Convergence of 'Similarities' and Making the Best of Probabilistic Evidence

Gwang-Yoon Goh
The Ohio State University

1. Ringe's (1992) Probabilistic Method

Ringe (1992) proposes a mathematical method of determining whether the similarities between the basic vocabularies of putatively related languages are the result of chance or not.¹ Although he claims that his method provides "a completely objective criterion of proof" of putative relationships (p.80), his approach has many problems from both linguistic and methodological standpoints some of which are so serious that they render his method partially or even totally incapable of producing accurate or meaningful results.

Above all, although his method can calculate the probability of there being a particular number of matchings of the same kind between the two relevant sounds in a comparable position, he doesn't provide a way of determining how likely a particular number of recurrent matchings (RMs) are to occur by chance. Thus, his method can only give us some sort of strong impression about language relationships. For example, 16 RMs occurring in the first positions of 70 word pairs of an English and German 100-word list might be impressive enough to make anybody believe the close relationship between English and German. However, it cannot answer what this high number (of RMs) means probabilistically or how different numbers of RMs from different pairs of languages can be compared.² This problem forces Ringe to appeal to historical arguments to explain the unexpected two RMs found between English and Turkish, making his argument rather circular (pp. 49-50).

Furthermore, Ringe's method can easily result in an undesirable conclusion, because his evidence is based on the number of RMs, whose probabilities cannot be nicely combined. For example, if closely-related languages show a smaller number of RMs than distantly-related languages, his method will give us a wrong prediction about the given relationships. Such a case is actually found in Whitman (1996), which adopts Ringe's method and applies it to three pairs of Algonquian languages: remotely-related Ojibwa and Yurok show 8 RMs, while closely-related Ojibwa and Cree on the one hand and Ojibwa and Arapaho on the other show 3 RMs and 4 RMs, respectively. Such problems can be solved by fully appreciating the wisdom of the traditional comparative method.

In short, Ringe's approach, despite its pioneering role, fails to attain its main goal of computing the chance probability for the relationship between putatively related languages. In this paper, I will propose an alternative method, which bases its probabilistic evidence mainly on the convergence of 'similarities'.

2. Making the best of probabilistic evidence

2.1. The Description of the Alternative Method

For a detailed description and explanation of the general probabilistic methods and procedures involved, which are summarized in (1), I refer the reader to Ringe (1992).

(1) General probabilistic methods and procedures involved

- (a) Compile a (Swadesh) word list for the two languages to be compared.
- (b) Choose positions for comparison.
- (c) Calculate the probability of all possible segment correspondences.
- (d) Tabulate the actual matchings.
- (e) Calculate the binomial distribution for n , a given number of trials, and p , the probability of a segment correspondence on any trial.³
- (f) Sort out RMs.

In this paper, "a set of singular facts" and "the convergence of singular facts" in the traditional comparative method (Meillet 1967: 14) are interpreted as a set of RMs and occurrences of multiple matchings [MMs] (i.e. the occurrence of more than one instance of RMs in a given word pair), respectively. Furthermore, the concept 'similarity' is defined probabilistically. A pair of sounds in a comparable position are 'similar' if their matching turns out to be an RM (i.e. if it falls in the 99th percentile of their expected range), because their correspondence is very difficult to explain unless they are assumed to be reflexes of the same sound.

The alternative method differs from Ringe's mainly in that RMs and their non-chance probabilities, which, in Ringe's method, are intended to be the main evidence for non-chance relationships, are mainly used here to identify 'similar' sounds between two compared languages. Thus, the non-chance probability of RMs, which cannot be determined against the given total word list, will be reflected in the calculation of the probability of MMs.

The further steps in the alternative method are as follows: first, using the method employed in Ringe (1992), determine what sounds are 'similar' on the basis of RMs actually found; second, determine how many MMs a given pair of languages show; third, calculate the probability of the convergence of 'similarities', on the basis of the frequencies of the sounds in each comparable position, the 'similar sounds' between two languages, and the number of MMs; finally, provide a probabilistic interpretation of the putative relationships.

2.2. The Probability of the Convergence of 'Similarities'

The probability that a MM occurs in a word pair can be calculated, as in (2). Moreover, the probability that a particular number of MMs will occur in the given 100-word list can be calculated by the formula for binomial distribution in (3).

(2) Probability [= P(n)] of an n-tuple RM in a word pair

- (a) $P(0)$ = probability that no RM occurs in any position
= $P_{12345} = (1-P_1) \times (1-P_2) \times (1-P_3) \times (1-P_4) \times (1-P_5)$
= $P_1 \times P_2 \times P_3 \times P_4 \times P_5$
 - P_n (or P_n) = probability that an RM (or no RM) occurs in n -th position
 - e.g. if the first positions show three RMs: d-t, s-s, k-k,
 $P_1 = (\#d/100 \times \#t/100) + (\#s/100 \times \#s/100) + (\#k/100 \times \#k/100)$
- (b) $P(1)$ = probability that any RM occurs in a word pair.
= $P(2) + (P_{12345} + P_{1234} + P_{12345} + P_{12345} + P_{12345}) = 1 - P(0)$
 - P_{12345} = probability that we have an RM only in the first position
= $P_1 \times (1-P_2) \times (1-P_3) \times (1-P_4) \times (1-P_5)$

- (c) **P(2)** = probability that any multiple RM occurs in a word pair.
 $= P(3) + (P_{12345} + P_{12345} + P_{12345} + P_{12345} + P_{12345} + P_{12345} + P_{12345} + P_{12345} + P_{12345} + P_{12345})$
 $= \{1 - P(0) - P''\}$, where P'' is the probability of any non-multiple RM.
 $= (1 - P_{12345} - [P_{12345} + P_{12345} + P_{12345} + P_{12345} + P_{12345} + P_{12345} + P_{12345}])$
- (d) **P(3)** = probability that any more-than-double RM occurs in a word pair.
 $= P(4) + (P_{12345} + P_{12345} + P_{12345} + P_{12345} + P_{12345} + P_{12345} + P_{12345} + P_{12345} + P_{12345} + P_{12345})$
- (e) **P(4)** = probability that any more-than-triple RM occurs in a word pair.
 $= P(5) + (P_{12345} + P_{12345} + P_{12345} + P_{12345} + P_{12345})$
- (f) **P(5)** = probability that any quintuple RM occurs in a word pair.
 $= P_{12345} = P_1 \times P_2 \times P_3 \times P_4 \times P_5$

- (3) Probability that a particular number of MMs occur in the given 100-word list

$$\frac{100!}{x! (100-x)!} \times (P)^x \times (1-P)^{100-x}$$
, where $P = P(n)$ and x = the number of MMs.

3. Investigation and Re-interpretation of the Data

3.1. English-German

The first 100-word list of English and German (cf. Ringe 1992: 83-85) shows many 'similar' sounds in each of the comparable positions, and the corresponding probabilities P_1 , P_2 , P_3 , P_4 , and P_5 (i.e. probability that any real match occurs in each chosen position in a word pair) can be given, as in (4)-(8).⁴

- (4) First position [cf. Ringe 1992: 19, 24]

similar sounds (frequencies)	number of matches
[1] s - s* (14 - 8)	5
[2] s - z (14 - 7)	6
[3] b - b (10 - 8)	5
[4] h - h (9 - 9)	6
[5] Ø - Ø (8 - 9)	8
[6] n - n (8 - 5)	5
[7] f - f (8 - 11)	8
[8] w - v (7 - 8)	4
[9] l - l (5 - 5)	4
[10] m - m (5 - 4)	3
[11] t - c (5 - 3)	3
[12] k - k (5 - 7)	3
[13] r - r (4 - 5)	3
[14] d - t (4 - 3)	2
[15] g - g (3 - 5)	3
[16] D - d (2 - 2)	2

Total: 16 pairs of similar sounds (= recurrent matchings) in 70 word-pairs

$$\bullet P_1 = (\#s/100 \times \#s/100) + \dots + (\#D/100 \times \#d/100) = 768/10000 = 0.0768$$

(5) Second position [cf. Ringe 1992: 25-6]

similar sounds (frequencies)	number of matches
[1] l - l (7 - 7)	5
[2] r - r (5 - 7)	4
[3] t - t (3 - 4)	3

Total: 3 pairs of similar sounds in 12 word-pairs

$$\bullet P_2 = (7/100 \times 7/100) + (5/100 \times 7/100) + (3/100 \times 4/100) = 96/10000 = 0.0096$$

(6) Third position [cf. Ringe 1992: 27-31]

similar sounds (frequencies)	number of matches
[1] Ø - Ø (18 - 10)	7
[2] n - n (15 - 17)	9
[3] t - s (13 - 11)	8
[4] r - r (10 - 13)	7
[5] l - l (8 - 7)	4
[6] d - t (6 - 7)	3
[7] m - m (5 - 5)	3
[8] s - z (3 - 3)	2
[9] s* - s*	3
[10] N - N (3 - 3)	3
[11] v - b (2 - 2)	2

Total: 11 pairs of similar sounds in 51 word-pairs

$$\bullet P_3 = (\#Ø/100 \times \#Ø/100) + \dots + (\#v/100 \times \#b/100) = 862/10000 = 0.0862$$

(7) Fourth position [cf. Ringe 1992: 32-3]

similar sounds (frequencies)	number of matches
[1] d - d (6 - 8)	3

Total: 1 pair of similar sounds in 3 word-pairs⁵

$$\bullet P_4 = 48/10000 = 0.0048$$

(8) Fifth position (= final syllable) [cf. Ringe 1992: 34-5]

similar sounds (frequencies)	number of matches
[1] «r «r (4 - 4)	4

Total: 1 pair of similar sounds in 4 word-pairs

$$\bullet P_5 = 16/10000 = 0.0016$$

On the basis of the probabilities for any RMs in each of the positions determined above, we can calculate the probability for convergence of similarities. For English and German, all the probabilities for all the different kinds of MMs are provided in order to show how each probability is calculated. However, $P(2)$ (= probability that any multiple match occurs in a word pair) alone will often be enough to provide the probabilistic interpretation of the putative relationship. Thus, $P(2)$, $P(3)$, $P(4)$, and $P(5)$ are given in (9), (10), (11), (12), respectively.

(9) $P(2)$ (= probability that any multiple RM occurs in a word pair)

$$\begin{aligned} P(2) &= P(3) + \text{Probability of any double RM in any positions} \\ &= 1 - \{P(0) + P''\}, \text{ where } P'' \text{ is the probability of any non-multiple RM.} \\ &= 1 - \{P_{12345} + (P_{12345} + P_{12345} + P_{12345} + P_{12345} + P_{12345})\} \\ &= 1 - 0.990935836 = 0.009064164 < 0.009 < 0.01 \end{aligned}$$

$$P_{12345} = 0.8301804843; P_{12345} = 0.0690618082; P_{12345} = 0.0080469837$$

$$P_{12345} = 0.0783120573; P_{12345} = 0.0040040859; P_{12345} = 0.0013304174$$

(10) $P(3)$ (= probability that any more-than-double RM occurs in a word pair)

$$\begin{aligned} P(3) &= P(4) + \text{Probability of any triple match in a word pair} \\ &= P(4) + (P_{12345} + P_{12345} + P_{12345} + P_{12345} + P_{12345} + \\ &\quad P_{12345} + P_{12345} + P_{12345} + P_{12345} + P_{12345}) \\ &= 0.00011586 < 0.00012 \end{aligned}$$

(11) $P(4)$ (= probability that any more-than-triple RM occurs in a word pair)

$$\begin{aligned} P(4) &= P(5) + \text{Probability of any quadruple matching} \\ &= P(5) + (P_{12345} + P_{12345} + P_{12345} + P_{12345} + P_{12345}) \\ &= 0.00000045 < 0.0000005 \end{aligned}$$

(12) $P(5)$ (= probability that we have any quintuple RM in a word pair)

$$\begin{aligned} P(5) &= P_{12345} = P_1 \times P_2 \times P_3 \times P_4 \times P_5 \\ &= 0.0000000004880911565 < 0.0000000005 (= 5^{-10}) \end{aligned}$$

The first 100-word list of English and German shows 55 MMs: 3 word pairs show quadruple RMs, 17 word-pairs show triple RMs, and 35 word pairs show double RMs (cf. Ringe 1992: 35-7). The binomial distribution for each type of MMs with $P(2)$, $P(3)$, and $P(4)$, respectively, can be computed by using the formula in (3), as in (13):

(13) The binomial distribution for each type of MMs with $P(2)$, $P(3)$, and $P(4)$

(a) Distribution for a MM of $P(2)$ [0.01]

no. matches	%	cumulative
0	36.60323413	36.60323413
1	36.97296377	73.5761979
2	18.48648188	92.06267978
3	6.099916581	98.162596361
4	1.494171486	99.656767847
5	0.2897787124	99.9465465594
6	0.04634508026	99.99289163966
7	0.006286345663	99.999177985323
8.	0.0007381693771	99.9999161547001
9.	0.0000762195092	99.9999923742093
10.	0.000007006035694	99.999999380244994
11.	0.0000005790112144	99.9999999592562084
12.	0.00000004337710276	99.99999999263331116
:	:	:

(b) Distribution for a MM of P(3) [0.00012]

no. matches	%	cumulative
0	98.80710014	98.80710014
1	1.185827501	99.992927641
2	0.007044660715	99.999972301715
3	0.00002761838421	99.9999992009921
4	0.00000008037914355	99.9999999.....
:	:	:

(c) Distribution for a MM of P(4) [0.0000005]

no. matches	%	cumulative
0	99.99500012	99.99500012
1	0.004999752506	99.999999872506
2	0.0000001237439364	99.9999999962499364
3	0.000000000002021152	99.999999996251957552
:	:	:

As we can see from the given binomial distributions, the probability of the 55 MMs occurring with the probability P(2) [0.01] is extremely small so it can exclude almost any possibility of chance. This then is precisely the probabilistic evidence for the close non-chance relationship between English and German which can replace the 'strong impression' about the closeness of the language relationship. Furthermore, this evidence is much more decisive than the probabilistic evidence (based on the number of RMs) which Ringe (1992) attempts in vain to provide.

3.2. English-Latin

The numbers of the 'similar' sounds found in the first 100 words of the Swadesh list for English and Latin are given in (14)-(17).⁶ The corresponding probabilities P1, P2, P3, P4, and P5 can be calculated in the way described in (2a), as follows:

(14) First position

- 7 pairs of similar sounds in 31 word-pairs
- $P1 = (8/100 \times 22/100) + (9/100 \times 14/100) + (14/100 \times 9/100) + (8/100 \times 8/100) + (8/100 \times 7/100) + (4/100 \times 3/100) + (5/100 \times 2/100)$
 $= 570/10000 = 0.057$

(15) Second position

- 1 pair of similar sounds in 2 word-pairs
- $P2 = 3/100 \times 2/100 = 6/10000 = 0.0006$

(16) Third position

- 2 pairs of similar sounds in 12 word-pairs
- $P3 = (10/100 \times 16/100) + (13/100 \times 10/100) = 290/10000 = 0.029$

(17) Fourth and other positions

- No matches
- $P4 = 0; P5 = 0$

On the basis of the probabilities for any RM in each position, we can calculate the probability that any MM occurs in a given word pair for English and Latin. Here, the calculation of P(2) alone is enough to provide the necessary probabilistic interpretation of the putative relationship.

(18) P(2) (= probability that any multiple RM occurs in a word pair)

$$\begin{aligned} P(2) &= P(3) + \text{Probability of any double RM in any positions} \\ &= 1 - \{P(0) + P(1)\} \\ &= 1 - \{P(12345) + (P(12345) + P(12345) + P(12345) + P(12345) + P(12345))\} \\ &= 1 - 0.998297384 = 0.001702616 < 0.002 \end{aligned}$$

The first 100 words of the Swadesh list for English and Latin show 9 MMs (cf. Ringe 1992: 47). The binomial distribution for a MM with P(2) [0.002] can be computed, as in (19).

(19) The binomial distribution for a MM with P(2) [0.002]

no. matches	%	cumulative
0	81.8566805	81.8566805
1	16.40414438	98.26082488
2	01.627264824	99.888089704
3	0.106527691	99.994617395
4	0.005176947	99.999794342
5	0.000199193147	99.999993535147
6	0.000006320424	99.99999855571
:	:	:

The probabilistic interpretation of 9 MMs in the English and Latin word is that the occurrence of the given number of MMs is extremely difficult to explain by chance. Even though the non-chance probability for those 9 MMs is not so big as the one for the 55 MMs from English and German, it is still big enough to exclude almost any possibility of chance. Such a probability can be said to reflect our strong impression about the closeness of the non-chance relationship between English and Latin as well as the difference we feel between the relationships of the two pairs of languages (i.e. English and German, on the one hand, and English and Latin, on the other).

3.3. English-Turkish

The first 100-word list of English and Turkish (cf. Ringe 1992: 86-89) shows several 'similar' sounds in some of the comparable positions (i.e. the third and fourth positions). They are given in (20) through (23), along with the corresponding probabilities P1, P2, P3, P4, and P5.

(20) First position (cf. Ringe 1992: 14, 48-9)

- 2 pairs of similar sounds in 8 word-pairs: 6 [b-k] (10-17); 2 [y-s] (2-6)
- $P1 = (10/100 \times 17/100) + (2/100 \times 6/100) = 182/10000$

(21) Second position

- No similar sounds ⁷
- **P2 = 0**

The sounds in the third position and their frequencies are given in (22a), the expected chance matchings for all possible pairs are given in table 1 in (22b), and the numbers of matchings actually found are given in table 2 in (22c), as follows:

(22) Third position

(a) Frequencies of consonants

English		Turkish	
Ø	18	r	15
n	15	l	13
t	13	Ø, n, m	9 each
r	10	g*	8
l	8	t	7
d	6	s*, z	6 each
m	5	k	4
k	4	c*, p	3 each
Q, s, s*, g, N	3 each	j*, d	2 each
v	2	b, v, s, h	1 each
p, f, D, z	1 each		
total	100	total	100

(b) Expected chance matchings in the third consonants, English-Turkish

Tk Eg	r (15)	l (13)	Ø, n, m (9 each)	g* (8)	t (7)	s*, z (6 each)	k (4)	c*, p (3 each)	j*, d (2 each)	b, v, s, h (1 each)
Ø (18)	2.7	2.34	1.62	1.44	1.26	1.08	0.72	0.54	0.36	0.18
n (15)	2.25	1.95	1.35	1.2	1.05	0.9	0.6	0.45	0.3	0.15
t (13)	1.95	1.69	1.17	1.04	0.91	0.78	0.52	0.39	0.26	0.13
r (10)	1.5	1.3	0.9	0.8	0.7	0.6	0.4	0.3	0.2	0.1
l (8)	1.2	1.04	0.72	0.64	0.56	0.48	0.32	0.24	0.14	0.08
d (6)	0.9	0.78	0.54	0.48	0.42	0.36	0.24	0.18	0.12	0.06
m (5)	0.75	0.65	0.45	0.4	0.35	0.3	0.2	0.15	0.1	0.05
k (4)	0.6	0.52	0.36	0.32	0.28	0.24	0.16	0.12	0.08	0.04
Q (3)	0.45	0.39	0.27	0.24	0.21	0.18	0.12	0.09	0.06	0.03
s (3)	0.45	0.39	0.27	0.24	0.21	0.18	0.12	0.09	0.06	0.03
s* (3)	0.45	0.39	0.27	0.24	0.21	0.18	0.12	0.09	0.06	0.03
g (3)	0.45	0.39	0.27	0.24	0.21	0.18	0.12	0.09	0.06	0.03
N (3)	0.45	0.39	0.27	0.24	0.21	0.18	0.12	0.09	0.06	0.03
v (2)	0.3	0.26	0.18	0.16	0.14	0.12	0.08	0.06	0.04	0.02
p (1)	0.15	0.13	0.09	0.08	0.07	0.06	0.04	0.03	0.02	0.01
f (1)	0.15	0.13	0.09	0.08	0.07	0.06	0.04	0.03	0.02	0.01
D (1)	0.15	0.13	0.09	0.08	0.07	0.06	0.04	0.03	0.02	0.01
z (1)	0.15	0.13	0.09	0.08	0.07	0.06	0.04	0.03	0.02	0.01

Table 1

(c) Numbers found for matchings of the third consonants, English-Turkish ⁸

Tk Eg	r (15)	l (13)	Ø (9)	n (9)	m (9)	g* (8)	t (7)	s* (6)	z (6)	k (4)	c* (3)	p (3)	j* (2)	d (2)	b, v, s, h (1 each)
Ø (18)	3	2		3	2	1	2		3	1	1				
n (15)	4	1	1	1	2	2		2		1					1 (v)
t (13)			3	1	1	2	1			2			[2]		1 (s)
r (10)	1	2		2				2			1	1			1 (b)
l (8)	3	2				1	1				1				
d (6)		1	1	1				1	1						1 (h)
m (5)		1					1		1					[2]	
k (4)	2		1		1										
Q (3)		1				1		1							
s (3)				1	1		1								
s* (3)		[2]					1								
g (3)			1		1							1			
N (3)		1			1				1						
v (2)	1					1									
p (1)			1												
f (1)												1			
D (1)			1												
z (1)	1														

Table 2

(d) Similar sounds based on RMs

similar sounds (frequencies) number of matches

[1] t - j* (13 - 2) 2

[2] m - d (5 - 2) 2

[3] s* - l (3 - 13) 2

Total: 3 pairs of similar sounds in 6 word-pairs

$$\bullet \text{ P3} = (13/100 \times 2/100) + (5/100 \times 2/100) + (3/100 \times 13/100) = 75/10000$$

As for the sounds in the fourth position, their frequencies, the expected chance matchings for all possible pairs, and the numbers of matchings actually found are given in (23), as follows:

(23) Fourth and other positions > no RM > no similar sounds

$$\bullet \text{ P4} = 0; \quad \text{P5} = 0$$

(a) Frequencies of consonants

English		Turkish	
Ø	86	Ø	79
d	6	m	12
t	3	n	3
n, k	2 each	d, r	2 each
Q	1	z, k	1 each
-----		-----	
total	100	total	100

(b) Expected chance matchings for the fourth consonants, English-Turkish

Eg	Tk	Ø (79)	m (12)	n (3)	d, r (2 each)	z, k (1 each)
Ø (86)		67.94	10.32	2.58	1.72	0.86
d (6)		4.74	0.72	0.18	0.12	0.06
t (3)		2.37	0.36	0.09	0.06	0.03
n (2)		1.4	0.24	0.06	0.04	0.02
k (2)		1.4	0.24	0.06	0.04	0.02
Q (1)		0.79	0.12	0.03	0.02	0.01

Table 3

(c) Numbers found for matchings of the fourth consonants, English-Turkish

Eg	Tk	Ø (79)	m (12)	n (3)	d (2)	r (2)	z (1)	k (1)
Ø (86)		69	9	3	2	1	1	1
d (6)		5	1					
t (3)		3						
n (2)		1						
k (2)		1						
Q (1)						1		

Table 4

On the basis of the probabilities for any real matches (i.e. P1, P2 and etc.) in the first 100 words of the Swadesh list for English and Turkish, we can calculate the probability that any MM occurs in a given word pair, as in (24) below:

(24) P(2) (= probability that any multiple RM occurs in a word pair)

P(2) = P(3) + Probability of any double RM in any positions

= 1 - {P(0) + P"} }

= 1 - {P₁₂₃₄₅ + (P₁₂₃₄₅ + P₁₂₃₄₅ + P₁₂₃₄₅ + P₁₂₃₄₅ + P₁₂₃₄₅)}

= 0.0001365 < 0.00014

P₁₂₃₄₅ = 0.9744365; P₁₂₃₄₅ = 0.0180635; P₁₂₃₄₅ = 0 (< P2=0)

P₁₂₃₄₅ = 0.0073635; P₁₂₃₄₅ = 0 (< P4=0); P₁₂₃₄₅ = 0 (< P5=0)

No MM is found in the first 100 words of the Swadesh list for English and Turkish and this can be verified by the fact that every pair of similar sounds found in the list occurs in a different word pair, as in (25).

(25) Word pairs which show similar sounds

(a) First position (2 pairs of similar sounds in 8 word-pairs)

word pair	meaning	matching	word pair	meaning	matching
bark - kabuk	'bark'	(b - k)	blood - kan	'blood'	(b - k)
belly - karın	'belly'	(b - k)	bone - kemik	'bone'	(b - k)
bird - kus	'bird'	(b - k)	yellow - sarı	'yellow'	(y - s)
black - kara	'black'	(b - k)	you - sen	'you'	(y - s)

(b) Second position (3 pairs of similar sounds in 6 word-pairs)

word pair	meaning	matching	word pair	meaning	matching
night - gece	'night'	(t - j*)	fish - balık	'fish'	(s* - l)
hot - sıcak	'hot'	(t - j*)	ash - kül	'ash'	(s* - l)
human - adam	'human'	(m - d)			
woman - kadın	'woman'	(m - d)			

The binomial distribution for a MM with P(2) [0.00014] is given, as in (26):

(26) Binomial distribution for a MM with P(2) [0.00014]

no. matches	%	cumulative
0	98.60965778	98.60965778
1	1.380728511	99.990386291
2	0.009569788351	99.999956079351
3	0.00004377196013	99.99999985131113
:	:	:

What this binomial distribution in (26) means is that the occurrence of one MM is more than 98.6 % non-chance, which means that out of 1000 hundred-word lists (or out of 100,000 word pairs) we can expect one MM in 14 lists. This again means that although finding a MM or two in some of the lists is possible and expected, it will still be very difficult. This is compatible with the fact that no MM was found in the given 100-word list.

In addition, the given probabilistic interpretation supports our expectation about the putative relationship between English and Turkish based on the comparative method. Thus, unlike Ringe (1992), we don't have to appeal to any extra-probabilistic arguments such as historical arguments for English and Turkish.

4. Conclusion

In this paper, I have proposed an alternative probabilistic method for determining the (non-)chance relationship between putatively related languages. In particular, the probabilistic evidence (especially, P(2) and the corresponding binomial distribution) based on MMs representing the convergence of similarities has been used to provide a better probabilistic prediction and to deal with problems remained unsolved in Ringe (1992).

In short, supporting and debunking claims or hypotheses about language relationships should be one of the main goals of the probabilistic methods. Considering the fact that the probabilistic method is rarely necessary for closely related languages such as English and German, the further demonstration of validity of the current method as a sifting device in more difficult cases is a pressing need.

Notes

¹ He applies his method to four different pairs of languages, i.e. English-German, English-Latin, English-Turkish, and English-Navajo (Ringe 1992). He also uses his method to test the controversial Nostratic hypothesis (Ringe 1995) and to debunk Greenburg's multilateral comparison of 'Amerind' family (Ringe 1996).

² For a more detailed discussion of this problem, see Baxter and Ramer (1996).

³ Probability (= P') that there will be k matches in n random trials, for any number k:

$$\bullet P' = [n! / k! \times (n - k)!] \times P^k \times (1 - P)^{n-k} \quad (\text{cf. Paulos 1988})$$

n = the total number of trials (i.e. the number of word pairs in a given list).

P = probability of a segment correspondence (i.e. probability that two sounds will match) in a comparable position on any trial.

⁴ The comparable compositions are determined, as in Ringe (1992): first position = initial consonants; second position = second consonants of the initial clusters; third position = consonants right after the first-syllable vowel nucleus; fourth position = second consonants after the first-syllable vowel nucleus; fifth position = final syllables.

⁵ This number of matches falls just below the 99th percentile. See Ringe (1992: 33).

⁶ For the first 100-word list, see Ringe (1992: 83-85). As for the 'similar' sounds found in each comparable position, on the other hand, refer to Ringe (1992: 41, 14, 44-47).

⁷ No RMs are found since there are no initial clusters in Turkish.

⁸ The numbers in brackets are the matchings (i.e. RMs) which cross the 99th percentile threshold.

References

- Baxter, W. H. & Ramer, A. M. 1996. Review of *On Calculating the Factor of Chance in Language Comparison* by D. A. Ringe, Jr. 1992. *Diachronica* 13: 2. 371-384.
- Hony, H. C. 1957. *An Turkish-English Dictionary*. Oxford: Clarendon.
- Iz, F. & Hony, H. C. 1952. *An English-Turkish Dictionary*. Oxford: Clarendon.
- Meillet, A. 1925/1967. *The Comparative Method in Historical Linguistics* (trans. G. B. Ford, Jr.). Paris: Labrarie Honoré Champion.
- Paulos, J. A. 1988. *Innumeracy: Mathematical Illiteracy and Its Consequences*. New York: Hill and Wang.
- Ringe, D. A., Jr. 1992. *On Calculating the Factor of Chance in Language Comparison*. Transactions of the American Philosophical Society, vol. 82, Part 1.
- _____. 1995. Nostratic and the Factor of Chance. *Diachronica* 12: 1. 55-74.
- _____. 1996. The Mathematics of Amerind. *Diachronica* 13: 1. 135-154.
- Whitman, N. 1996. A Probabilistic Comparison of Four Algonquian Languages: A Follow-up to Ringe. unpublished manuscript. The Ohio State University.

The Phonetics and Phonology of Non-modal Vowels: A Cross-Linguistic Perspective

Matthew Gordon
University of California, Los Angeles

1. INTRODUCTION.* Some languages of the world have vowels characterized by non-modal phonation, e.g. breathy voiced vowels, voiceless vowels, or creaky (laryngealized) vowels. Depending on the language and on the phonation type, these non-modal vowels may either contrast with or be allophonic variants of modal voiced vowels. For example, creaky vowels are phonemic in Kedang (Samely 1991) and Jalapa Mazatec (Kirk et al. 1993), but occur allophonically in the vicinity of glottalized consonants in many languages, e.g. Georgian and Tzeltal (cf. Crothers et al. 1979). Breathly voiced vowels are phonemic in Gujarati (Fischer-Jørgensen 1967) but occur allophonically in the vicinity of /h/ in many languages. Similarly, voiceless vowels contrast with voiced vowels on the surface in Turkana (Dimmendaal 1983), but occur allophonically in Japanese (Han 1961).

Non-modal vowels have a quite different distribution from modal vowels. First, they are quite rare cross-linguistically, both as phonemic segments which contrast with modal vowels, and as non-contrastive allophones of modal voiced vowels. For example, Maddieson's (1984) survey of 317 languages includes only two with contrastive laryngealized/creaky voice (Sedang and Southern Nambiquara), two with contrastive voiceless vowels (Ik and Dafla), and one with phonemic breathy voiced vowels (Tamang). Another characteristic property of non-modal vowels which differentiates them from modal vowels is their limited distribution. For example, voiceless vowels are often limited to word-final position, and creaky vowels tend to occur adjacent to glottalized consonants. In other languages, non-modal vowels are the synchronic manifestations of other types of contrasts, e.g. segmental, tonal, or durational ones.

Given the limited distribution of non-modal vowels relative to modal vowels there are a couple of basic questions which come to mind. First, why do non-modal vowels typically play a limited role in the phonology of most languages? Second, is the distribution of non-modal vowels predictable on phonetic grounds? This goal of this paper is to provide answers to these questions and to formalize these answers in an Optimality Theoretic grammar.

2. THE RARITY OF NON-MODAL VOWELS CROSS-LINGUISTICALLY. I conjecture that the rarity of contrastive non-modal vowels has a perceptual basis; non-modal vowels are perceptually less robust than modal vowels and are therefore eschewed by many languages. It has been shown by Silverman (1995, 1998) that, non-modal phonation reduces the ability of vowels to manifest tonal contrasts in a salient manner. Given this fact, it is thus not surprising that many tone languages (e.g. Jalapa Mazatec) restrict overlap between tonal and phonation contrasts.

However, beyond the inherent incompatibility between non-modal voicing and tone discussed by Silverman, there is reason to believe that non-modal vowels share properties which make them *inherently* less salient than modal vowels, even in the absence of tonal contrasts. This reduced salience can inhibit the recovery of contrastive place information in the vowel. Let us now consider the acoustic properties which make non-modal vowels less salient than modal vowels.

First, non-modal vowels are characterized by less overall acoustic intensity than modal vowels as shown for different non-modal phonation types, e.g. breathy vowels in Kui and Chong (Thongkum 1987), creaky vowels in Chong (Thongkum

1987) and voiceless and creaky vowels in Hupa (Gordon 1998). Decreasing the acoustic intensity results in a decrease in loudness, the auditory correlate of intensity; it thus follows that non-modal vowels are less salient than modal vowels.

Furthermore, non-modal voicing often alters the spectral properties of vowels, including formant structure, as demonstrated instrumentally for Kedang (Samely 1991) and Chong (Thongkum 1987), and also qualitatively evident in many languages, e.g. those in which vowels in the vicinity of glottal stop and /h/ have noticeably different qualities than vowels in other environments not associated with non-modal phonation (cf. Blankenship 1997 for extensive discussion of the spectral properties of non-modal vowels). The perturbation of formant structure potentially makes recovery of vowel quality contrasts more difficult. In summary, given the reduced intensity of non-modal phonation and its influence on vowel quality, it would thus not be surprising that many languages design their phonologies to limit the distribution and the role of non-modal vowels.

In addition to the general paucity of contrastive non-modal vowels cross-linguistically, other language specific facts suggest that phonologies avoid non-modal vowels for perceptual reasons. First, breathy voiced vowels in Kedang (Samely 1991), and both breathy and creaky vowels in Jalapa Mazatec (Kirk et al. 1993) are phonetically much longer (up to 50% longer) than their modal voiced counterparts. An interesting aspect of Jalapa Mazatec is that non-modal voicing does not persist throughout the entire duration of phonemic breathy and glottalized vowels. Rather, non-modal voicing occurs principally on the first half of non-modal vowels; the second half of non-modal vowels is phonetically characterized by modal voicing. Acoustic measurements of fundamental frequency and formants suggest that non-modal voicing is also largely confined to the first half of the vowel in Kedang. Crucially, Kedang is not a tone language; thus, the realization of non-modal voicing cannot be attributed to the presence of tonal contrasts. Rather, *segmental* properties of the vowel are driving the phonetic realization of non-modal voicing in Kedang. It thus seems to be the case that both Jalapa Mazatec and Kedang are sensitive to a constraint requiring that at least some portion of the duration of a non-modal vowel be characterized by modal voicing. This requirement is sensible in light of the reduced salience of non-modal voicing for reasons mentioned earlier. Leaving a modal voiced portion enhances the salience of a non-modal vowel.

Laryngeal timing patterns in Hupa provide further evidence that languages are sensitive to the reduced salience of non-modal vowels. In brief (for detailed discussion, see Golla 1970, Gordon 1998), laryngeal features underlying associated with preconsonantal obstruents spread onto a preceding vowel in Hupa. Vowels preceding preconsonantal ejectives, i.e. constricted glottis consonants, are realized with creak, the acoustic manifestation of constricted glottis in vowels; vowels preceding preconsonantal voiceless obstruents are realized as voiceless vowels. The crucial facts for the present discussion are as follows. First, laryngeal features do not spread onto short vowels. Second, laryngeal features only spread onto the last half of a preceding *long* vowel. Thus, it is never the case that a vowel is obscured by non-modal voicing for its entire duration. Thus, the Hupa pattern of laryngeal spreading is governed by the same restrictions governing the realization of underlying non-modal vowels in Kedang and Jalapa Mazatec. Realizing non-modal voicing on the last half of a long vowel in Hupa still leaves a portion of modal voicing from which place information may be easily recovered. Laryngeal features cannot spread onto a short vowel, since this would completely obscure the vowel. Crucially, as in Kedang, there are no tonal contrasts present in Hupa which would

block laryngeal spreading in non-modal voicing; it is thus the desire to realize place information saliently which is driving the data.

The durational patterns of non-modal voicing can be modeled in a constraint-based grammar using a few constraints. First, there are two constraints against non-modal vowels, one prohibiting non-modal short vowels (*NON-MODAL SHORT V), the other against non-modal long vowels (*NON-MODAL LONG V). Let us assume that the constraint against non-modal long vowels is violated once for each half of the vowel which is non-modal, i.e. one violation for each timing position associated with non-modal voicing. Thus, a fully non-modal long vowel violates this constraint twice, while a partially non-modal long vowel violates it once. In Hupa, we must also assume a constraint which forces non-modal voicing to spread from a preconsonantal consonant onto the preceding vowel. The relevant constraint is motivated by the requirement that laryngeal features of a consonant not be completely overlapped by the consonant constriction (see Gordon 1998 for more discussion). Here I will simply formulate the constraint as *SPREAD LARYNGEAL F; this constraint requires that laryngeal features (creak or voicelessness depending on the consonant) spread from preconsonantal obstruents onto an adjacent vowel. By ranking *SPREAD LARYNGEAL F below *NON-MODAL SHORT V but above *NON-MODAL LONG V, we get the Hupa facts.

We can also account for the Kedang and Jalapa Mazatec patterns if we assume that non-modal vowels in these languages are underlyingly linked to one timing position reflecting their phonemic quantity as short vowels, but two timing positions on the surface, reflecting their substantially longer surface duration. By ranking the constraint against insertion of timing positions not present underlyingly, phrased here as *DEP-X, following McCarthy and Prince's (1995) Correspondence Theory, below *NON-MODAL SHORT V but above *NON-MODAL LONG V, we account for the fact that, in Kedang and Jalapa Mazatec, non-modal vowels are phonetically quite long, but non-modal for only portion of the vowel.

3. VOICELESS VOWELS AND THE ROLE OF ARTICULATORY FACTORS. Thus far, I have provided a perceptually-driven explanation for why non-modal vowels have such a limited distribution cross-linguistically in comparison to modal voiced vowels. This account makes the prediction, borne out in the data presented thus far, that, certain languages will disprefer non-modal voicing on short vowels.

Interestingly, as it turns out, there are many languages which devoice short but not long vowels, a pattern which runs opposite to the predictions made by the perceptually driven explanation offered in the previous section. The presence of both patterns cross-linguistically, devoicing of short but not long vowels, and devoicing of long but not short vowels, deserves explanation. In sections 3.2-3.5, I will address the asymmetries related to vowel length, as well as other asymmetries gleaned from a typology of approximately 50 languages (see the Appendix). Approximately half of the typology is drawn from Crother et al.'s database (1979), while most of the remaining languages are mentioned or discussed in either Cho (1993), Vine (1981), or Jun et al. (1997).

3.1. THE STATUS OF NON-MODAL VOICING: PHONETIC OR PHONOLOGICAL. Before preceding with the typology, it is appropriate to address the question of whether non-modal voicing is a phonological or phonetic phenomenon or perhaps both, depending on the language. Clearly in languages with an underlying contrast between non-modal voiced vowels and modal voiced vowels (e.g. Sedang, Gujarati, Jalapa Mazatec), non-modal vowels are a synchronic phonological

feature¹. However, as pointed out earlier, the number of languages with underlying or even surface contrastive phonation type for vowels is quite small. In the majority of languages in which they occur, non-modal vowels are a surface non-contrastive property, and thus less clearly belong to the phonology.

The issue of the phonological vs. phonetic status of non-modal vowels has been most thorough investigated for voiceless vowels, e.g. in relatively recent work by Vine (1981), Cho (1993), Tsuchida (1994), Jun and Beckman (1993), Jun et al. (1997). In the majority of languages for which vowel devoicing has been the subject of intensive acoustic analysis, vowel devoicing appears to be a gradient rather than a categorical phenomenon; languages falling into this category include Japanese (Han 1961, Beckman 1982, Tsuchida 1994), Montreal French (Gendron 1966, Cedergren and Simoneau 1985), Greek (Dauer 1980), Turkish (Jannedy 1995) and Korean (Jun and Beckman 1993, 1994, Jun et al. 1997, 1998)². In these languages, vowel devoicing operates on a continuum with token to token variation in the presence or absence or degree of devoicing. On one end of the continuum is a voiced vowel, at the other extreme is vowel deletion; various degrees of devoicing fall in between these two extremes. The likelihood of devoicing is a function of various phonetic factors: position of stress/accent, distance from prosodic boundaries, vowel height, surrounding consonants, and speech rate. The gradient nature of vowel devoicing is even suggested in many grammars which describe devoicing as optional but not required in a given environment (e.g. Tongan, Acoma, Tubu, Boraana Oromo, Kawaiisu, Big Valley Shoshoni, Mokilese, Cocama) or in languages where the span of devoicing can vary in length (Acoma, Southern Paiute). It is possible that instrumental work would demonstrate that, in a great many, perhaps most, languages, devoicing is a gradient phenomenon.

On the other hand, there are many languages in which vowel devoicing behaves like a phonological phenomenon. In some languages, voiceless vowels contrast on the surface with voiced vowels. For example, the word-internal contrast between short and long vowels is realized as a contrast between voiceless and voiced vowels in word-final position in Oromo and in Woleaian. Similarly, in Hupa as discussed in section 2, vowel devoicing is contrastive before many syllable-final consonants. If we adopt the standard assumption that contrastive properties are phonological, vowel devoicing would clearly fall under the purview of phonology in Oromo, Woleaian and Hupa. Furthermore, in other languages, vowel devoicing interacts with other phenomena which are typically assumed to be phonological. For example, vowel devoicing influences stress assignment in Awadhi, pitch accent placement in Tunica, and debuccalization and tone shift in Comanche. Furthermore, in Tongan, one of the prerequisites for vowel devoicing is that vowels be in morpheme-final position; such morphological conditioning would suggest that vowel devoicing is not merely a low level phonetic phenomenon. In summary, vowel devoicing thus appears to operate at a relatively deep level of the grammar in a fair number of languages³.

Perhaps not surprisingly, languages in which vowel devoicing clearly appears to be phonological display vowel devoicing in the same environments (e.g. domain finally, adjacent to voiceless consonants) in which vowel devoicing is most likely to occur in languages where it has been demonstrated to be gradient. Even if one assumes a sharp distinction between vowel devoicing as a phonetic process vs. devoicing as a phonological one, the striking similarity between the distributions of phonetic and phonological vowel devoicing suggests that examination of phonetic devoicing may also provide insight into phonological devoicing. For this reason, the typology in this paper includes cases of vowel devoicing which are clearly

phonological as well as others which may not be. This paper will not address the issue of where to draw the line between phonological and phonetic processes. Crucially, because substantially the same phonetic factors condition devoicing in all languages, many aspects of the analysis of devoicing are more likely than not to be quite similar for all languages with voiceless vowels.

3.2. THE LENGTH ASYMMETRY. Of the 32 languages in the survey with contrastive vowel length in environments targeted by devoicing, devoice short but not long vowels occurring in the same environment (in many languages, only high vowels devoice; see section 3.4): e.g. Awadhi, Big Smokey Valley Shoshoni, Bulu, Mbay, Cheyenne, Cocama, Gadsup, Galla, Ik, Inuit, Oneida, Goajiro, Tarascan, Zuni, Japanese, Kawaiisu, Mokilese, Sámi, Sara, Shina, Bagirmi, Tongan, Tubu, Tunica, Turkish, Woleaian). Four languages (Boraana Oromo, Papago, Southern Paiute, Ket⁴) possess voiceless short vowels and also devoice a portion of long vowels in certain environments⁵, one (Acoma) possesses voiceless short vowels and long vowels which are completely voiceless, and one language (Hupa) has partially voiceless long vowels but lacks voiceless short vowels. In one, Cheyenne, long vowels partially devoice in final position but do not devoice at all in word-medial environments in which short vowels devoice.

3.3. ENVIRONMENT OF DEVOICING. In virtually all languages in the survey, vowel devoicing is found at least in final position; in many languages, devoicing also occurs in other environments as well. Crucially, the occurrence of devoicing in non-final environments almost always implies devoicing in final position. The languages I know of which are exceptional in this regard are Inuit (Crothers et al. 1979), Quechua (Crothers et al.), Turkish (Jannedy 1995), Azerbaijani (Crothers et al.) and Montreal French (Gendron 1966, Cedergren and Simoneau 1985); vowels in these languages resist devoicing in final position, but allow it in other environments. However, in three of these languages (Turkish, Azerbaijani and Montreal French), final vowels are stressed, thereby explaining their failure to devoice (see section 3.5). In Inuit, phrase final position is typically associated with a high tone which also often blocks devoicing cross-linguistically (see section 3.5).

Up to this point, I have been intentionally vague in defining "final position". The reason for this is that the domain of devoicing varies from language to language; however, these domains follow an implicational hierarchy. Devoicing in final position of a smaller domain (e.g. word) implies devoicing in larger domains (e.g. phrase, utterance); the reverse of this statement is not necessarily true. In 20 languages in the survey (Ik, Dafla, Cocama Galla, Bagirmi, Turkana, Sara, Tubu, Mbay, Malagasy, Campa, Tarascan⁶, Ticuna, Ket, Ainu⁷, Island Carib, Zuni, Washkuk, Goajiro, Woleaian⁸), voiceless vowels occur word-finally (and of course, by implication, finally in larger domains as well). In 14 languages, devoicing is characteristic only of final position of larger domains, e.g. phrase or utterance (Alabama, Papago, Greek, Tarascan, Totonac, Chontal, Gadsup, Oneida, Apinaye⁹, Mixtec, Nyangumata¹⁰, Boraana Oromo, Cheyenne, Kawaiisu¹¹). Note that, from most descriptions, it is impossible to make distinctions among larger domains such as the phonological phrase, intonational phrase or utterance.¹² Interestingly, only one language, Cocama, regularly devoices initial vowels in addition to final vowels; in both environments, devoicing only affects vowels adjacent to a voiceless consonant.

After final position, the next most common position in which vowels devoice is adjacent to voiceless consonants. Word-medial devoicing is found in 19 languages

in the survey (Mandarin, Brazilian Portuguese, Malagasy, Mixtec, Quechua, Goajiro, Azerbaijani, Inuit, Chontal, Montreal French, Cheyenne, Mokilese, Big Valley Shoshoni, Japanese, Turkish, Korean, Tongan, Cocama, Papago). It is interesting to note that devoicing of final vowels in most languages (29 of 36), occurs not only after voiceless but also after voiced consonants¹³. In only 6 languages with final devoicing (Japanese, Korean, Tongan, Turkana, Cocama, Mixtec) must the vowel both be final *and* next to a voiceless consonant for devoicing to occur¹⁴. In 8 languages with word-medial devoicing it is sufficient to have a voiceless consonant on only one side of a vowel to trigger devoicing. In 5 of these languages, the triggering consonant is on the right side of the vowel (Big Valley Shoshoni, Comanche, Southern Paiute, Goajiro, Quechua), in 3, it is on the left side (Acoma, Mandarin, Chontal). In 9 languages with word-medial devoicing (Cheyenne, Mokilese, Japanese, Turkish, Korean, Montreal French, Tongan, Papago, Malagasy), devoicing is described as affecting vowels (almost) exclusively between two voiceless consonants. Further asymmetries between different voiceless consonants will be discussed in section 4.

3.4. THE HEIGHT ASYMMETRY. Vowel devoicing is also sensitive to vowel height; in many languages, high voiceless vowels but not mid and low voiceless vowels occur (Greek, Korean, Turkish, Dafla, Montreal French, Mokilese, Brazilian Portuguese, Mandarin, Campa, Mixtec, Ainu, Azerbaijani, Gadsup, Inuit, Ticuna).¹⁵ Similarly, in Tongan, the set of environments in which non-high vowels devoice is a subset of the environments in which high vowels devoice.¹⁶ I know of no language which devoices non-high vowels but not high vowels.

3.5. THE STRESS AND TONE/INTONATION ASYMMETRY. Two other asymmetries in vowel devoicing relate to the closely related properties of tone and accent. In all languages in the survey for which data on accent location is reported and in which the other necessary preconditions for devoicing are present, accented vowels resist devoicing (Montreal French, Turkish, Tongan, Comanche¹⁷, Cheyenne, Brazilian Portuguese, Azerbaijani, Quechua). In keeping with this pattern, in Papago, the set of environments in which stressed vowels devoice is a subset of those in which unstressed vowels devoice. I know of no language with devoicing of stressed but not unstressed vowels.

A final asymmetry is that many tone and pitch accent languages fail to devoice high-toned vowels (Japanese¹⁸, Cheyenne, Acoma¹⁹). Furthermore, in some languages with stress, intonational pitch accents (Greek, Boraana Oromo, Tunica) and high boundary tones (Inuit) can inhibit devoicing. The stress and tone/intonational asymmetries are presumably closely related since accented syllables often carry high pitch accents cross-linguistically, as in Japanese.

4. AN ARTICULATORY ACCOUNT OF VOWEL DEVOICING. The asymmetries in vowel devoicing discussed can, in large part, be explained in terms of a combination of articulatory overlap between neighboring glottal gestures and aerodynamic considerations. First, let us consider the patterns which are compatible with a gestural overlap account of devoicing. The reasoning given here basically follows that of Dauer (1980), Jun and Beckman (1993, 1994), Jun et al. (1997, 1998). The types of vowels which devoice cross-linguistically are those which are likeliest to be produced with voicing gestures which are durationally shortest. Trivially, phonemic short vowels are phonetically shorter than phonemic long vowels. It is also well known that high vowels are shorter than non-high vowels (Lehiste 1970) and that unaccented vowels are shorter than accented ones.

Because of their shorter duration, the glottal adduction gestures associated with phonemic short vowels, unstressed vowels, and high vowels are more likely to be overlapped by the glottal gestures of neighboring segments. When the neighboring gestures are abduction gestures, as in the case of voiceless consonants, they threaten to overlap the adduction gestures for voicing of the vowel. When sufficient overlap occurs, vowel devoicing results. The overlap account makes the prediction that devoicing is more likely to occur when a vowel is surrounded on both sides by voiceless consonants. This prediction is borne out by a number of instrumental studies, e.g. Jun and Beckman (1994), Jun et al. (1997, 1998) on Korean, Han (1961) on Japanese, Jannedy (1995) on Turkish, Dauer (1980) on Greek. It also is supported by the fact that devoicing in many languages is only triggered when a vowel is surrounded by voiceless consonants. The overlap account also predicts that devoicing is most likely to occur in the vicinity of voiceless consonants with the greatest glottal abduction gestures and in the vicinity of voiceless consonants whose glottal abduction peaks are timed to occur near the vowel. In general, this prediction is also borne out. Voiceless consonants with the largest glottal openings, (fricatives - Löfqvist and Yoshioka 1980), and those whose peak glottal abductions fall close to a vowel (aspirated stops - Kagaya 1974, Pétursson 1976), tend to trigger devoicing most. Thus, in languages with unaspirated stops, fricatives are most likely to trigger devoicing than stops; e.g. fricatives but not stops trigger devoicing in Comanche. In Mokilese, devoicing is most likely next to an /s/. In Goajiro, devoicing of vowels occurs before voiceless fricatives and affricates, which presumably also often have relatively large glottal openings (cf. Kagaya). In Southern Paiute word-medial devoicing is triggered by a following fricative or geminate stop; geminates have been shown to have greater glottal apertures than singletons (Pétursson 1976). In Turkish, devoicing is more likely in the neighborhood of phonologically unaspirated stops than fricatives; however, phonetically, as Jannedy (1995) points out, the "unaspirated" stops of Turkish are characterized by substantial aspiration which perhaps accounts for the preferential devoicing of vowels following stops. This hypothesized link between aspiration duration and likelihood of devoicing is compatible with the fact that /k/, the stop with the longest aspiration duration cross-linguistically, is the only consonant to trigger devoicing in Tunica. However, aspiration duration is not the entire story, as Jannedy points out, since a preceding /p/ is more likely to trigger devoicing than /k/ in Turkish even though /p/'s aspiration duration is shorter than /k/'s.

Certain languages show place asymmetries in the set of fricatives which trigger devoicing; e.g. /h/ but not /s/ or /f/ triggers devoicing of /a/ in Tongan, /s/ is much more likely than /f, θ, x/ to trigger devoicing of high vowels in Greek, /h/ but not /s/ triggers devoicing of short vowels in Big Smokey Valley Shoshoni. In Mandarin, devoicing is most common after aspirated affricates and after voiceless fricatives *other than uvulars*. Although we lack the relevant articulatory data on these languages, it is a reasonable hypothesis that these language specific differences in the likelihood of devoicing near certain fricatives may be due to language specific differences in the relative width (and perhaps timing) of glottal abduction gestures of different fricatives: fricative(s) with greater glottal openings in a given language are more likely to trigger devoicing in that language.

Interestingly, in Korean, the asymmetry between stops and fricatives depends on whether the stop or fricative appears on the right or left side of the potential target of vowel devoicing (Jun and Beckman 1994, Jun et al. 1997, 1998). Fricatives and aspirated stops are more likely to trigger devoicing than fortis and

lenis stops when they precede a vowel, an expected pattern given the greater glottal apertures of fricatives and aspirated stops. However, following a vowel, all stops, including fortis and lenis stops trigger devoicing more often than fricatives, even though fricatives have greater glottal openings than fortis and lenis stops. Jun and Beckman suggest that the closing gesture into a stop might be faster than the oral constriction gesture made for a fricative; this greater velocity of the oral closing gesture could lead to a more abrupt increase in oral pressure which could inhibit voicing in the preceding vowel. One might hypothesize that the asymmetry between Korean and those languages in which devoicing is triggered by a following fricative and not a following stop (Comanche, Goajiro, and Southern Paiute--singletons) is due to language specific differences in the relative magnitude and timing of glottal opening gestures in the two classes of consonants.

In summary, although the glottal overlap story does not account for all cases of devoicing²⁰, it nevertheless offers a coherent explanation for many of the devoicing asymmetries. There are two robust asymmetries, however, which do not fall out directly from a gestural overlap account without recourse to other factors. First, there is the tendency for low-toned vowels to preferentially devoice over high-toned vowels pointed out in section 3.5. The existence of this asymmetry is presumably linked to the inherent inability of voiceless segments to carry tone phonetically. Thus, devoicing both high and low toned vowels would lead to neutralization of a tonal contrast, either a lexical contrast in the case of languages with lexical tone, or a semantic contrast, in the case of sentence (or phrase-level) intonation, e.g. questions vs. statements. In order to maintain the contrast, a language could thus devoice either high toned or low toned vowels but not both. Perhaps low toned vowels devoice, because the articulatory gestures involved in producing low tone are more compatible with the glottal abduction gestures associated with devoicing. An explanation for the tone asymmetry must await instrumental research.

The next asymmetry left unexplained by the gestural account of devoicing is the fact that devoicing of final vowels is so prevalent, in fact, even more prevalent than the devoicing of word-medial vowels in the vicinity of voiceless consonants. Furthermore, devoicing in final position typically takes place even when the preceding consonant is not voiceless. Strikingly, as pointed out in section 3.3., final devoicing itself respects an implicational hierarchy. The occurrence of devoicing in final position of a given domain implies devoicing in final position of smaller domains. Thus, utterance final devoicing in a language implies phrase final devoicing which implies word-final devoicing. This implicational hierarchy of devoicing can be explained in terms of the decline in subglottal pressure throughout the course of an utterance (Dauer 1980); this drop in subglottal pressure results in a decrease in the volume-velocity of air flow through the glottis which in turns inhibits devoicing. Subglottal pressure is lesser in final position of larger domains than in final position of smaller domains; hence, the likelihood of devoicing increases the larger the domain. Because subglottal pressure is lowest utterance finally, vowel devoicing is most common in this environment; cross-linguistic devoicing patterns reflect this fact. Those vowels whose glottal adduction gestures are inherently hypoarticulated either in terms of magnitude or duration, e.g. short vowels, high vowels, unstressed vowels and perhaps low toned vowels, are most susceptible to devoicing in final position. The gradual nature of the decline in subglottal pressure throughout the utterance is also compatible with the fact that devoicing in many languages (e.g. Acoma, Big Valley Shoshoni, Turkana, Nyangumata) affects not only the final vowel but also may extend farther back from the end of the domain in gradient fashion.

Interestingly, the decline in subglottal pressure is in direct competition with another common cross-linguistic property of final position: final lengthening (cf. Wightman et al. 1992). A priori one might expect, by analogy with the blocking of devoicing by phonemic long vowels and accented vowels in final position in many languages, that the additional phonetic length of final vowels would block devoicing. However, the gestures associated with final lengthening are different from those associated with other types of lengthening. Final lengthening does not involve an increase in gestural magnitude, unlike lengthening associated with accent (Beckman et al. 1992) or presumably, phonemic length. It is thus not surprising that, whereas phonemic long vowels and accented vowels inhibit final devoicing, final lengthening typically does not. In fact, final voiceless vowels, like their non-final counterparts, are usually described as being quite short, shorter than even non-final voiced vowels. Thus, the subglottal pressure decline not only inhibits final lengthening, it also appears to induce final *shortening*.

5. ARTICULATORY/AERODYNAMIC VS. PERCEPTUAL FACTORS. In summary, a combination of articulatory overlap and the decline in subglottal pressure in final position account for many of the devoicing patterns found cross-linguistically. Interestingly, the articulatory and aerodynamic factors which induce devoicing are in conflict with the perceptual factors militating against devoicing. Devoicing of vowels is articulatorily and aerodynamically natural under certain conditions as shown in the last section; however, voiceless vowels are perceptually less salient than voiced vowels as argued in section 2. The conflict between articulatory/aerodynamic factors and perceptual considerations is evident when we compare Hupa, in which long but not short vowels devoice, with the many languages (e.g. Cheyenne) in which short but not long vowels undergo devoicing. This conflict can be modeled in the grammar by assuming different ranking of the relevant constraints in the two language types. In Hupa, *NON-MODAL SHORT VOWELS is ranked above the relevant constraint forcing devoicing, whereas, in Cheyenne, *NON-MODAL SHORT VOWELS is ranked lower than the relevant constraint driving vowel devoicing between voiceless consonants. By shifting the rankings slightly we get other patterns. For example, the Acoma pattern, in which both short and long vowels devoice, is derived by ranking both *NON-MODAL SHORT VOWELS and *NON-MODAL LONG VOWELS below the constraint driving devoicing. If, on the other hand, we rank both *NON-MODAL SHORT VOWELS and *NON-MODAL LONG VOWELS above constraints requiring devoicing, we get a language without devoicing of any vowels. The rankings which generate the attested patterns are shown in (1). For expository purposes, the set of constraints which force devoicing are collapsed as a single constraint devoice. Although space limitations preclude doing so in this paper, these constraints can easily be divided into narrower (or broader, as in Hupa) constraints capturing the further asymmetries discussed in this paper.

(1) Ranking

Devoicing Targets

*NON-MOD V, *NON-MOD VV >> DEVOICING	No vowels
*NON-MOD V >> DEVOICING >> *NON-MOD VV	Long vowels
*DEVOICING >> *NON-MOD V, *NON-MOD VV	Short + long vowels
*NON-MOD VV >> DEVOICING >> *NON-MOD V	Short vowels

6. SUMMARY AND CONCLUSIONS. I have argued that the distribution of non-modal vowels is governed by the conflict between perceptual demands, on the one hand, and articulatory and aerodynamic considerations, on the other hand.

Although non-modal vowels are perceptually dispreferred to modal vowels, non-modal vowels may be articulatorily and aerodynamically preferred under certain conditions. The conflict between ease of perception and ease of articulation can be modeled in a constraint based grammar through different rankings of constraints.

Appendix: Languages in the survey of voiceless vowels. (Numbers following the data of Crothers et al.'s publication are the reference number of the relevant language in their database.)

Language	Source(s)	Language	Source(s)
Acoma	Miller 1965	Korean	Jun and Beckman 1993, 1994, Jun et al. 1997, 1998
Ainu	Crothers et al. 1979:370	Malagasy	Crothers et al. 1979:530
Alabama	Crothers et al. 1979:770	Mandarin	Crothers et al. 1979:455
Apinaye	Crothers et al. 1979:985	Mbay	Caprile 1968
Awadhi	Saksena 1971	Mixtec	Crothers et al. 1979:875
Azerbaijani	Crothers et al. 1979:325	Mokilese	Harrison and Albert 1976
Bagirmi	Gaden 1909	Nyangumata	Crothers et al. 1979:600
Big Smokey Valley Shoshoni	Crapo 1976	Oneida	Crothers et al. 1979:760
Boraana Oromo	Voigt 1984, Stroomer 1995	Papago	Saxton et al. 1983
Bulu	Alexandre 1962	Portugese, Brazilian	Crothers et al. 1979:205
Campa	Crothers et al. 1979:925	Quechua	Crothers et al. 1979:895
Cheyenne	Davis 1962	Sámi	Nielsen 1926
Chontal	Crothers et al. 1979:800	Sara	Vine 1981
Cocama	Faust and Pike 1959	Shina	Masica 1991
Comanche	Charney 1993	Southern Paiute	Sapir 1930
Dafla	Ray 1967	Taraskan	Crothers et al. 1979:810
French, Montreal	Gendron 1966, Cedergren and Simoneau 1985	Ticuna	Crothers et al. 1979:950
Gadsup	Crothers et al. 1979:655	Tongan	Feldman 1978
Galla	Vine 1981	Totonac	Crothers et al. 1979:795
Goajiro	Crothers et al. 1979:920	Tubu	Lukas 1953
Greek	Dauer 1980	Tunica	Haas 1946
Hupa	Golla 1970, Gordon 1998	Turkana	Dimmendaal 1983
Ik	Heine 1975	Turkish	Jannedy 1995
Inuit	Crothers et al. 1979:685	Tzeltal	Crothers et al. 1979:805
Island Carib	Crothers et al. 1979:910	Washkuk	Crothers et al. 1979:640
Japanese	Han 1961, Beckman 1982, Jun and Beckman 1993, Tsuchida 1994	Woleaian	Sohn 1975
Kawaiisu	Zigmond et al. 1990	Zuni	Crothers et al. 1979:830
Ket	Crothers et al. 1979:385		

*Thanks to Adam Albright, Victoria Anderson, Katherine Crosswhite, Bruce Hayes, Sun-Ah Jun, Pat Keating, Robert Kirchner, Peter Ladefoged, Pam Munro, Donca Steriade, Motoko Ueyama and Jie Zhang, and audiences at UCLA and at the 24th meeting of the Berkeley Linguistics Society for helpful discussion of the issues presented in this paper. Any misconceptions or inaccuracies are my own responsibility.

¹It is more difficult to make a case for positing underlying voiceless vowels in any language. Dafla (Ray 1967) and Turkana (Dimmendaal 1983) appear to be the strongest cases for languages with underlying voiceless vowels.

²Note that /h/ is not considered here as a voiceless vowel strictly speaking, since it

lacks place features unlike "true" voiceless vowels.

⁹It has been claimed (e.g. by Cho 1993 of Comanche) that some languages have both phonological and phonetic devoicing.

¹⁰In Ket, the short vowels are not explicitly described as voiceless, but they are described as overshoot and virtually absent, a description which suggests that they might in fact often be voiceless.

¹¹In Alabama, which lacks short vowels in final position, the last half of utterance final vowels devoice.

¹²The set of consonants triggering vowel devoicing in word-final position is a subset of those triggering devoicing in phrase-final position in Tarascan.

¹³In Ainu, devoicing is more likely to occur in phrase-final position than in word-final position.

¹⁴In Woleaian, the environment for devoicing might be more accurately described as clitic group, since a following article or deictic inhibits devoicing.

¹⁵In Apinaye, phrase-final vowels optionally either lengthen, devoice or become creaky.

¹⁶Vowels may devoice before silence which I interpret to mean utterance finally.

¹⁷In Kawaiisu, likelihood of devoicing is described as gradient; devoicing is most likely in utterance final position and less likely the smaller the domain in which the vowel is final.

¹⁸In Big Valley Shoshoni and Southern Paiute, the level of juncture which triggers devoicing is not explicitly stated in the sources consulted.

¹⁹In Totonac, the only utterance final short vowels occur after voiceless consonants.

²⁰In Malagasy, devoicing is more likely to affect word-final vowels after voiceless consonants.

²¹Among the high vowels, less peripheral vowel qualities devoice over more peripheral ones, e.g. in Ticuna and Ainu. There seems to be a slight tendency for /i/ to devoice over /u/, e.g. Mixtec, Gadsup, Greek, Turkish, although the opposite pattern is found in Brazilian Portuguese and Tunica. In Japanese, according to Han (1962), /u/ is more likely to reduce than /i/, an asymmetry which Han links to the lesser intrinsic duration of /u/.

²²Interestingly, in Tongan, /a/ devoices in certain environments, but mid-vowels do not.

²³Comanche also observes a restriction against two consecutive voiceless vowels; when two potential undergoers of devoicing are adjacent, devoicing only affects the first vowel. This is plausibly related to the alternating stress pattern in Comanche reported by Charney (1993). Primary stress in most words falls on the initial syllable with secondary stresses on alternating moras thereafter, but some words have primary stress on a non-initial syllable; in such words, pretonic syllables do not devoice (Charney 1993). Cho (1993) attributes this pattern to the leftward spread of a high tone onto preceding vowels.

²⁴High toned syllables could alternatively be considered accented syllables in Japanese.

²⁵Creaky vowels, i.e. those marked with what Miller (1965) refers to as the glottal accent, also do not devoice. Creak also blocks devoicing in Southern Paiute (Sapir 1930).

²⁶See Jun et al. (1997, 1998) for a detailed token by token analysis of devoicing in Korean, including discussion of results which are not compatible with an account based purely on gestural overlap.

REFERENCES

- ALEXANDRE, P. 1962. Sur la voyelle suffixielle du bulu. *JAL* 1.243-252.
 BECKMAN, M. 1982. Segment duration and the "Mora" in Japanese. *Phonetica* 39.113-135.
 BECKMAN, M.; EDWARDS, J. and FLETCHER, J. 1992. Prosodic structure and tempo in a sonority model of articulatory prominence. *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, ed. by G. Docherty and D. Ladd. Cambridge: Cambridge University Press.
 BLANKENSHIP, B. 1997. The time course of breathiness and laryngealization in vowels. UCLA Ph.D. dissertation.

- CEDERGREN, H.J. and SIMONEAU, L. 1985. La chute des voyelles hautes en français de Montreal: 'As-tu entendu la belle syncope?' Les tendances dynamiques du français parlé a Montreal, ed. by Lemieux, M. and Cedergren, H.J. 57-144.
- CHARNEY, J. 1993. A grammar of Comanche. Lincoln: U. of Nebraska Press.
- CHO, Y. 1993. The phonology and phonetics of 'voiceless' vowels. Proceedings of the 19th Annual Meeting of the Berkeley Linguistics Society.
- CRAPO, R. 1976. Big Smokey Valley Shoshoni. Desert Research Institute Publications in the Social Sciences 10.
- CROTHERS, J.; LORENTZ, J.; SHERMAN, D.; and VIHMAN, M. 1979. Handbook of phonological data from a sample of the world's languages. Stanford.
- DAUER, R. 1980. The reduction of unstressed high vowels in Modern Greek. Journal of the International Phonetics Association. 10.17-27.
- DAVIS, I. 1962. Phonological function in Cheyenne. IJAL 28.36-42.
- DIMMENDAAL, G. 1983. The Turkana language. Dordrecht: Foris.
- FAUST, N. and PIKE, E. 1959. The Cocama sound system. Publicações do Museu Nacional, Série Linguística Especial 1, 10-55.
- FELDMAN, H. 1978. Some notes on Tongan phonology. Oceanic Linguistics 17(2).133-139.
- FISCHER-JØRGENSEN, E. 1967. Phonetic analysis of breathy (murmured) vowels in Gujarati. Indian Linguistics 28.71-139.
- GADEN, H. 1909. Essai de grammaire de la langue baguirmienne. Paris: E. Leroux.
- GENDRON, J-D. 1966. Tendances phonétiques du français parlé au Canada. Québec: Les Presses de L'Université Laval.
- GOLLA, V. 1970. Hupa Grammar. Ph.D. dissertation. U. of California, Berkeley.
- GORDON, M. 1998. Laryngeal timing and correspondence in Hupa. UCLA Working Papers in Phonology 2.
- HAAS, M. 1946. A grammatical sketch of Tunica. Linguistic Structures of Native America. New York: The Viking Fund. 337-366.
- HAN, M. 1961. Japanese phonology: an analysis based upon sound spectrograms. Ph.D.diss. U. of Texas, Austin.
- HARRISON, S.; and ALBERT, S. 1976. Mokilese reference grammar. Honolulu: University of Hawaii Press.
- JANNEDY, S. 1995. Gestural phasing as an explanation for vowel devoicing in Turkish. OSU Working Papers in Linguistics 45.56-84.
- JUN, S-A.; and BECKMAN, M. 1993. A gestural-overlap analysis of vowel devoicing in Japanese and Korean. Paper presented at the 67th Annual Meeting of the Linguistic Society of America. Los Angeles, CA.
- JUN, S-A; and BECKMAN, M. 1994. Distribution of devoiced high vowels in Korean. Proceedings of the 1994 International Conference on Spoken Language Processing, vol. 1.479-482.
- JUN, S-A; BECKMAN, M.; NIIMI, S.; TIEDE, M. 1997. Electromyographic evidence for a gestural-overlap analysis of vowel devoicing in Korean. Journal of Speech Sciences 1.153-200.
- JUN, S-A; BECKMAN, M.; LEE, H-J. 1998. Fiberscopic evidence for the influence on vowel devoicing of the glottal configurations for Korean obstruents. UCLA Working Papers in Phonetics 96.
- KAGAYA, R. 1974. A fiberscopic and acoustic study of the Korean stops, affricates and fricatives. Journal of Phonetics 2.161-180.
- KIRK, P.; LADEFOGED, P.; and LADEFOGED, J. 1993. Quantifying acoustic properties of modal, breathy, and creaky vowels in Jalapa Mazatec. American Indian Linguistics and Ethnography in Honor of Laurence C. Thompson. UM Occasional Papers in Linguistics 10.435-50.
- LEHISTE, I. 1970. Suprasegmentals. Cambridge, MA: MIT Press.
- LÖFQVIST, A; and YOSHIOKA, H. 1980. Laryngeal activity in Icelandic obstruent production. Haskins Laboratories Status Report on Speech Research 63.275-92.
- LUKAS, J. 1953. Die Sprache der Tubu in der Zentralen Sahara. Berlin: Akademie-Verlag.
- MADDISON, I. 1984. Patterns of sounds. New York Cambridge University Press.

- MASICA, C. 1991. The Indo-Aryan languages. New York: Cambridge U. Press.
- MCCARTHY, J.; and PRINCE, A. 1995. Faithfulness and reduplicative identity. University of Massachusetts Occasional Papers in Linguistics 18, ed. by Beckman, J.; Urbanczyk, S.; and Walsh, L. Amherst, MA: Graduate Linguistics Student Association, University of Massachusetts, 249-384.
- MILLER, W. 1965. Acoma grammar and texts. Berkeley: U. of California Press.
- NIELSEN, K. 1926. Lærebok i Lappisk I: Grammatikk. Oslo: A.W. Brøggers Boktrykkeris Forlag.
- PÉTURSSON, M. 1976. Aspiration and activité glottale. *Phonetica* 33.169-98.
- RAY, P. 1967. Dafla phonology and morphology. *Anthropological Linguistics* 9.9-14.
- SAKSENA, B. 1971. Evolution of Awadhi. Delhi: Motilal Banarsidass.
- SAMELY, U. 1991. Kedang, (Eastern Indonesia), some aspects of its grammar. Hamburg: Helmut Buske Verlag.
- SAPIR, E. 1930. Southern Paiute, a Shoshonean language. *Proceedings of the American Academy of Arts and Sciences* 65 (1-3).
- SAPIR, E. 1912. The Takelma language of southwestern Oregon. Republished in Golla, V. (ed.). 1990. The collected works of Edward Sapir VIII. New York: Mouton de Gruyter.
- SAXTON, D.; SAXTON, L.; and ENOS, S. 1983. Papago/Pima-English, English-Papago/Pima dictionary. Tucson, AZ: University of Arizona Press.
- SILVERMAN, D. 1995. Phrasing and recoverability. *Garland's Outstanding Dissertations in Linguistic series*.
- SILVERMAN, D. 1998. Laryngeal complexity in Otomanguean vowels. *Phonology* 14(2). 235-262.
- SOHN, H. 1975. Woleaian reference grammar. Honolulu: U. of Hawaii Press.
- STROOMER, H. 1995. A grammar of Boraana Oromo (Kenya). Cologne: Rüdiger Köppe Verlag.
- THONGKUM, T. 1987. Phonation types in Mon-Khmer languages. *UCLA Working Papers in Phonetics* 67. 29-48.
- TSUCHIDA, A. 1994. Fricative-vowel coarticulation in Japanese devoiced syllables: acoustic and perceptual evidence. *Working Papers of the Cornell Phonetics Laboratory* 9. 183-222.
- VINE, B. 1981. Remarks on African 'shadow vowels'. *Harvard Studies in Phonology* 2.383-427.
- VOIGT, R. 1984. Zu den stimmlosen Vokalen des Borana. *Africana Marburgensia* 17(2).59-69.
- WIGHTMAN, C.W., S. SHATTUCK-HUFNAGEL, M. OSTENDORF, AND P.J. PRICE. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America* 92.1707-17.
- ZIGMOND, M.; BOOTH, C.; and MUNRO, P. 1990. Kawaiisu: a grammar and dictionary. University of California Publications in Linguistics 119. Berkeley: University of California Press.

Location and Direction in Waikurúan Languages¹

Verónica Grondona
University of Pittsburgh

1. Introduction

The Waikurúan language family comprises two branches: Waikurúan and Southern Waikurúan. The Waikurúan branch includes Mbayá, formerly spoken in the Brazilian and Paraguayan Chaco, and its descendant Kadiwéu, now spoken by about 1,500 people in western Mato Grosso do Sul, Brazil. The Southern Branch includes Mocoví, Pilagá, Toba, and [†]Abipón. Mocoví is spoken in northern Santa Fe and southern Chaco provinces in Argentina, by approximately 7,000 speakers. There are about 4,000 speakers of Pilagá scattered in the northeastern part of Chaco and in eastern Formosa provinces in Argentina. Toba, with 25,000 speakers, is spoken in southern Paraguay and eastern Bolivia, and in the eastern part of Chaco and Formosa provinces in Argentina (approximately 15,000 Toba speakers live in Argentina). Abipón, now extinct, was spoken in eastern Chaco province in Argentina and was very closely related to the other languages in this branch.

This paper is organized as follows. Section 2 presents a discussion of deictic classifiers in Waikurúan languages, including a table with the forms in each of the languages and example sentences. In § 3 locative/directional verbal morphemes are discussed, a table with locative/directional morphemes is presented, and examples from various Waikurúan languages are provided. Section 4 summarizes the conclusions.²

2. Deictic classifiers

All the Waikurúan languages have a set of nominal, deictic classifiers, which precede the noun in the noun phrase. They mark absence/presence of the noun they modify, as well as motion(coming/going) and position (standing/sitting/lying). They are marked for gender and number (although number in some languages is optional if it is marked on other elements in the noun phrase). They can occur with additional morphology marking distance relative to the speaker (e.g. Plg: *-ča* 'distal', *-ho?* 'proximal', *-mʔe* 'no reference to distance' (Vidal 1997:70); Mcv: *-kerawk* 'far', *-keram* 'farther').³ In some of the languages they can function as third person pronouns (sometimes with additional morphology).⁴

This set of noun classifiers is referred to in this paper as deictic classifiers because of their deictic meanings and their function as noun classifiers. However, they have received different labels in the literature of Waikurúan languages. Klein (1979) describes them as noun classifiers in Toba, Vidal (1997) considers *da?* 'vertically extended', *ñi?* 'sitting/non-extended' and *di?* 'lying/horizontally

extended' as positional classifiers, and *naʔ* 'coming/proximal', *soʔ* 'going away/past', and *gaʔ* 'absent/distal' as deictic classifiers. Sandalo (1995) for Kadiwéu, and Ceria and Sandalo (1995) for Waikurúan and Proto-Waikurúan refer to them as demonstratives.

The deictic classifiers in Waikurúan languages are presented in table 1.

		Kdw	Mcv	Plg	Tb	Abp ⁵
Sg.M	<i>absent</i>	i-ka	(e-)ka	ga?	ka	e-ka
	<i>present</i>					
	<i>mov. coming</i>	i-n:a	(e-)na	na? (~no?)	na	e-na
	<i>going</i>	i-jo	(e-)so	so?	so	e-ha
	<i>pos. standing</i>	i-d:a	(e-)da	da?	ra	e-ra
	<i>sitting</i>	i-n:i	(e-)ñi	ñi? (~ño?)	ñi	e-ñi
	<i>lying</i>	i-d:i	(e-)ji	di? (~dyo?)	ji	e-ri
Sg.F	<i>absent</i>	a-ka	(a-)ka	(ha-)ga?	a-ka	a-ka
	<i>present</i>					
	<i>mov. coming</i>	a-n:a	(a-)na	(ha-)na? (~no?)	a-na	a-na
	<i>going</i>	a-jo	(a-)so	(ha-)so?	a-so	a-ha
	<i>pos. standing</i>	a-d:a	(a-)da	(ha-)da?	a-ra	a-ra
	<i>sitting</i>	a-n:i	(a-)ñi	(ha-)ñi? (~ño?)	a-ñi	a-ñi
	<i>lying</i>	a-d:i	(a-)ji	(ha-)di? (~dyo?)	a-ji	a-ri
Pcl/Pl	<i>absent</i>	i-d:i-wa	ka-wa	ga?	ka:-/ka-wa	e-k-o(a)
	<i>present</i>					
	<i>mov. coming</i>	i-d:i-wa	na-wa	na: (~na?)	na:-/na-wa	e-n-o(a)
	<i>going</i>	i-d:i-wa	sa-wa	sa?	so:-/so-wa	?
	<i>pos. standing</i>	i-d:i-wa	da-wa	dya?	ra:-/ra-wa	e-r-o(a)
	<i>sitting</i>	i-d:i-wa	ña-wa	ña?	ñi:-/ñi-wa	e-ñ-o(a)
	<i>lying</i>	i-d:i-wa	ja-wa	dya?	ji:-/ji-wa	e-ri-o(a)

Examples (1a-f) from Mocoví show the same noun *ʔalo* 'woman' modified by different deictic classifiers.

- | | | | |
|--------|---|---------------|-------------------------|
| (1) a. | a-ka
F-class(absent) | ?alo
woman | ‘that woman (absent)’ |
| b. | a-na
F-class(coming) | ?alo
woman | ‘that woman (coming)’ |
| c. | a-so
F-class(going) | ?alo
woman | ‘that woman (going)’ |
| d. | a-da
F-class(standing) | ?alo
woman | ‘that woman (standing)’ |
| e. | a-ñi
F-class(sitting) | ?alo
woman | ‘that woman (sitting)’ |
| f. | a-ji
F-class(lying) | ?alo
woman | ‘that woman (lying)’ |

In example (2) from a Mocoví text, different deictic classifiers are used with the different nouns in the sentence, depending on the absence/presence, motion or position of the nouns they modify. The deictic classifier *so* 'going' modifies the noun *qom* 'person' since it refers to a person walking on a field; the deictic classifier *ka* 'absent' modifies *lepetaganagat* 'food' since it refers to the food the man is looking for, therefore absent; and finally the deictic classifier *ji* 'lying/horizontally extended' modifies the noun *no?we:naga* 'field'.⁶

- (2) **so** *qom* *nakitetako?* (Mcv)

Ø-**so** *qom* Ø-n-akite-tak-o?

M-class(going) person 3ag-hither-look.for-prog-pst

ka *lepetaganagat* *yowo:tako?* *ki* **ji** *no?we:naga.*

Ø-**ka** l-epetaganagat y-owo:-tak-o? *ke* Ø-**ji** n-o?we:naga

M-class abs-food 3ag-walk-prog-pst prep **M-class** abs-field

(absent)

(lying)

'That man was looking for food, (he) was walking on the land (=field).'

Examples (3) and (4) from Pilagá show the different deictic classifiers modifying different nouns, again depending on their absence/presence, motion or position.

- (3) **so** *serak ya-čangi* **ha-ñi** *kaxa* **di** *ganaat* (Plg)

class(going) name 3sg-put **F-class** box **class** knife

(sitting)

(standing)

'Seraki put the knife in the box.' (Vidal 1997:59)

- (4) **ñi** *mayo?* *netaye qa?li?* **ha-da** *epaq* (Plg)

class(sitting) bird loc before **F-class(standing)** tree

'The bird was on the tree.' (Vidal 1997:76)

In example (5) from Klein (1979) the classifier *ra* 'standing' modifies the noun *lma?* 'house' referring to a house that is being built, and is already off the ground. However, in example 6 the house is just about to be built and is barely off the ground, so the classifier *ji* 'lying' is used.

- (5) *hi?ottak* **ra** *lma?* (Tb)

he.is.building **class(stdng)** house

'He is building a house.' (Klein 1979:89)

- (6) hiʔottak ʒi lmaʔ (Tb)
 he.is.building **class(lying)** house
 'He is building a house.' (it is just about to be built, is only barely off the ground) (Klein 1979:89)

In example (7) the deictic classifier *so* 'going' modifies the nouns *piʔoq* 'dog' and *roʔo* 'hat' when describing a situation in which the dog is biting on a hat and the dog is moving. Example 8 refers to a similar situation, in which a dog is biting on a hat. However in this case both the dog and the hat are out of sight, so the deictic classifier *(ha-)ka* 'absent' is used for the same nouns.

- (7) **so** piʔoq hinaktapigi **ha-so** roʔo (Tb)
 class(going) dog is.biting.on **F-class(going)** hat
 ‘The dog is biting on a hat.’ (and the dog is moving in such a way that
 both it and the hat are almost out of sight of the speaker) (Klein 1979:87)

- (8) **ka** pi?oq hinaktapigi **ha-ka** ro?o (Tb)
 class(absent) dog is.biting.on **F-class(absent)** hat
 ‘The dog is biting on a hat.’ (and both the dog and the hat are out of sight
 of the speaker) (Klein 1979:87)

Example (9) from a Kadiwéu text shows the classifier *ika* ‘absent’ used when referring to a situation in the past, in which the various nouns are ‘absent’ from the visual field.

- (9) **ngika** **ika** **ejewa jęgi** **bagalei:gači** (Kdw)
ng-i-ka **jotigide i-ka** **ejewajęgi** **bga+le+y-i:gači**
prox-M-class **ancient M-class** **Kadiwéu** **compl+?+3sgS-teach**
(absent) **(absent)**

ika ly:onig:i datematiqatema (Kdw)
i-ka l-y:o-nig:i y-d:-atemati-qan-t+e-ma
M-class(absent) 3poss-son-Mdim 3sgS-theme-tell-val-rel+3sgCL-ben
‘As for the Kadiwéus, the ancient people used to teach their sons telling
stories to them.’

(Lit: 'These ancient people these Kadiwéus used to teach their sons telling stories to them.')(Sandaló 1995:87)

To summarize, then, deictic classifiers in Waikurúan languages have very similar meanings, very similar functions, and very similar forms. They mark absence/presence in the visual field as well as motion and position of the noun

they modify; they precede the noun in the noun phrase, and they are marked for gender and (optionally) number.

3. Locative/directional verbal morphemes

All the Waikurúan languages show a set of locative/directional (loc/dir) verbal morphemes that mark the location and/or direction of the action expressed by the verb. Although the exact position of these morphemes varies slightly from language to language, in all the Waikurúan languages these loc/dir morphemes are part of the verb form and they follow person number and aspect markers within the verb form. In Toba and Kadiwéu some of them can co-occur, however it is not clear whether this is the case for Mocoví, Pilagá and Abipón. In Mocoví and Kadiwéu these loc/dir morphemes are described as clitics. In Toba and Abipón they are described as suffixes. It is not clear whether they are clitics or suffixes in Pilagá.

Examples (10)-(14) from Mocoví show some locative/directional morphemes added to the same verb root, *añogot* 'hide'.

- | | | | | |
|------|---|-------------------|------------------|-------|
| (10) | lwis nanogočicit | ada | qoʔpaq | (Mcv) |
| | lwis Ø-n-añogot+igit | a-da | qoʔpaq | |
| | Luis 3ag-hither-hide+ behind | F-class(standing) | tree | |
| | 'Luis hides behind the tree.' | | | |
| (11) | lwis nañogotowgi | ñi | ʔimek | (Mcv) |
| | lwis Ø-n-añogot+owgi | Ø-ñi | ʔimek | |
| | Luis 3ag-hither-hide+ tds.the.inside | M-class(sitting) | house | |
| | 'Luis hides inside the house.' | | | |
| (12) | lwis nañogočiʔñot | ʃi | nkiʔyagala | (Mcv) |
| | lwis Ø-n-añogot+iʔñot | Ø-ʃi | n-kiʔyagala | |
| | Luis 3ag-hither-hide+ under | M-class(standing) | abs-table | |
| | 'Luis hides under the table.' | | | |
| (13) | lwis nañogohlek | ñi | ʔimek lelaq | (Mcv) |
| | lwis Ø-n-añogot+lek | Ø-ñi | ʔimek l-elaq | |
| | Luis 3ag-hither-hide+ on | M-class(sitting) | house 3poss-roof | |
| | 'Luis hides on the roof of the house.' | | | |

- 'Luis hides on the other side of the river.'

Examples 15-19 show sentences from Mocoví, Pilagá, Toba, Abipón and Kadiwéu, with the same loc/dir morpheme meaning ‘up/upwards’ (+*šigim* (Mcv)/-*sem* (Plg)/-*šigem* (Tb)/-*hegem* ~ -*ihegem* (Abp)/+*bigim* (Kdw) ‘upwards’).

- 'We lift the bag up to the tree.'

- 'He plans to build a house.' (Vidal 1997:92)

- (Klein 1981:228)

- 'You are standing (=you are up, you are standing up)' (Najlis 1966:40)

Table 2
Locative/directional verbal morphemes in the Waikuruan languages

	Kdw	Mcv	Plg	Tb	Abp
'downwards' 'hither'	+nigi ~ +ni:	+fi dwn(wds) +o ~ +two		-fi -get	-afi ~ -ifi down
'inside'	+n going inside	+kena tds.here +fiigi +awgi in (?)		-ngi -igi	-oa
'inwards, into'	+w ~ +wgi inwds	+owgi ~ +iwgi into +lek on, over		-wo -lek	-alge~elge
'on'		+aŋa on/to.other.side	-lege on (?)	-ata	
'on/to.other.side.of'				-ŋoga across	
'outwards'	+ke	+eg ~ +weg		-wek	-ooge ~ -ook out(wds) -aa outside
'towards'	+gi:	+igi	-ge	-ot	
'under'		+ot		-ŋigem	-hegem ~ -ihegem up(wds)
'upwards'	+bigim	+ŋigim	-sem	-peget	
'with'	+wag going together	+e'e		-ga:ma tds.water/town	-hagam ref.to.water
'towards water'	+ko going straight			-teg straight	
'straight'	+ka absent	+igit behind	-ma ?	-ŋa	-aigit around, near
	+we bckwds	+pege? upto	-pe ?	-asop next.to	
	+kwak apart		-ya ?	-wag tds.fire	
	+jo going		-wo ?	-pe circular	
	+get going against				

'These morphemes in Kadiwéu are the dirli clitics in CLITIC 2 position in the verb form and they all express 'motion'. All others express 'direction' and occur are the dirli clitics in the CLITIC 2 position in the verb form.

- (19) nekenigo walokoditibigimed:i nalagate (Kdw)
 neke-nigo w-alokon-d-t+bigem+e-d: nalagate
 dog-class 3sgS-run-atel-rel+upwds+3sgCL-theme mountain
 'The dog ran up the mountain.' (Sandalo 1995:55)

Examples (20)-(23) show sentences with the loc/dir morpheme 'out' or 'outwards'. Mocoví + *weg* 'out(wards)' in example (20), Pilagá -*gek* 'outwards' in (21), Toba -*wek* 'out' in (22) and Abipón -*ge* 'outwards' in (23).

- (20) yim sa:**wek** ke ĵi noʔwe:naga (Mcv)
 yim s-a:+**weg** ke Ø-ĵi noʔwe:naga
 1pron 1ag-go+**outwds** prep M-dem(lying) field
 'I go (out) to the fields.'
- (21) naega? awa-pya-**gek** na? l-apat (Plg)
 interrog 2sg-cut-**outwds** class(prox) 3poss-meat
 'What do you cut meat with?' (Vidal 1997:79)
- (22) senoganagawek (Tb)
 s-enogan-ag-**wek**
 1S-go-1pl-out
 'Let's get out of here, we are leaving for outside.' (Klein 1981:232)
- (23) ñatagaoge (Abp)
 ñ-atagao-ge
 1S-spit-**outwds**
 'I spit (outside)'

In summary, then, all the Waikurúan languages have a set of locative/directional morphemes encoding the location and/or direction of the action expressed by the verb. Although not all the loc/dir morphemes occur in all the languages, most of them occur in two or more Waikurúan languages, and in those cases the forms are very similar. The structure of the verb form in the Waikurúan languages is very similar, and these loc/dir morphemes occur in very similar positions within the verb form.

4. Conclusions

All the Waikurúan languages have a set of deictic classifiers which precede the noun in the noun phrase and express the absence/presence, motion (coming/going) and position (standing/sitting/lying) of the noun they modify. These deictic classifiers have very similar forms, functions, and meanings. However, a few questions still remain unanswered. What is the form of the third person pronoun

in all the Waikurúan languages. In those languages for which the deictic classifiers can be used as third person pronouns, do they occur by themselves or do they require additional morphology, and if they do require additional morphology, what are those morphemes?

These languages have a set of loc/dir morphemes which have similar forms, similar meanings, and which occur in similar positions within the verb form. However, the set of loc/dir morphemes is not complete yet. It is not clear whether they can co-occur in all the Waikurúan languages, and if they can, what the restrictions are, if any. Further research is needed to find out how these loc/dir interact with the noun phrases in the sentences. In Mocoví, they seem to modify the argument structure of some of the verbs they attach to, and it is not clear whether this is the case as well in other Waikurúan languages.

This is only the first step in a comparative study of location and direction in Waikurúan languages. A few questions still remain to be answered. These will be the basis for future research.

Notes

¹ This work was supported in part by the Center for Latin American Studies of the University of Pittsburgh, and the Tinker Foundation.

² List of abbreviations: Abp, Abipón; Kdw, Kadiwéu; Mcv, Mocoví; Plg, Pilagá; Tb, Toba; Wkr, Waikurúan; 1, first person; 2, second person; 3, third person; abs, absolutive; ag, agent; CL, clitic; class, classifier; dem, demonstrative; F, feminine; loc, locative; M, masculine; non-ag, non-agent; O, object; pcl, paucal; pl, plural; poss, possessive; pron, (independent) pronoun; S, subject; sg, singular; stdng, standing; tds, towards; val, valency; †, extinct language.

³ Phonemic inventory of the Waikurúan languages:

Kadiwéu: stops and affricates: p, b, b:, t, d, d:, j, č, k, g, g:, q; fricatives: ɣ; nasals: m, m:, n, n:; approximants: l, l:, w, w:, y, y:; vowels: a, a:, e, e:, i, i:, o, o:.

Mocoví: stops and affricates: p, t, d, č, j, k, q, ʔ; fricatives: (ɸ), s, š, ɣ, ɣ; nasals: m, n, ñ; approximants: l, r, λ, h; vocoids: w, y. Vowels: i, i:, e, e:, a, a:, o, o:.

Pilagá: stops and affricates: p, t, d, k, q, ʔ, g, č; fricatives: s, (x), ɣ; nasals: m, n, ñ; approximants: l, r, w, y, h; vowels: a, e, i, o.

Toba: stops and affricates: p, t, č, j, k, q, g, ʔ; fricatives: s, š, ɣ; nasals: m, n, ñ; approximants: l, λ, r, w, y, h; vowels: a, a:, e, e:, i, i:, o, o:.

Abipón: stops and affricates: p, t, č, k, q, g, {ʔ}; fricatives: ɣ; nasals: m, n, ñ; approximants: l, r, w, y, h; vowels: a, e, ě, i, o.

() marks a phoneme which occurs only in loanwords; { } marks a phoneme that seems to have existed in Abipón but is not marked in the sources.

⁴ In Pilagá there is an additional classifier *hen* 'general classifier', used only when pointing out an entity physically proximate to the speaker. It can be used also with mass nouns with no specific reference and with nouns such as 'sky', 'land/earth', 'moon' or 'sun'. (Vidal 1997:82-83)

- (1) qomí? sa-liena-k hen lapat (Plg)
pron.1pl 1-eat-pl class meat
'We eat meat.' (Vidal 1997:82)

- (2) wʔo hen noop (Plg)
exist class water
'There is water.' (=pointing at it) (Vidal 1997:82)

⁵There are some discrepancies among the sources for Abipón. However, I have taken Najlis (1966) as a source for the classifiers listed in the table since her work is based on that of the other sources.

⁶ The examples from Mocoví are from my own fieldwork. Examples from other languages are taken from various sources. The source of each example is provided in parenthesis in the last line. Whenever possible, morpheme-by-morpheme glosses are provided for all the examples, even if they are not given in the original sources. The language that each example corresponds to is provided in parenthesis in the first line. The examples are organized as follows: the first line is a phonetic or phonemic transcription (depending on the source); the second line provides a morpheme-by-morpheme breakdown; the third line shows morpheme-by-morpheme glosses; and the fourth line is a free translation.

References

- Banfi, Cristina. 1995. Indigenous Languages in Argentina. ms.
- Bigot, Margot. 1992. La lengua Qom (Toba) del Chaco, Argentina. En ACTAS: Segundas Jornadas de Lingüística Aborigen. Instituto de Lingüística. Facultad de Filosofía y Letras. Universidad de Buenos Aires. Buenos Aires. 177-189.
- Buckwalter, Alberto. 1980. Vocabulario Toba. Chaco, Argentina.
- Ceria, Verónica & Filomena Sandalo. 1995. A Preliminary Reconstruction of Proto-Waikurúan with Special Reference to Pronominals and Demonstratives. *Anthropological Linguistics* 37. 169-91.
- Dobritzshoffer, Martin S.J. 1967 (1784). Historia de los Abipones Tomos I. Traducción de Clara Vedoya de Guillén. Departamento de Historia, Facultad de Humanidades, Universidad Nacional del Nordeste. Resistencia: Chaco.
- Dobritzshoffer, Martin S.J. 1968 (1784). Historia de los Abipones Tomos II. Traducción de Edmund Wernicke. Departamento de Historia, Facultad de Humanidades, Universidad Nacional del Nordeste. Resistencia: Chaco.
- Dobritzshoffer, Martin S.J. 1970 (1784). Historia de los Abipones Tomos III. Traducción de Clara Vedoya de Guillén. Departamento de Historia, Facultad de Humanidades, Universidad Nacional del Nordeste. Resistencia: Chaco.
- Gronadona, Verónica. 1998. Locative/Directional Verbal Morphemes and Valency Changes in Mocoví. Paper presented at the 1997 Annual Meeting of the Society for the Study of Indigenous Languages of the Americas (SSILA). New York. January 1998.
- Hernández, Isabel. 1992. Los Indios de la Argentina. Editorial Mapfre. Madrid, España.
- Kaufman, Terrence. 1994. The Native Languages of South America. Atlas of the World's Languages. Routledge.
- Klein, Harriet M. 1978. Una Gramática de la Lengua Toba: Morfología Verbal y Nominal. Montevideo Uruguay: Universidad de la República.
- _____. 1979. Noun Classifiers in Toba. In Madeleine Mathiot (ed). *Ethnolinguistics: Boas, Sapir and Whorf Revisited*. Mouton: The Hague.
- _____. 1981. Location and Direction in Toba: Verbal Morphology. *IJAL* 47.227-235.
- _____. 1986. Styles of Toba Discourse. In Joel Sherzer (ed). *Native South American Discourse*. Mouton de Gruyter. Berlin, Germany.
- _____. 1987. Future Precedes the Past: Time in Toba. *Journal of the International Linguistic Association* 38.173-185.
- Lafone Quevedo, Samuel A. 1891. Principios de Gramática Mocoví. Buenos Aires: Imprenta de Pablo E. Coni e Hijos.
- _____. 1896. Idioma Abipón. Buenos Aires: Imprenta de Pablo E. Coni e Hijos.
- Martínez Sarasola, Carlos. 1992. Nuestros Paisanos los Indios. Emecé Editores. Buenos Aires Argentina.
- Najlis, Elena L. 1966. Lengua Abipona, Tomos I y II. Buenos Aires, Argentina: Universidad de Buenos Aires, Facultad de Filosofía y Letras, Centro de Estudios Lingüísticos.
- Sanchez Labrador, José S. I. 1760. Gramática Eyiguayegí-Mbayá. Según el manuscrito del siglo XV. In B. Susnik. 1971. Familia Guaycuru. Asunción, Paraguay: Museo Etnográfico "Andrés Barbero" (Lenguas Chaqueñas 1.) 1-166.
- Sandalo, Filomena. 1995. A Grammar of Kadiwéu. PhD Dissertation. Department of Linguistics University of Pittsburgh.
- Vidal, Alejandra. 1997. Noun classification in Pilagá (Guaykuruan). *Amazonian Linguistics* 1. 58-111.

Linguistic Systems and Social Models: A Case Study from Japanese

Yoko Hasegawa
University of California, Berkeley

1. Introduction

It is widely recognized by researchers in pragmatics and sociolinguistics that certain linguistic subsystems of a given language cannot be accounted for without recourse to the social organizations of the speech community, e.g. addressing and kinship terms, linguistic politeness.* Anthropologists, ethnographers, and sociologists propose a social model and commonly use linguistic subsystems to support their arguments; linguists, in turn, use such a model to explicate various linguistic phenomena. Thus, a mutual dependency exists between sociological and linguistic investigations. Japanese is one of the languages that have been most extensively investigated in this respect.

Because social models are usually constructed explicitly or implicitly for a particular issue with a particular readership in mind — e.g. in the case of Japanese, to plan the Allied Occupation of post-war Japan, to explain Japan's rapid economic growth in the 60s, to solve US-Japan trade imbalances in the 80s — the adequacy of each model is relative to the objectives of the investigation. It is, therefore, justifiable, in principle, for researchers to use linguistic data *selectively* to meet their needs. However, when a model is applied to unintended areas, this selectivity of linguistic data can lead to a distorted view of the speech community. This, in fact, is the case with the Japanese language. The present study (i) reviews several major works covering both the Japanese language and Japanese society, and (ii) demonstrates that the prevailing group model of Japanese society is inconsistent with the notion of *self* as encoded in the Japanese language.

2. Fluidity of Japanese self

Japanese society has frequently been characterized in terms of *groupism* and *contextualism*. The former claims that Japan is overwhelmingly group-oriented, and the latter contends that the notion of Japanese self is relative and interpersonal, not a fixed reference point residing within each individual. It is the latter that has direct relevance to an understanding of the Japanese linguistic phenomena. This section summarizes some major works representing those views.

Groupism vis-à-vis individualism refers to such characteristics as 'the individual's identification with and immersion into the group, conformity and loyalty to group causes, selfless orientation towards group goals, and consensus and the lack of conflict among group members' (Yoshino 1992:19). The Japanese are said to be 'extremely sensitive to and concerned about social interaction and relationships' (Lebra 1976:2). 'Even in intimate groups there are strong pressures to conformity, which many have seen as the source of the deepest psychological

malaise in Japanese society' (Smith 1983:56). To provide an anecdotal example of Japanese group consciousness, Nakane (1970:2-3) points out that the Japanese commonly introduce themselves with their affiliation, rather than with their personal attributes. For example, a Japanese person is more likely to say 'I'm from X University' or 'I belong to Y Company' rather than 'I'm a psychologist' or 'I'm an engineer.'¹

Since Ruth Benedict's 1946 book, *The Chrysanthemum and the Sword*, Japanese society has often been characterized as *hierarchical*: it involves vertical stratification by institution or group of institutions, rather than horizontal stratification by class or caste, and each group is vertically organized based on the relationships between paternalistic superiors and their subordinates (Nakane 1970). The society assumes loyalty from below and benevolence from above.

Underlying this vertical society is said to be the Japanese personality called *amae* 'dependence/indulgence' (Doi 1973). *Amae* is 'the feelings that normal infants at the breast harbor toward the mother — dependence, the desire to be passively loved, the unwillingness to be separated from the warm mother-child circle and cast into a world of objective "reality"' (Bester 1973:7). This attitude of dependence is reported to be carried into adulthood, and dependence on others' benevolence is encouraged during socialization processes of the Japanese (DeVos 1985:165). This type of dependency is considered to occur in group settings: subordinates, who play the child role, can seek dependence on their superior, and the superior, who plays the parent role, is expected to display benevolence (Yoshino 1992:18).

Closely linked to Japanese group orientation is the notion of *uti*. The translational approximation of *uti* is 'inside', but *uti* is also used to refer to the speaker's own home, house, or household. *Uti* is also commonly associated with 'insider(s)'. *Uti* and its antonym *soto* 'outside/outsider' are said to constitute 'a major organizational focus for Japanese self, social life, and language' (Bachnik 1994:3), and without these concepts, much of Japanese behavior is 'inexplicable (at least from a Western perspective)' (Wetzel 1994:74). *Uti* and *soto* are also the key words to connect groupism and contextualism — the latter concerning the notion of Japanese self.

Characterized as relational and social, Japanese self is said to be *situationally* defined (Araki 1973). Advocates of this view claim that in Japan, relationships between individuals are prioritized over individual self — which is more central to the Western notion of self (Bachnik 1994:18), and that 'the identification of self and other is always indeterminate in the sense that there is no fixed center from which, in effect, the individual asserts a noncontingent existence' (Smith 1983:81). Furthermore, 'proper use of Japanese teaches one that a human being is always and inevitably involved in a multiplicity of social relationships. Boundaries between self and other are fluid and constantly changing, depending on context and on the social positioning people adopt in particular situations' (Kondo 1990:31).

To recapitulate, compared with 'the Western notion of self,' Japanese self is claimed to be unstable, constantly shifting, and context dependent. It is imperative to discuss here what the term 'Western notion of self' is used to refer to. Because most works I consulted use 'Western self' as a reference point to which Japanese self is compared and explained, I could not find elaborate discussions of the Western self. However, it appears that what is widely assumed is the Cartesian notion of self: 'I am transparent to myself: My self-knowledge is mediated neither by inference nor by any teleological element such as a passing purpose or project. I know my own identity directly and completely, whereas others know it only inferentially and relative to certain sets of purposes' (Boër & Lycan 1986:139). Like most authors cited in the present work, I use the term *self* in this sense.

The group model with fluidity of self is customarily utilized to account for linguistic phenomena in Japanese, most notably for address/kinship terms, extended use of donatory verbs, and the polite/honorific systems.

Regarding kinship terms, the unmarked choice for the word corresponding to *mother* is *okaasan*. It can be used to refer to the addressee's as well as a third-person's mother. It can also be used for the speaker's own mother in a conversation with in-group members or in informal conversations with outsiders. In formal conversations, on the other hand, *haha* must be employed rather than *okaasan* to refer to one's own mother. The same distinction is made for father, grandparents, siblings, and other close relatives. This demonstrates that an appropriate choice of kinship terms depends on context — whether or not the addressee is an insider and the formality of the conversation.

An example to illustrate fluidity of boundaries between the self and others is drawn from the usage of donatory verbs. There are two kinds of verbs corresponding to *give*: *kure-* and *age-*. With *kure-*, the inherent destination of the transfer is the speaker, i.e. the self. To indicate that Okada gave money to her, the speaker would say:

- (1)a okada-san ga okane o kasite **kure-ta.**²
 Okada NOM money ACC lend-INF give-PAST
 Lit. 'Mr. Okada gave [me a favor of] lending money.'

Kure- can also be used when the recipient of the transfer is regarded as an insider, as shown in (1b).

- (1)b okada-san ga haha ni okane o kasite **kure-ta.**
 Okada NOM mother DAT money ACC lend-INF give-PAST
 Lit. 'Mr. Okada gave [my mother a favor of] lending money.'

Thus, a common analysis treats insiders as *extended self*, i.e. the boundary between the self and others is considered to be shifted.³

Fluidity of in-group/out-group boundaries can be observed in the proper usage of honorific/humble forms of the predicate. Speaking with a colleague about their company president, an honorific form should be used when the president is

- (6) akiko wa haha wa byooki da to omo-tta.
 Akiko TOP mother TOP ill COP QUOT think-PAST
 'Akiko thought her mother was ill.'

The relevant notion of self here is absolute and cannot be relative or context dependent. Furthermore, the boundary between the self and others cannot be fluid. That is, as shown in (3b, 4b, 5b), typical in-group members such as *haha* 'Mother' cannot be regarded as extended self. This very fact indicates that the Japanese language is extremely self-conscious, however primordial such a notion of self may be.⁵ And it is difficult to envision that such a language has been formed by a speech community lacking the concept of individual self, akin to the Western notion of self, contrary to those who contend that Japan is a 'selfless' society. This is a well-known fact in Japanese linguistics, but it has never received serious attention in research on Japanese society. The prevailing relational model of self is inconsistent with this aspect of the Japanese language.

4. Relational vs. absolute self

So far we have discussed that the Japanese language appears to involve both relational and absolute notions of self. Are they both indispensable? I would argue that recognition of the absolute self is an integral part of the Japanese language, but the relational self is not.

First, linguistic expressions based on the absolute self are learned by children automatically and unconsciously, whereas expressions requiring the notion of relational self must be taught explicitly by adults. First graders are not confused as to whose opinion is represented in (7), with a psych predicate *omow*- 'think'; it is the speaker's, not her mother's or teacher's or whoever's is present in the discourse.

- (7) aki-tyan kuru to omo-u.
 Aki-HYPOCORISTIC come QUOT think-NPAST
 '[I] think Aki will come.' NOT 'Aki thinks [she/someone else] will come.'

On the other hand, the use of *haha* to refer to one's own mother is usually taught in the upper grades of elementary education.

Second, most dialects of Japanese provide neither dual kinship terms nor honorific/humble forms of verbals. The elaboration of the honorific system began in Late Old Japanese by members of the Japanese aristocracy living in and around Kyoto, the capital city at the time. Unlike the Tokyo dialect, which borrowed a considerable number of honorific expressions from the Kyoto dialect, most other dialects do not have the elaborate honorific systems that are frequently cited in the literature on Japanese society (cf. Shibatani 1990:123-26).

Third, expressions that presuppose shifting (relational) self as illustrated in (2) are artificial, not a natural part of language as linguists conceive it. Manipulation of honorific and humble forms according to the situation is not learned naturally. Many companies provide their employees with honorific-language lessons

as part of on-the-job training. Those employees, most of whom hold college degrees, have not acquired such a language in their twenty-plus years of life. Books teaching how to use honorifics properly are always in great demand and found in virtually all bookstores in urban communities.

While it is indeed amazing that the Japanese continue to maintain their elaborate honorific language at such a high cost, as far as the Japanese language is concerned, the presence of individual (absolute) self is unquestionable.

With a few exceptions,⁶ the researchers who advocate a multiple or relational notion of Japanese self have attempted to derive it based on behavioral observations; as Bachnik (1992:152) asserts, 'Japanese choose appropriate behavior situationally, from among a range of possibilities, resulting in depictions of the Japanese self as "shifting" or "relational."' This statement is a non sequitur, for holders of individual self can and should also behave appropriately in various situations.

Bachnik's idea of relational self is not as simple as it sounds, however. She argues that the crucial question for defining self is 'to look away from "fixing" a unified typology of self [i.e. individual (absolute) self vs. relational self] to the process of "fixing" a series of points along a sliding scale for a self which is defined by shifting. If movement between different modes is central to the organization of self, then the question of how movement is initiated and defined at a given "point" along the sliding scale is *the* central organizing factor for the double continuum [e.g. *uti* and *soto*] which social behavior is identified, not as a general set of behaviors which transcends situations, but rather as a series of particular situations which generate a kaleidoscope of different behaviors *which are nonetheless ordered and agreed upon*' (155, emphasis in original). Bachnik's notion of self appears to be constantly shifting, but at any given moment, it must be identifiable at some point along a sliding scale. Let us now examine whether or not this is the case.

As mentioned earlier, the verb *kure*- 'give' is used when the favor is made for the benefit of the speaker, and the speaker is grateful for it. It can also be used when the beneficiary is an insider, e.g. Mother. In the theory of relational self, this is made possible because the boundary between the self and others is shifted: a third-person beneficiary (the speaker's mother in this case) is included in the notion of relational self. We have also observed that the use of psych predicates, such as *omow*- 'think', is restricted to the first-person subject. Let us examine the case in which these two kinds of predicates co-occur in a single sentence.

- (8) okada-san ga haha ni okane o kasite kure-ru
 Okada NOM mother DAT money ACC lend-INF give-NPAST
 to omo-u.

QUOT think-NPAST

'I think Mr. Okada will give my mother [a favor of] lending money.'

NOT 'My mother thinks Mr. Okada will give [her a favor of] lending money.'

In (8), the destination of the favor of lending money is the speaker's mother, but the subject of *omow*- remains constantly the first person, *watasi*. That is, the use of psych predicate is still impossible with a third-person subject. This fact indicates that the self and the mother must belong to different conceptual categories — one permits a direct expression about a mental state based on the guaranteed accessibility to the source, but the other does not, due to the lack of such accessibility.⁷

5. The group model revisited

As mentioned earlier, social models are formed with respect to particular purposes, implicit or explicit. Models are ideological in nature: ideal but visionary speculation. Various characteristics are frequently abstracted out to form a monolithic and coherent model. It is, therefore, worth exploring the contexts in which such a model was formulated. This section will investigate how the prevailing model of groupism and contextualism, which is so popular in pragmatics and sociolinguistics literature about Japanese, came about.

Since Basil Hall Chamberlain (1850-1935), Western observers have depicted Japanese national cultural traits in an *Orientalist tradition*, which shares ontological assumptions about the West as the universal reference point and the exotic and inferior other (Minear 1980:508). Those observers were fascinated with some exotic characteristics of Japan. Benedict (1946:2) depicted Japan as fundamentally paradoxical: 'the Japanese are both aggressive and unaggressive, both militaristic and aesthetic, both insolent and polite, rigid but adaptable, submissive and resentful of being pushed around, loyal and treacherous, brave and timid, conservative and hospitable to new ways.'

While Japan's constructed unity was created by Western observers, such an imaginary community has been supported and extended by the Japanese themselves, for representation and dissemination of various ideologies and myths are necessary parts of construction of a unified nation. This strategic 'Japaneseness,' characterized by Miller (1982) as 'self-Orientalism,' maximizes national interests and minimizes individualism.

Iwabuchi (1994) analyzes that in the process of Japan's self-Orientalization, which has been done most extensively in the last fifty years, images of the 'West' have been discursively created.⁸ Western nations were considered superior and to be emulated, but they were also condemned as individualistic and selfish. He goes on to assert that it was in this context that self-Orientalisation was utilized as a strategy to construct and self-assert Japan's national cultural identity, and 'the West' was necessary for Japan's invention of tradition, the suppression of heterogeneous voices within Japan, and the creation of a modern nation whose people are loyal to Japan.

Both the West's Orientalist characterization of Japan and Japan's self-characterization tend to use the other to the self and to repress the heterogeneous voices within: heterogeneous voices of people within the nation have been re-

pressed through the homogenizing discourses of an imaginary 'us' versus 'them,' although Japan is neither static nor homogenous nor is it closed as a society (Iwabuchi 1994).

In the late 1970s, researchers began casting doubt upon the validity of the group model of Japanese. In comparison with people of other societies, are the Japanese more group-oriented and accustomed to vertical organization? Do they place more emphasis on consensus and social harmony, and value more deeply group membership or social solidarity? Do they have underdeveloped egos, and lack an autonomous sense of self-interest? (Mouer & Sugimoto 1986:11-14).

Rosenberger (1992:13) asserts that the Japanese are 'not essentialized individuals despite economic and social changes. ... Japanese self emerges as neither entirely collective nor completely individualistic. In fact, [the essays in the same volume] imply that whether or not Japanese are becoming individualistic in an American sense is the wrong question, one itself rooted in Western dichotomies. The more appropriate question is, what shifts occur as Japanese people make Western lifestyles and concepts of individuality part of their own processes of self and social relationship.' I counter her opinion with the contention that the real questions here are (i) what concepts the term *individuality* signifies in each work on the relationship between Japanese society and its language, and (ii) what evidence the advocates of the group model provide to support their claims.

The most essential feature of self is self-awareness, which is generated and fostered through self-other interaction and the symbolic processing of information (Lebra 1992:105). Symbolic processing of information in Japanese requires a clear demarcation between the self and others, whereas such a demarcation is less significant in English, which permits sentences like 'Mom wants to go shopping,' without any evidential expressions. Therefore, there is no basis to assume that the Japanese have a less-developed concept of self than do Westerners. And, if the Japanese were 'much more likely than Westerners to operate in groups or at least to see themselves as operating in this way' (Reischauer 1977:125), the ultimate cause should be sought elsewhere than in the cognitive domain.

Concerning evidence for the group model, there is a fundamental problem in methodology: the claims are frequently, if not always, based on anecdotes, without providing a clear argument to what extent such anecdotes are representative of a larger population, as eloquently discussed by Mouer and Sugimoto (1986:130-33). For example, it has been pointed out that the Japanese travel abroad in groups; they bathe together in groups; in Japanese festivals a group of young men carry a portable shrine (ibid.). In the presence of countless counterexamples, such anecdotes cannot be used as evidence.

6. Conclusion

We have considered whether the notion of Japanese self as encoded in the Japanese language is fluid, constantly shifting as the speech situation changes, as claimed by many researchers on Japanese society and its language. All literature I

have reviewed on this topic attempts to drive the Japanese concept of self based on behavioral observations. Whether cognitive notions such as self can be deduced from observed behavior alone is highly questionable. However the Japanese behave non-linguistically, it is indisputable that all competent speakers of Japanese possess a clear and rigid concept of self, without which idiomatic Japanese is impossible. This self is individualistic, not part of a group or interpersonal. Japanese self is the one whose mental states the speaker has direct access to, and that the speaker is privileged to express such mental states without evidential sources — essentially identical with the Cartesian notion of self. This notion is learned by children at an early age, much earlier than they learn proper behavior in society. They learn without explicit instruction that *okaasan wa kaimono ni ikitai* 'Mom wants to go shopping,' a direct statement about the mother's mental state, is anomalous. The Japanese language forces its users to establish a category of self which is disjointed from all other people.

NOTES

* I am indebted to Cindi Sturtz, Anna Wierzbicka, and Ikuko Yuasa for valuable discussions and insightful comments.

¹ This generalization on Japanese preference of group over individual attributes neglects the crucial aspect of self-introduction; one selects introductory statements according to the presumed knowledge of the addressee and the purpose of the subsequent conversations. Usually a Japanese introduces him-/herself in the way Nakane describes when it is already known that the speaker is a scholar or a company employee. Only then, stating one's affiliation becomes natural and relevant. (Because being a company employee is a default in modern urban Japan, the phenomenon mentioned by Nakane is commonly observed.) It sounds strange, however, to say 'I'm from XYZ University' unless the addressee already knows that the speaker is a college student or professor or somehow affiliated with a university.

Smith (1983:82) reports that many Japanese men have two different cards, the business card that includes rank in his organization, business address, and telephone number, and the personal card that bears only his name, home address and telephone number. The latter is said to be given out only to those in the closed circle of one's acquaintances. Although I myself do not know anyone practicing this, it indicates that one's affiliation is not an automatic part of identification.

² Abbreviations: INF = infinitive; TOP = topic; NPAST = nonpast; EVID = evidential; COP = copula; QUOT = quotative; CONJ = conjunctive.

³ Another, equally common, analysis posits three concentric circles: a fixed self as the innermost circle, in-group surrounding the self, and out-group representing the rest of universe. In this analysis, *kure-* is considered to be used for an inward transfer, e.g. from an outsider to the self or an insider, or from an insider to the self.

⁴ The direct representation of subjective experiences (represented consciousness) other than the speaker's own yields what Banfield (1982) refers to as an *unspeakable sentence*, i.e., one which cannot naturally occur in spoken language. Kuroda (1973) calls the style where such expressions appear only with first-person subjects the *reportive style*. In reportive style, only the speaker is entitled to express one's own psychological state. In the *nonreportive style*, by contrast,

a third-person subject is permitted for such expressions.

⁵ Naturally, keen observers have recognized that the 'individual Japanese [is] a very self-conscious person' (Reischauer 1950:143).

⁶ Lebra (1992) proposes three dimensions of self: the interactional self, the inner self, and the boundless self. The interactional self is relative, multiple, and variable in accordance with where and how self stands with respect to others. The inner self is less relative, more stable, fixed, and purely subjective. The boundless self is embedded in the Buddhist version of transcendentalism, disengaging from dichotomies between subject and object, self and other, inner and outer realms, existence and non-existence, and so forth.

⁷ One may wonder why the use of *kure-* is possible in (8) when the destination of the favor of lending money is the speaker's mother, not the speaker herself, because in the theory of relational self, the mother is considered to be included in the speaker's extended self, and yet she is excluded from the potential subject of *omow-*. As mentioned in note 3 above, it is possible to characterize the use of *kure-* differently: a fixed self as the innermost circle, in-group surrounding the self, and out-group representing the rest of universe. This latter analysis does not conflict with the fact that the first person alone can be the subject of psych predicates.

⁸ An example of discursively created images of the 'West' can be drawn from Kunihiro (1976), who claims that the cognitive behavior of the Japanese is different from that of Westerners. According to him, Westerners employ the dualistic Aristotelian logic, but the Japanese do not rely on such reasoning processes.

REFERENCES

- Araki, Hiroyuki. 1973. *Nihonjin no koodooyooshiki* (Japanese behavioral patterns). Tokyo: Kodansha.
- Bachnik, Jane. 1992. *Kejime*: Defining a shifting self in multiple organizational modes. In N. Rosenberger, ed., *Japanese Sense of Self*, 152-72. Cambridge: Cambridge University Press.
- Bachnik, Jane. 1994. Uchi/Soto: Challenging our conceptualizations of self, social order, and language. In J. Bachnik and C. Quinn, eds., *Situated Meaning: Inside and Outside in Japanese Self, Society, and Language*, 3-37. Princeton: Princeton University Press.
- Benedict, Ruth. 1946. *The Chrysanthemum and the Sword*. Boston: Houghton Mifflin.
- Banfield, Ann. 1982. *Unspeakable Sentences: Narration and Representation in the Language of Fiction*. Boston: Routledge and Kegan Paul.
- Boër, Steven, and William Lycan. 1986. *Knowing Who*. Cambridge, Mass.: MIT Press.
- DeVos, George. 1985. Dimensions of self in Japanese culture. In A. Marsella, G. DeVos, and F. Hsu, eds., *Culture and Self*, 141-84. New York: Tavistock.
- Doi, Takeo. 1973. *The Anatomy of Dependence*. Trans. by John Bester. Tokyo: Kodansha International.
- Iwabuchi, Koichi. 1994. Complicit exoticism: Japan and its other. *Continuum: The Australian Journal of Media & Culture* 8:49-82.
- Kondo, Dorinne. 1990. *Crafting Selves: Power, Gender, and Discourses of Identity in a Japanese Workplace*. Chicago: University of Chicago Press.

- Kunihiro, Masao. 1976. The Japanese language and intercultural communication. *The Japan Interpreter* 10: 267-83.
- Kuroda, S.-Y. 1973. Where epistemology, style and grammar meet: A case study from Japanese. In S. Anderson and P. Kiparsky, eds., *Festschrift for Morris Halle*, 377-91. New York: Holt, Rinehart and Winston.
- Lebra, Takie S. 1976. *Japanese Patterns of Behavior*. Honolulu: University of Hawaii Press.
- Lebra, Takie S. 1992. Self in Japanese culture. In N. Rosenberger, ed., *Japanese Sense of Self*, 105-20. Cambridge: Cambridge University Press.
- Miller, Roy A. 1982. *Japan's Modern Myth: The Language and Beyond*. New York: Weather Hill.
- Miner, Richard H. 1980. Orientalism and the study of Japan. *Journal of Asian Studies* 30:507-17.
- Mouer, Ross, and Yoshio Sugimoto. 1986. *Images of Japanese Society: A Study in the Structure of Social Reality*. London: KPI.
- Nakane, Chie. 1970. *Japanese Society*. London: Weidenfeld and Nicolson.
- Reischauer, Edwin. 1950. *The United States and Japan*. Cambridge, Mass.: Harvard University Press.
- Reischauer, Edwin. 1978. *The Japanese*. Cambridge, Mass.: Harvard University Press.
- Rosenberger, Nancy. 1992. Introduction. In N. Rosenberger, ed., *Japanese Sense of Self*, 1-20. Cambridge: Cambridge University Press.
- Shibatani, Masayoshi. 1990. *The Languages of Japan*. Cambridge: Cambridge University Press.
- Smith, Robert. 1983. *Japanese Society: Tradition, Self and the Social Order*. Cambridge: Cambridge University Press.
- Wetzel, Patricia. 1984. Uchi and soto: Social deixis in Japanese. Ph.D. dissertation, Cornell University.
- Wetzel, Patricia. 1994. A movable self: The linguistic indexing of uchi and soto. In J. Bachnik and C. Quinn, eds., *Situated Meaning: Inside and Outside in Japanese Self, Society, and Language*, 74-87. Princeton: Princeton University Press.
- Yoshino, Kosaku. 1992. *Cultural Nationalism in Contemporary Japan*. London: Routledge.

Phonetic Assessment of Tone Spreading

Sung-A Kim

University of Texas at Austin

1. Introduction. The alignment of fundamental frequency (f_0 henceforth) contour with other units in speech is an important issue not only for modeling f_0 contour in speech analysis, but also for theoretical understanding of tone and intonation systems. The f_0 -syllable alignment has been investigated by two groups of researchers from different perspectives. The first line of research has been conducted in tone studies within a framework of theoretical linguistics. Tones can be spread, shifted, or copied, all involving changes of alignment of tones with their host units (Hyman and Schuh 1974, Schuh 1978). Major portions of metrical and autosegmental phonology of tone/accent languages are devoted to various complex tone-syllable alignment issues. The f_0 -syllable alignment has been formalized by resorting to the notion of 'association lines' in autosegmental phonology (Goldsmith 1979, 1990, among others).

The second line of research has been mainly conducted in the field of phonetics. It concerns physical aspects of alignment of f_0 contour with the segmental units in speech. For example, Bruce (1977) found that f_0 realizations in Swedish are characterized by the relative stability of certain tonal targets. However, a review of these tone/accent studies displays a general lack of communication between the groups. This paper tries to fill this gap by investigating one of the most common tone alternations: tone spread.

Tone spread is described as a phonological process whereby "a tone moves beyond its original segmental domain to replace or displace the tone of the following syllable or syllables" (Schuh, 1978: 231). In autosegmental phonology, the tone spread is formalized as a process in which two tone-bearing units share the same tone, as schematically shown in figure 1.



Figure 1. Phonological analysis of tone spreading

Recent works on fundamental frequency realization cast a doubt on the nature of the phonological analysis of tone spreading. It is agreed that the f_0 peak corresponding to high tone tends to be delayed until the later part of the tone-bearing unit or the onset of the next syllable (Silverman and Pierrehumbert 1990, Prieto et al 1995, Arvaniti et al 1998, Kim 1998, among others) because of a sluggish cessation of f_0 movement (Ohala 1978, Fujisaki 1988). Furthermore, phonetic studies on implementation of tones have shown that only a single peak is found in the context of tone spreading (Xu and Wang 1997, Kim 1998). Findings of the phonetic studies raise a serious question of how tone spread is distinguished from f_0 peak delay described above. Will f_0 peak delay and tone spreading be isomorphic in nature? This issue has not been addressed before, simply assuming that phonological spreading would be realized by an f_0 plateau laid upon two tone-bearing units at the phonetic implementation level (Pierrehumbert and Beckman 1988). This paper attempts to answer this question and proposes a phonetic

assessment for the facts previously described as tone spreading in Yoruba and Chichewa.

The remainder of the paper is organized as follows: Facts of the tone alternation in Yoruba and Chichewa are introduced in section 2. A brief sketch of the previous literature on f0 peak realization will be presented in section 3. The experimental method and the result will be addressed in sections 4 and 5, respectively. Implications of this study will be discussed in the conclusion.

2. Tone Alternation in Yoruba and Chichewa. Yoruba has a three-tone system, which is composed of high, mid, and low tones. Among the three tones, mid tone is phonologically inert in the sense that it does not interact with adjacent high or low tones. It is not subject to tone spreading rules (Pulleyblank 1988, Laniran 1992). In comparison, high tone spreads to the next low-toned syllable but not to the mid-toned syllable as shown in (1).

- | | | |
|---------------|--------|----------------------------|
| (1) a. /Láyò/ | [Layô] | 'a personal name for male' |
| /Débò/ | [Debô] | 'a personal name for male' |
| b. /lókò/ | [lókò] | 'a farm' |

In (1a), high tone on the first syllable spreads to the next low-toned, and the second syllable bears a contour tone. On the other hand, such a contour does not occur where the high tone is followed by a mid tone, as shown in (1b). The tone alternation described above occurs regardless of the position in which the high tone occurs in a phrase.

The tone alternation in Chichewa, a language with high and low tones, differs from the tone alternation in Yoruba in the sense that it is position-sensitive. In this language, low tones are assumed to be phonologically inert and thus low tones do not play a role in tone patterns. It is a high tone that spreads and deletes. In Chichewa, high tones generally spread forward one syllable as in (2a) and (2b), but not onto or within the disyllabic phrase-final foot as in (2c). High-toned vowels created by tone spread are underscored in (2).

- | | |
|--|------------------------------|
| (2) a. tinapátsá mwaná <u>dé</u> ⁿ guilli | 'We gave the child a basket' |
| b. chigawé ⁿ <u>gá</u> iichi | 'this terrorist' |
| c. chigawée ⁿ ga | 'terrorist' |

Kanerva (1989) formulates the position-sensitive tone spread rule for the data in (2) as nonfinal high tone-spreading as depicted in (3).

- (3) Nonfinal doubling:
- $$\begin{array}{c}
 H \\
 | \quad \diagdown \\
 \mu \quad \mu_i \quad \mu \quad \mu
 \end{array}$$
- condition: μ_i is not in phrase-final foot.

According to this rule, a high tone spreads onto the following mora if the mora is not in the phrase-final foot. In (2a) and (2b), there occurs high tone spreading because the target mora is not in the phrase-final foot. On the other hand, the high tone in (2c) does not spread to the following mora since the target mora is included in the phrase-final foot. The same tone alternation is also observed by Mtenje

(1988). The phonological analyses of the tone alternation in Chichewa can be summarized as follows: Tone spreading occurs in pre-penult positions, while it does not in penult positions in Chichewa.

Before moving on to the next section, let us briefly discuss a problem with the spreading analysis given in (3). The problem arises from the special reference to the phrase-final element. According to the phonological analysis, the spreading rule must refer three moras ahead before tone-spreading takes place. It makes this rule typologically peculiar in the sense that local tone spreading is triggered by an element at the end of a phrase and by counting the syllables from the end of the phrase. This pattern leads to a violation of the locality condition in an otherwise well-motivated generalization. Phonological rules are usually assumed to be subject to conditions of locality where the trigger and the target should be structurally adjacent to each other (McCarthy and Prince 1986, Odden 1994). The tone-spreading as described is a violation of this well-motivated constraint.

The summary of tone alternations in the two languages is tabulated in (4). In Yoruba, high tone spreading is assumed to occur in a high-low context (henceforth, HL context), but not in a high-mid context (henceforth, HM context). In Chichewa, high tone spreading is supposed to occur in pre-penult positions but not in penult positions.

Languages	Locality	Phrase Position	So-called Spreading Case
Yoruba	local	Insensitive	HL contexts
Chichewa	non-local	Sensitive: Spreading only in pre-penult positions	Pre-penult positions

(4) Summary of Phonological Analysis of Tone Alternation. I explored an instrumental analysis of the facts previously described as tone spreading, which suggests another description of the facts and another analysis. This ultimately proposes that the position-sensitive tone alternations are not tone spreading, but a by-product of f0 peak alignment.

3. Theoretical Background: F0 Peak Delay. Compared to the unusualness of the phonological analysis in Chichewa given in section 2, the position-sensitive processes are not typologically unusual, from the perspective of the phonetic timing of the f0 peak. A number of studies on the phonetic realization of accent and tone have found that phonetic prominence, specifically pitch prominence, may not always align with the accented or tone-bearing syllable (Steele 1986, Silverman and Pierrehumbert 1990 for English, Prieto et al. 1995 for Mexican Spanish, Liberman et al. 1993 for Igbo, and Liberman 1996 for Yoruba).

Also, it has already been well established that the f0 delay does not occur where the syllable is close to a prosodic phrase edge. Steele (1986) and Silverman and Pierrehumbert (1990) show that two factors, rhyme duration and upcoming prosodic contexts, are the main source of peak location variation in English. That is, when a vowel is lengthened because of slow speech, the f0 peak is correspondingly delayed. Given an identical phrase position, there is a positive correlation between vowel duration and f0 peak delay relative to the vowel onset. In contrast to this, the f0 peak is aligned early in the syllable where the syllable is close to a prosodic edge in English. Prieto et al. (1995) report a similar result for Mexican Spanish. In addition, they observe that the same finding is obtained when

f0 peak delay is measured relative to accented syllable onset as well as vowel/rhyme onset. Kim (1998) reveals that the distance from syllable onset to f0 peak highly correlates with the duration of tone-bearing syllables in Chichewa. Of particular interest is the lack of correlation between the syllable duration and distance between syllable offset and f0 peak (offset-to-f0 peak) in Chichewa. This means that offset-to-f0 peak does not systematically vary with any consistency. In other words, the f0 peak simply stays close to the syllable offset. The high correlation between f0 peak relative to the tone-bearing syllable onset and the syllable duration indicates that the f0 peak is moving closely with the syllable offset. Let us call this pattern 'peak delay.' As Ohala (1978) and Fujisaki (1988) argue, the f0 peak delay appears to be due to a function of sluggish cessation of an f0 movement. In other words, f0 is a function of the strain of the muscles such as cricothyroid and sternohyoid, the mass of the thyroid cartilages and stiffness of the cricothyroid joint. Even if the neural commands for producing a pitch target is issued simultaneously with those for producing the syllable that carries it, the pitch target would be attained more slowly than the segmental targets. Therefore, it is possible that f0 peak is realized on the following syllable.

The peak delay pattern in f0 represents a regular timing relation between the laryngeal gesture that leads to the f0 peak and landmarks in the syllable (syllable onset, syllable offset). Since the groundbreaking work of Bruce (1977) on Swedish, it has been known that at least some f0 patterns are characterized by relative phonetic stability of certain tonal targets. Bruce finds that the local f0 peak corresponding to Swedish word accent is constantly aligned with the segmental material and the f0 value is constant across multiple repetitions of the same utterance by the same speaker. Bruce interprets this to mean that, for the Swedish accentual distinction, "reaching a certain pitch level at a particular point in time is the important thing, not the movement (rise or fall) itself (1977:132)." The same kind of stability is observed elsewhere as well. For instance, Huffman (1993) shows that syllable landmarks are important in the timing of velar gestures. The regular and proportional timing between gestures of different articulators is also founded in work within the framework of task dynamics (Tuller and Kelso 1984 and Browman and Goldstein 1990 among others).

4. Experimental Design. The experiment was designed to compare the f0 alignment of the so-called spreading case to the non-spreading case in each language. Consider the sentences in (5) from the Yoruba corpus used in the experiment.

(5) Corpus in Yoruba

- | | |
|--------------------------------------|---|
| a. <u>Ó</u> <u>m̀dómò</u> lò ìyén | b. <u>Mámámíá</u> lè lò ìyén. |
| He intentionally use that one | mother my is able to use that one |
| "He intentionally uses that one." | "My mother is able to use that one." |

The two sentences in (5) are almost identical to each other except for the target words, indicated by underlines. The underscored word in (5a) contains a sequence of a high tone followed by a low tone (henceforth, HL), while the one in (5b) has a sequence of high tone followed by a mid tone (henceforth, HM). Therefore, the target word in (5a) corresponds to the spreading case, whereas the one in (5b) corresponds to the non-spreading case. If there is tone spreading only in (5a), clear differences in the f0 realization between the spreading and non-spreading cases are

predicted by the phonological analysis. The predictions made by phonological analysis of tone-spreading are summarized in (6).

	F0 realization	Number of Associated Syllable(s)
Spreading	F0 plateau (f0 peak laid upon two syllables)	2
Non-spreading	no plateau (single f0 peak)	1

(6) Prediction Borne out of Phonological Analysis of Tone Spreading.

A crucial difference between spreading and non-spreading cases is whether there is a tone target in the syllable next to the underlying high tone-bearing syllable. As shown at (6), the so-called spreading case is characterized by the existence of an additional tone target. In order to test the prediction born out of the phonological analysis of tone spreading, I compared the f0 peak alignment in the two target words in (5). If there is a tone spread in pre-penult positions, then we should find f0 peak to be timed with the syllable following the high tone-bearing syllable, while the f0 peak is timed with the high tone-bearing syllable itself if there is no tone spread.

The same reasoning applies to the tone spreading case in Chichewa. For Chichewa, the sentence in (7) was used in the experiment. Notice that the sentence has four high tones. Among those, the third high tone and the fourth high tone underlined are the main concern here. It is the fourth high tone that is located in the phrase-final foot, while the third high tone is not. For the sake of convenience, let us call the third and the fourth high tones the pre-penult and penult tones respectively. In order to minimize segmentally induced perturbation on f0, target words indicated by italics are composed of sonorants, except the *k* in *kun'ér'a*. Also notice that the pre-penult and the penult tone-bearing vowels are identical (i.e., *e*) to control for the intrinsic f0 of vowels.

- (7) Ml^oda áma yená kun' é^h *a*.
 The watchman must to goof off
 "The watchman must goof off."

If there is a spreading only in pre-penult positions in Chichewa, then the f0 alignment in the pre-penult positions should differ from the one in penult positions. If there is no tone spread, f0 is expected to be determined by reference to the tone-bearing syllable regardless of the positions.

In the experiment, a male Chichewa speaker and a female Yoruba speaker participated at the recording. The Chichewa speaker and the Yoruba speaker uttered the sentence in (5) and (7) respectively. To induce a broad range of f0 values and syllable duration, the speaker was asked to vary loudness and speech rate. There were three conditions with respect to loudness as used in Liberman et al. (1993). A total of 576 tokens (1 speaker x 2 sentence types x 3 loudness levels x 2 speech rates x 2 target words x 24 repetitions=576) was obtained for Chichewa. A total of 120 tokens was obtained for Yoruba. The speech signal recorded on a digital audio tape-recorder was digitized at a sampling rate of 22 kHz and segmented from waveform and spectrogram display, using Sound Scope, a Macintosh-based sound analysis by GW Instruments.

The following measurements were taken for the two target words in spreading and non-spreading cases in each language.

(8) Measurement Points

- The onset and offset of the tone-bearing syllable,
- the onset and offset of the tone-bearing rhyme,
- the onset and offset of the syllable following the tone-bearing syllable,
- the onset and offset of the rhyme in the syllable following the tone-bearing syllable,
- the f0 peak corresponding to high tones in both spreading and non-spreading cases.

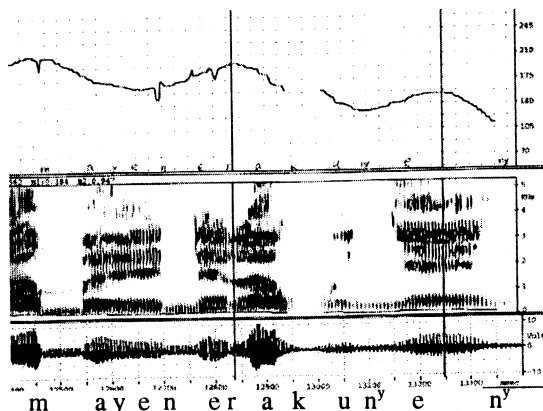


Figure 2. A pitch track of pre-penultimate and penultimate high tones in Chichewa.

Figure 2 shows the waveform display, f0 contour, and spectrogram corresponding to the target words *amayenera* and *kun'en'a*. Segmental transcription is given at the bottom of Figure 2.

The following null hypotheses are tested in the study. If those null hypotheses are rejected, then it can be said that tone spreading analysis is confirmed in each language.

(9) The Hypotheses:

- In Yoruba, the temporal locations of the f0 peak are determined by reference to high tone-bearing syllable in both HL and HM contexts.
- In Chichewa, the temporal locations of the f0 peak are determined by reference to the high tone-bearing syllable in both pre-penult and penult positions.

In order to test these hypotheses, I compared models of the temporal relation using multiple regression. The relevant variables used in the multiple regression are given in (10):

(10) Variables

- a. Peak delay: Temporal distance between f0 peak and the onset of the tone-bearing syllable.
- b. T-Syllable duration: Duration of the tone-bearing syllable.
- c. N-syllable duration: Duration of the syllable following the tone-bearing syllable.

F0 peak relative to the onset of the tone-bearing syllable in (10a) is the dependent variable and position and the syllable duration in (10b) and (10c) are the independent variables in the multiple regression. Penult position and pre-penult position are coded as 1 and 0 respectively.

5. Results and Discussion.

5.1. Yoruba. First, let us discuss the results of Yoruba. With regard to the f0 peak in pitch tracks, only a single f0 peak is observed in both HM and HL contexts. There was no f0 plateau laid upon the two syllables in general. The f0 peak is delayed until the syllable following the high tone-bearing syllable. As a result, the longer the high-toned syllable, the longer the f0 peak is delayed. At first glance, it appears that there is no difference in f0 timing between the HL context (i.e., spreading case) and HM context (i.e., non-spreading cases) with regard to the f0 peak alignment. A close examination, however, reveals that the two contexts differ from each other in the syllable with which the f0 peak has a constant relation. Consider the table in (11) where the results of simple regression models in HM context are summarized. The one in (11a) indicates the regression model where the peak is assumed to be timed with the tone-bearing syllable itself, while the one in (11b) represents the regression model where the peak is supposed to be timed with the syllable next to the tone-bearing syllable. We can compare models using relative measures of goodness of fit. One of such measures is the Pearson R^2 values, which is indicated in the second column in the table. For example, the R^2 value of 0.684 indicates that the model accounts for 68% of the variation in the data.

a. Peak = $0.049 + 0.676 * T_{\text{syll}}$	R^2 : .684	HM context (non-spread)
b. Peak = $0.001 + 0.022 * N_{\text{syll}}$	R^2 : .002	HM context (non-spread)

(11) Simple regression models for the prediction of f0 peak location

In the table in (11), higher R^2 value in (11a) indicates that f0 peak is timed with the tone-bearing syllable in non-spreading context. The extremely low R^2 value in (11b) shows that the f0 peak has almost no relation with the syllable following the tone-bearing syllable.

In the spreading context, however, the opposite result is obtained, as shown in (12). The equation in (12b) shows the result of the case where the f0 peak is assumed to be timed with the next syllable in the spreading context. This equation model has higher R^2 value than the one in (12a). This means that the temporal location of f0 peak is best predicted by reference to the syllable next to the tone-bearing syllable in spreading cases.

a. $\text{Peak} = 0.134 + 0.901 * \text{Tsyll}$	$R^2: .592$	HL context (spread)
b. $\text{Peak} = -0.027 + 0.823 * \text{Nsyll}$	$R^2: .621$	HL context (spread)

(12) Simple regression models for the prediction of f0 peak location

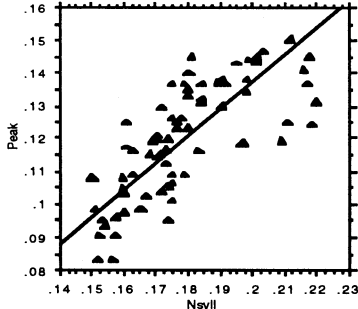
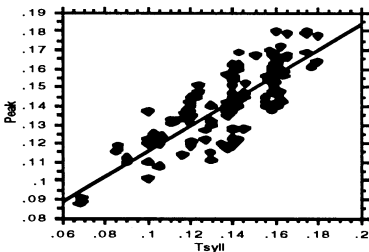
If we compare the tables in (11) and (12), it is clear that there is an asymmetry between HM and HL contexts. In a non-spreading context, f0 peak has a constant relation with the high tone-bearing syllable itself. On the other hand, in a spreading context, it is constantly timed with the syllable next to the tone-bearing syllable. The equation models in (11a) and (12b) are graphically shown in (13a) and (13b).

(13) Regression equations in non-spreading and spreading cases in Yoruba

a. HM context

b. HL context

Regression Plot



To sum up, Yoruba results appear to support the phonological analysis of tone spreading as they display a difference in f0 timing between spreading and non-spreading cases.

5.2. Chichewa. In the previous section, Yoruba results show that there is a clear difference in f0 timing between spreading and non-spreading cases. In comparison, Chichewa results show that there is no difference in f0 timing between spreading and non-spreading cases. F0 peak is timed with the high tone-bearing syllable in both cases in Chichewa. The table in (14) shows two regression models in pre-penult positions whereas the one in (15) contains two regression models in penult positions. It should be noted that regression equations in (14a) and (15a) have higher R^2 values than the ones in (14b) and (15b).

a. $\text{Peak} = .003 + 1.355 * \text{Tsyll}$	$R^2: .775$	pre-penult (i.e. spreading)
b. $\text{Peak} = .028 + .968 * \text{Nsyll}$	$R^2: .381$	pre-penult (i.e. spreading)

(14) Simple regression models for the prediction of f0 peak location

a. $\text{Peak} = -.027 + .637 * \text{Tsyll}$	$R^2: .616$	penult (i.e. non-spreading)
b. $\text{Peak} = .063 + .293 * \text{Nsyll}$	$R^2: .217$	penult (i.e. non-spreading)

(15) Simple regression models for the prediction of f0 peak location

If there is tone spreading only in penult positions, then it is predicted that the f_0 peak in pre-penult positions has the most constant relations with the syllable next to the high tone-bearing syllable, as we found in Yoruba cases. Notice that the R^2 value in the N-syllable based model (i.e., equations in 14b and 15b) is much lower than the one in the T-syllable based model (i.e., equations in 14a and 15a). When the peak delay is plotted against the duration of the high tone-bearing syllable as in figure 3, quite high R^2 values (0.775 in pre-penult positions and 0.616 in penult position) were obtained in both pre-penult and penult position. Therefore, the tone spreading analysis is rejected in Chichewa.

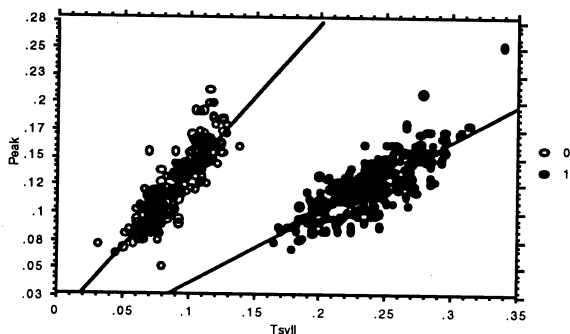


Figure 3. Peak delay as a function of duration of tone-bearing syllable in Chichewa.

The findings of this study are important both empirically and theoretically. Empirically, they provide instrumental data about the little-studied f_0 peak delay phenomenon in African tone languages. Theoretically, the experimental results for Chichewa indicate that the tone alternation in the language deserves a phonetic account rather than a phonological analysis. The phonetic account does not require that Chichewa be an exception to phonological constraint on locality. More importantly, the contrast between tone alternations in the two languages provides an empirical ground to tease apart phonetic implementation of tone (e.g. tone alternation in Chichewa) from true phonological rules of tone (e.g. tone alternation in Yoruba).

Acknowledgments

I would like to thank Larry Hyman, Björn Lindblom, Scott Myers, Tony Woodbury, and Yi Xu for helpful discussions and comments. I also thank Amalia Arvaniti for sending her papers. Any error is mine.

References

- ARVANITI, R. LADD and INEKE MENNEN. 1998. Stability of tone alignment: The case of Greek prenuclear accents. To appear in *Journal of Phonetics*, 26.1
- BROWMAN, CATHERINE and LOUIS GOLDSTEIN. 1990. Tiers in articulatory phonology, with some implications for casual speech. *Papers in laboratory phonology I*, ed. by John Kingston and Mary Beckman, 341-376. Cambridge, MA: Cambridge University Press
- BRUCE, GOASTA. 1977. *Swedish word accents in sentence perspective*. Lund: Gleerup.

- FUJISAKI, HIORTA. 1988. A note on the physiological and physical basis for the phrase and accent components in the voice fundamental contour. *Vocal physiology: Voice production, mechanisms and functions*, ed. by Osamu Fujimura, 347-355. New York: Raven Press.
- GOLDSMITH, JOHN. 1979. *Autosegmental phonology*. New York: Garland.
- _____. 1990. *Autosegmental and metrical phonology*. Oxford: Blackwell Publishers.
- HUFFMAN, MARY. 1993. Phonetic patterns of nasalization and implications for feature specification. *Phonetics and phonology 5: Nasals, nasalization and the velum*, ed. by Mary Huffman and R. Krakow, 303-327. New York: Academic Press.
- HYMAN, LARRY M, and RUSSELL G. SCHUH. 1974. Universals of tone rules: Evidence from west Africa. *Linguistic Inquiry* 5.81-115.
- KANERVA, JONNI. 1989. *Focus and phrasing in Chichewa phonology*. Stanford, CA: Stanford dissertation.
- KIM, SUNG-A. 1998. Position-sensitive tone alternation in Chichewa: Phonological analysis vs. phonetic timing. To appear in the proceedings of the 34th Chicago Linguistic Society.
- LANIRAN, YETUNDE. 1992. *Intonation in tone languages: The phonetic implementation of tones in Yoruba*. Ithaca, NY: Cornell University dissertation.
- LIBERMAN, MARK. 1996. Tone, accent and relative prominence. Talk given at the University of Texas at Austin.
- _____; J. Schults; Soonhyun Hong; and Vincent Okeke. 1993. The phonetic interpretation of tone in Igbo. *Phonetica* 50.147-160.
- MCCARTHY, JOHN, and ALAN PRINCE. 1986. *Prosodic morphology*. Amherst and Waltham, MA: University of Massachusetts, and Brandeis University, MS.
- MTENJE, AL. 1988. On tone and tone transfer in Chichewa reduplication. *Linguistics* 26.125-155.
- ODDEN, DAVID. 1994. Adjacency parameters in phonology. *Language* 70.2.289-330.
- OHALA, JOHN J. 1978. The production of tone. *Tone: A linguistic survey*, ed. by Victoria Fromkin. 5-39. New York: Academic Press.
- PIERREHUMBERT, JANET and MARY BECKMAN. 1988. *Japanese tone structure*. Cambridge, MA: MIT Press.
- PRIETO, PILAR; JAN VAN SANTEN; and JULIA HIRSCHBERG. 1995. Tonal alignment patterns in Spanish. *Journal of Phonetics* 23.429-451.
- PULLEYBLANK, DOUGLAS. 1988. Vocalic underspecification in Yoruba. *Linguistic Inquiry* 19.233-270.
- SCHUCH, RUSSELL G. 1978. Tone rules. *Tone: A linguistic survey*, ed. by Victoria Fromkin. 221-256. New York: Academic Press.
- SILVERMAN, KIM, and JANET PIERREHUMBERT. 1990. The timing of prenuclear high accents in English. *Papers in laboratory phonology I*, ed. by John Kingston and Mary Beckman, 72-106. Cambridge, MA: Cambridge University Press.
- STEELE, S.A. 1986. Nuclear accent f0 peak location: Effects of vowel, rate, and the number of syllables following. *Journal of Acoustical Society of America Supplement* 1.80; s 51.
- TULLER, BETTY, and SCOTT KELSO. 1984. The timing of articulatory gestures: Evidence for relational invariants. *Journal of the Acoustical Society of America* 76.1030-1076.
- XU, YI, and EMILY WANG. 1997. *Intonation contours decomposed: From bottom up*. Evanston, IL: Northwestern University, MS.

Motivating Morpho-Syntactic Changes in Turkic Subordination*

Jaklin Kornfilt
Syracuse University

1. Introduction. It has often been claimed that diachronic change and first language acquisition reflect similar principles and tendencies, since (presumably) the former is to a great part due to the latter. As Kiparsky (1968:175) states, "the transmission of language is discontinuous, and a language is recreated by each child on the basis of the speech data it hears." Thus, a child acquiring the language of her parents might construct a grammar that is different than the one her parents' generation had internalized. For example, "[a] highly constrained rule may be generalized, or a difficult rule may fail to be learned." (Jeffers & Lehiste 1979: 98) Or, a given stage in the development of a language might have a high degree of exceptionality with respect to certain syntactic categories (cf. Lightfoot 1979) or with respect to certain rules (cf. Kiparsky 1968 and 1978), making the lexicon or the rule system opaque. Children would then construct a grammar which would be more transparent, with concomitant changes in the lexicon and/or in the rule system.

The claim that acquisition influences diachronic change is illustrated very clearly for Turkic languages in work by Slobin (in particular, in Slobin 1986).¹ The following quote illustrates this point of view:

Those parts of a language that are most stable over time should be acquired relatively early in the course of development and should be relatively easy to process. Conversely, the parts of the grammar most susceptible to change – either through internal change or under the influence of borrowing from other languages – should be those parts which are acquired late and which are relatively difficult to process. (Slobin 1986: 273)

In this paper, I shall assume the essential correctness of the general claim, without however accepting the conclusions and generalizations about the Turkic languages insofar as the proposed explanations for the putative diachronic changes are concerned. More specifically, mine differ from Slobin's proposals substantially when considering his particular claims about the historical developments in the Turkic languages, and I establish tendencies of change that are diametrically opposed to the psycholinguistic principles which Slobin claims underlie both historical changes in the Turkic languages and the acquisition of Turkish.

The relevant psycholinguistic principles are: 1. "If a semantic configuration can be expressed by a single, unitary form (synthetic expression) or by a combination of several separate forms (analytic expression), prefer the analytic expression." (Slobin 1986:278-79) 2. "... If a clause has to be reduced, rearranged, or otherwise deformed when not functioning as a canonical main clause ..., attempt to use or approximate the full or canonical form of the clause." (Slobin 1986:279) In addition, the assumption is made that left-branching structures are more difficult to parse than right-branching ones, and that therefore the former tend to be replaced by the latter. These principles and tendencies are designed to explain various phenomena of diachrony as well as of first language

acquisition. In this paper, I shall focus primarily on diachronic phenomena. I shall conclude that at least some of the diachronic changes observed go counter to principles 1 and 2, and I shall hypothesize that the changes are actually motivated by a tendency which is opposed to principle 2, i.e. a tendency towards overtly differentiating between main and subordinate clauses, irrespective of the means that are used to achieve such differentiation.

2. Facts of diachrony, typology and acquisition. There are primarily two phenomena concerning language change which are discussed in Slobin (1986). One is the observation that in many Turkic languages spoken in the former Soviet Union there has been a strong tendency to form subordinate clauses as in Russian (and in Indo-European in general), with free (i.e. non-affixal) subordinate particles and fully finite clauses. Note that subordinate clauses in Turkic are typically "nominalized", i.e. are not fully finite. A few representative examples from Modern Standard Turkish follow to illustrate this claim.

In fully finite clauses, Turkish has a wide variety of tense and/or aspect markers that are suffixed to the verb stem and produce so-called simple conjugation forms. These suffixes are followed by subject agreement markers drawn from a number of paradigms; the choice of these agreement markers is dictated by the tense/aspect suffix. The subject of these finite clauses is in the Nominative case which is not marked morphologically.

- (1) ben her gün sinema-ya gid-er -im
 I (NOM) everyday cinema -DAT go -AOR -1SG
 'I go to the movies every day'
- (2) ben yarın sinema-ya gid-eceğ-im
 I (NOM) tomorrow cinema-DATgo -FUT -1SG
 'I shall go to the movies tomorrow'

In subordinate clauses, we find essentially two nominalization markers on the stems, generating gerundive clauses. One type is roughly a factive gerundive (FG), while the other corresponds roughly to a subjunctive (which I shall gloss as "action nominal" [AN]).

- (3) ben [Hasan-in sinema-ya git-me -sin-i]
 I Hasan-GENcinema-DAT go -AN-3SG-ACC
 iste -di -m
 want -PST-1SG
 'I wanted for Hasan to go to the movies'
- (4) ben [Hasan-in sinema-ya git-tiğ-in -i]duy-du -m
 I Hasan-GENcinema-DATgo-FG -3SG-ACC hear-PST-1SG
 'I heard that Hasan went to the movies'

Both gerundive patterns share the following properties: their subject is marked for the Genitive case (in contrast to the nominative subject of finite clauses), the subject agreement suffixes come from a nominal possessive agreement paradigm

(rather than from any one of the verbal paradigms as in finite clauses), and the gerundive suffixes are placed in the same morphological slot as the tense/aspect suffixes in finite verbs. As a consequence of the last property, such gerundive verbs lack the numerous tense/aspect differentiations that can be expressed on simple finite verbs. While such semantic distinctions can be made in gerundive clauses, as well, periphrastic forms must be used, since the simple gerundive verb cannot express these distinctions by itself, due to the morphological reasons mentioned above.

From the perspective of Slobin's operational principles, we assess such gerundive clauses as follows: If nominalization affixes are viewed as markers of subordination, the resulting forms are not analytic, thus violating principle 1. Furthermore, in such clauses, the subject (when existent) is not in the Nominative case, as it is in fully finite main clauses, but in the Genitive; the gerundive markers show up instead of (and in the morphological slots for) the finite tense and aspect suffixes; furthermore, the subject agreement markers on the gerundive are from a different paradigm than the verbal agreement markers found on fully finite predicates. All of these factors conspire to make subordinate clauses quite different from main clauses, thus violating principle 2. Based on these operational principles, we would therefore expect that such clauses would be difficult to process; the change toward Indo-European patterns in their complex constructions in the Turkic languages used in the formerly Soviet areas would then be explained by these principles rather than, say, by direct political pressures. Here, I shall address this particular claim only in passing. Note also that we would expect such nominalized clauses to be difficult to acquire by children; as a matter of fact, it appears that Turkish children do acquire subordination and complementation about two or more years later than their peers who acquire Indo-European languages like English, French and Russian (cf. Slobin 1986: 273-274, 277).

A second point concerning diachronic change is made by Slobin with respect to a particular type of complex construction, namely relative clauses. He claims that these are particularly complex and complicated in Turkish, and that at least some of these complexities and difficulties have been simplified in most of the other Turkic languages. It is this construction which I will be addressing primarily in this paper, and I will show that, firstly, (most of) the supposed simplifications and neutralizations in the other Turkic languages are not new developments but rather are reflexes of older "simple" constructions, and that secondly, the supposed complexities of Turkish are innovations rather than reflexes of "old" patterns as assumed in Slobin's work. If so, the patterns in question cannot be viewed as universally and cognitively difficult—at least not within an approach like Slobin's which seeks to explain directionality in historical change via general psycholinguistic principles. To illustrate claim and counter-claim, let me turn to a discussion of relative clauses.

2.1. Relative clauses in Turkish. Modern Standard Turkish has relative clause constructions which exhibit properties similar to the complex constructions which we saw above. Those constructions have nominalized subordinate clauses which have argument function. In addition to whatever cognitive complexity is contributed by the morphological as well as the syntactic properties of nominalized clauses, relative clauses exhibit the following complication: relative clauses whose target is a subject or part of a subject are of a different form than those whose target is a non-subject.² This dichotomy is illustrated by the following examples.

Relative clauses whose target is a subject have modifier clauses headed by a "subject participle" (SP):

- (5) su -yu iç -en misafir
 water -Acc. drink-SP guest
 'the guest who drinks the water'
- (6) su -yu iç -miş (ol-an) misafir
 water -Acc. drink-PERFP be-SP guest
 'the guest who has drunk the water'

Relative clauses whose target is a non-subject have modifier clauses headed by an "object participle" (OP). While this terminology is one used in most of the recent literature on the topic of Turkish relative clauses (cf. Underhill 1972, Hankamer & Knecht 1976), this so-called object participle is nothing other than the familiar factive gerundive, which is not surprising from a typological perspective, since the canonical relative clause pattern is one that involves indicatives rather than subjunctives or infinitives.

- (7) misafir -in iç -tiğ-i su
 guest -GEN drink-OP -3SG water
 'The water which the guest drinks'
- (8) misafir -in iç -miş *(ol-duğ -u) su
 guest -GEN drink-PERFP be -OP -3SG water
 'The water which the guest has drunk'

Notice that relative clauses whose target is a non-subject are rather different from their counterparts whose target is a subject. The modifier clause has a subject, and that subject is overtly marked for genitive case. Furthermore, the participle bears overt agreement morphology with its subject, while the subject participle does not. Even where the same perfective participle is used in the two constructions, there is a difference: the perfective participle can be used by itself, as a non-agreeing predicate, in subject relative clauses as in (7). In contrast, it requires the presence of the copula bearing the appropriate agreement and non-subject participle forms when used in non-subject relative clauses as in (8).

Slobin points out that the non-subject type of relative constructions are "exceptionally rare in Turkish", a fact which he attributes to its complexity (Slobin 1986: 284). This complexity would be a function of a violation of his principle 2 by this construction, namely the fact that the modifier clause in non-object relative clauses is even less similar to a canonical main clause than a subject relative clause. Hence, while both types of relative clauses would be difficult to parse and to acquire (since they are both synthetic rather than analytic, thus violating principle 1, and are both left-branching), the subject relative construction would be evaluated as somewhat simpler, and one would therefore predict that it would be more stable diachronically than non-subject relative clauses. Furthermore, the choice that must be made between two participial forms is said to place a heavy burden on the speaker, and that therefore other Turkic languages show a "tendency to neutralize the surface distinction between subject

and nonsubject relatives, with use of a single form for all types" (Slobin 1986: 286). This neutralization towards a single form would have to develop in favor of the form which is claimed to be cognitively simpler, namely the subject participle and its concomitant syntactic properties: no genitive case, no non-verbal agreement form on the verb.

Indeed, this is exactly what Slobin claims to have happened:

Significantly, the form chosen, repeatedly, is the -An form.
(Slobin 1986: 286)

Some of the examples given in Slobin (1986) to illustrate this diachronic and typological claim are repeated below (I shall use the general gloss of participle (P), where the distinction between subject and non-subject participle is not clear or not relevant):

- (9) Turkmen: gel -en adam
 come-P man

'The man who came' (Slobin 1986: 287; his 19)

This example would be equivalent in Turkish.

- (10) Turkmen: men-in yaz -an kitab-ım
 I -Gen. write-P book-1SG

'The book that I wrote' (Slobin 1986: 287; his 19)

The Turkish equivalent would be as follows:

- (11) ben-im yaz -dığ-ım kitap
 I -Gen write-OP -1SG book
 'The book that I wrote'

We note that, indeed, Turkmen appears to have "neutralized" the two participial forms into one, and that this single form is the same as the Turkish subject participle and not its non-subject participle.

In the next section, I turn to older stages of Turkic, and I show that the morphological complexity of relative clauses in Modern Turkish is not a reflex of previous stages of Turkic which got neutralized in a number of other Turkic languages, but rather that they are innovations. The patterns in other Turkic languages, as exemplified by Turkmen, appear to be in part reflexes of older stages and in part due to individual developments, as we shall see later. In any event, they are not due to the kind of "neutralization" posited by Slobin. In what follows, I shall limit myself to relative clauses.

3. Old Turkic relative clause constructions. In this section, we shall see that the dichotomy between subject and non-subject relative clauses did not exist in the earliest Turkic texts that are available to us. While there were a number of participial morphemes did exist in relative clauses, they had primarily aspectual

(as well as secondarily temporal) functions, similar to corresponding (or sometimes identical) morphemes in main clauses. The Old Turkic documents consist of stone monuments from the 7th century AD (usually referred to as the Orkhon monuments) and Old Uighur manuscripts from the 8th century AD.

The assumption is usually made in Turkological literature that in constructions corresponding to relative clauses, the participles are "indifferent" between active and passive (e.g. v. Gabain 1941:76). This means that the syntactic and morphological properties of relative clauses are the same, irrespective of whether the target of relativization is a subject or a non-subject. In the examples that follow, I have noted between parentheses whether they correspond to a modern Turkish subject relative (SP) or to a non-subject relative (OP).³ In either instance, there is no agreement morpheme on the participle, nor is the subject of the modifier clause (in those examples that do have a subject) marked with the genitive case (although Old Turkic did have a genitive marker for possessors).⁴

- (12) il tut -sıq yir
tribes rule -P place
'the territory from which the tribes should be ruled' ("OP")
(KTS 4; Tekin 1968:176) [future/necessitative]
- (13) tinsi ogli ayti -gma tag
Tinsi ogli call -P mountain
'the mountain called Tinsi ogli' ("OP")
(TII S 2-3; Tekin 1968:176) [factive, potential]
- (14) il biri -gmä täNri
state grant -P deity
'Heaven who has granted the state' ("SP")
(KT E 25; Tekin 1968: 176)
- (15) kör-ür köz-üm kör-mäz tag bol -ti
see -P(=AOR) eye -1SG see -P(=NEG-AOR) like become -PST 'my
seeing eyes became as if they did not see' ("SP")
(KT N 10; Tekin 1968: 177) [present state, action]
- (16) sigun tart -ar qaNli
maral pull -P(=Aor.) cart
'The cart pulled by a maral; the cart which a maral pulled'
(Hüen-tsang, translations; v. Gabain 1941:76)
- (17) äcü -miz ...tut -miş yir
predecessor -1PL rule -P place
'The place(s) which our predecessors ruled; the place(s)
which were ruled by our predecessors' ("OP")
(KT E 19; Tekin 1968: 179) [perfective]

- (18) *igid -miş qagan*
 nourish -P kagan
 'The kagan who has nourished (you)' ("SP")
 (BK E 22; Tekin 1968: 179) [perfective]
- (19) *kāl -mä -dük öd*
 come-NEG -P time
 'The time that hasn't come; the future' ("SP")
 (v. Gabain 1941:76) [factive; past]
- (20) *bar-duq yir -dä*
 go -P place -LOC
 'In the lands where you went' ("OP") [factive; past]
 (KT E 24; BK E 20; Tekin 1968: 178-79)

We note that in these examples, a rather large number of participle morphemes is displayed. These clearly express a variety of aspectual and also temporal functions. Interestingly enough, two of these morphemes are also found in Modern Turkish relative clauses: *-miş* and *-duq*. The first, as we saw earlier, can function as a subject participle (when it is the only verbal form used), while the latter is the canonical non-subject participle. Yet, in the Old Turkic data, either form could be used in either function (note, in particular, examples (17) versus (18) on the one hand, and (19) versus (20), on the other. What is constant in the functions of these morphemes is therefore their aspectual, modal and/or temporal functions, and not any putative function as a clue for the grammatical relation of the relativization target.

Comparing the grammar of relative clauses in Old Turkic with that of Modern Turkish, we note that the latter introduced a number of innovations: reduction in the number of participle markers; shift of the main function of these markers from modal/aspectual towards a function as a clue for the grammatical relation of the relativization target; genitive marking on the embedded subject (when there is one); overt agreement markers on the participle (when there is a subject to agree with, i.e. when the target is a non-subject).

We conclude, then, that most of the properties exhibited by relative clauses in Modern Turkish that are claimed in Slobin (1986) to be perceptually difficult, which make the construction difficult to acquire by children, and which are also claimed to be historically unstable are actually later innovations. The older cognate forms are much "simpler" and conform to principle 2, i.e. they are much more similar to canonical main clauses than are their Modern Turkish counterparts. Hence, the system proposed in Slobin (1986) would make the prediction that Old Turkic relative clauses should have been, relatively speaking, stable, and that constructions like those in Modern Turkish should not have arisen in the course of later developments—obviously an incorrect prediction.

I now turn to a brief discussion of relative clause constructions in other Turkic languages; however, due to space limitations, I shall only look at Turkmen as a representative.

4. Other (contemporary) Turkic languages. As we saw earlier, Turkmen relative clauses have only one participial form, and that is the form used as a

subject participle in Modern Turkish—the form claimed to be "simpler", in that it does not require agreement morphology. The following example is representative; it is quite similar to the one cited from Slobin (1986) earlier, and it corresponds to a Turkish non-subject relative clause.

- (21) Rus qoşunın -iN yerleş -en yer -in -e
 Russian soldiers -Gen.settle -P place -3SG -DAT
 Kulikovomeydanı diyilyer
 Kulikovofield called
 'The place where the Russian soldiers settled is called the Kulikovo
 field'
 (Alekseev & Karcov 1964, as cited in Schönig 1992/93:328)

Examples of this type are claimed to arise because the "[s]urface distinction" between the two types of relativization patterns in Turkish is "neutralize[d]" (Slobin 1986:286).

Note, however, that the Turkmen relative clause is every bit as "complex" as its Modern Turkish counterpart. The embedded subject is in the genitive case, thus violating principle 2, i.e. it is not in the "canonical" form for subjects. Furthermore, while the participle is indeed bare, i.e. devoid of agreement morphology, the construction itself does have agreement—namely on the head of the relative clause. In this context, it is important to realize that the head noun is not interpreted as possessed by the subject; in other words, the embedded subject is not a possessor. The agreement on the head noun (or head noun phrase) of the relative clause has the very same function as the agreement on the participle in its Turkish counterpart, namely to serve as subject agreement (rather than possessive agreement).

It appears, then, that there is no reason for claiming that in Turkmen, cognitively and perceptually complex and difficult Turkish patterns were neutralized in favor of "simpler" constructions. Rather, a certain Old Turkic pattern of relative clauses which would count as "simple", due to its resemblance to canonical main clauses and its lack of synthetic morphemes, underwent similar, but not identical, developments in Turkish and Turkmen towards what appear to be more "complex" constructions.

I draw the conclusion at this point that the historical development of the grammar of relative clauses in both Turkish and Turkmen cannot be explained by the proposals in Slobin (1986), and that these diachronic facts that were taken by him to illustrate and corroborate his principles actually are problematic for those principles. In concluding this discussion, I now turn to some hypotheses about the reasons for the developments we saw above that took place in the grammar of relative clauses in some Turkic languages.

5. Concluding hypotheses. In Turkish (as well as a number of other Turkic languages), the temporal/aspectual distinctions among certain participles were neutralized, and some of these nominalization morphemes were used instead to mark differences in terms of (non-)subjecthood of the relativization target, as mentioned. In a number of other Turkic languages, the neutralization did not only

concern temporal/aspectual distinctions, but the number of the morphemes as well, which was reduced to just one, as appears to be the case with Turkmen.

What is common in both types is an impoverishment in temporal/aspectual functions, and thus the common innovation is a clearer-cut array of differences between subordinate and main clauses (contra Slobin's principle 2) as compared to previous stages of all the Turkic languages. In this sense, Turkic languages do become more similar to Indo-European languages, i.e. the subordinate clause is clearly marked by means of overt markers and is thus set apart from the main clause, and subordinate clauses have different syntactic properties from those of main clauses. However, this is achieved not by means used in Indo-European (i.e. by free subordination morphemes), but by "Turkic" means, i.e. by nominalization (and case) suffixes (contra principle 1).

It appears, then, that neither principle 1 nor principle 2 is explanatory in its current form and needs to be refined.

The fact that the diachronic developments in complementation and subordination structures in Turkic, driven (so I claim) by the necessity to introduce clear distinctions between root and subordinate clauses, are very natural and led to stable and morphologically transparent syntactic states is shown by the fact that, despite very strong and politically enforced Arabic and Persian influences (e.g. right-branching structures, non-affixal complementizers, finite subordinate clauses) during the Ottoman centuries, Turkish still exhibits primarily Turkic properties in its syntax of subordination. The distinctions between "subject" and "non-subject" relative clauses are robust. Consequently, whatever changes towards Indo-European forms of subordination are found in the Turkic languages spoken in the formerly Soviet areas must, to a large part, be due to political (and perhaps socio-economic) pressures and cannot (exhaustively or even primarily) be explained by "deep-seated tendencies in Turkic, going back to the earliest records" (Slobin 1986:282). If so, the acquisition facts in Turkish (as claimed to be contrasting with those in the Indo-European languages) remain unexplained and must be investigated anew.⁵

Abbreviations

ACC	Accusative
AN	Subjunctive gerund (=action nominal)
AOR	Aorist
DAT	Dative
FG	Factive gerund
FUT	Future
GEN	Genitive
NEG	Negation
NOM	Nominative
OP	Object participle (=FG: Factive gerund)
P	Participle
PERFP	Perfect participle
PL	Plural
PST	Past
SG	Singular
SP	Subject participle

Notes

*I would like to thank the participants of the Berkeley Linguistic Society for their questions and comments. I am particularly grateful to Dan Slobin for his comments after the presentation of this paper, and to Claus Schöning and to Karl Zimmer for enlightening discussion. I also thank the various funding sources within Syracuse University for their travel support which made the presentation of this paper possible.

¹This does not mean that I attribute this point of view to the current thinking of Dan Slobin. As a matter of fact, he has informed me that he now thinks differently about the relationship between first language acquisition and historical change. In this paper, my main objective was not to criticize Slobin; however, since his work is exceptional in having clearly articulated concrete and specific proposals concerning the relationship between acquisition and diachrony, I addressed some of the claims he made due to their general interest.

²This is only a rough generalization. There are some additional factors that also contribute to the choice of the particular pattern of relative clauses. For more detailed discussion and attempts at explaining the criteria governing the choice of the different patterns of relative clauses, the reader may consult a number of works that have addressed this issue; e.g. Barker, Hankamer & Moore (1990), Dede (1978), Hankamer & Knecht (1976), Kornfilt (1997), Zimmer (1987 and 1996). For our present purposes, the rough characterization of the two patterns given in the text will suffice.

³I have chosen the designation "OP" whenever the target of relativization is a non-subject. In Modern Turkish, whenever the subject relativization pattern is used in such an instance, the participle must bear a passive marker. None of the following examples exhibits the passive marker, thus justifying the designation "OP" in the instances it was used and where the translation appears to suggest application of passive. Note that the translations (which, in most instances, I took from the secondary sources noted with the examples) are only idiomatic and are not designed to convey any particular syntactic analysis. Note also that Old Turkic did have a passive morpheme.

⁴In the following list of examples, the capital letters after the examples designate a particular monument. KT: Kül Tigin; TII: the second monument erected for Tonyukuk. BK: Bilge Kagan. The designations for the directions (i.e. E: East, N: North, S: South) refer to the various faces of the monuments. Finally, the numbers after these letters refer to the lines.

⁵In this context, it is interesting to note that some recent acquisition studies report results that are surprising from the point of view of the principles proposed in Slobin (1986), while being rather expected in the light of the diachronic developments in Turkic subordination. Somashekar, Foley and Gair (1998) report that in the course of repetition experiments concerning Tulu, a south Indian Dravidian language, children were given as stimuli right-branching, fully finite correlative construction—the relative clause type closest to "canonical" main clauses. However, young children would instead utter left-branching, nominalized relative clauses—structures very similar to the Turkic relative clauses discussed in this paper, i.e. structures claimed to be cognitively much more difficult. It is mysterious at this point why Tulu children should be able to acquire such constructions earlier than their Turkish counterparts, and an evaluation of these

acquisition facts must await further study. Nevertheless, these facts are challenging in the light of the discussion presented in this paper.

Bibliography

- Barker, Chris., Jorge Hankamer and John Moore. 1990. *Wa and Ga* in Turkish. UCSC manuscript.
- Dede, M. 1978 "Why should Turkish Relativization Distinguish between Subject and Non-Subject Head Nouns?" BLS 4: 67-77.
- von Gabain, Annemarie. 1941 *Altürkische Grammatik*. (Third edition, 1974.) Wiesbaden: Otto Harrassowitz.
- Hankamer, J. & L. Knecht. 1976. "The Role of the Subject/Non-Subject Distinction in Determining the Choice of Relative Clause Participle in Turkish". NELS VI (Montreal Working Papers in Linguistics 6): 123-135.
- Jeffers, Robert J. and Ilse Lehist. 1979, *Principles and Methods for Historical Linguistics*. (Paperback edition, 1982.) Cambridge, MA: MIT Press.
- Kiparsky, Paul. 1968. Linguistic Universals and Linguistic Change. Universals in Linguistic Theory, ed. by Emmon Bach and Robert Harms, 171-202. New York: Rinehart and Winston.
- Kiparsky, Paul. 1978. Historical Linguistics. A survey of Linguistic Science, ed. by William Orr Dingwall, 33-57. Stamford: Greylock publishers.
- Kornfilt, J. 1997. Synchrony and Diachrony of Relative Clauses and Possessive Phrases in Turkish. Manuscript, Syracuse University.
- Lightfoot, David. 1979. *Principles of Diachronic Syntax*. Cambridge: Cambridge University Press.
- Schönig, Claus. 1992/93. Relativsatzbautypen in den sogenannten altaischen Sprachen". *Acta Orientalia Academiae Scientiarum Hung.* Tomus XLVI (2-3), 327-338. Budapest: Akadémiai Kiadó.
- Slobin, Dan Isaac. 1986. The Acquisition and Use of Relative Clauses in Turkic and Indo-European Languages. *Studies in Turkish Linguistics*, ed. by Slobin, Dan Isaac and Karl Zimmer, 273-291. Amsterdam: John Benjamins.
- Somashekar, Shamitha, Claire Foley and James W. Gair. 1998. Tulu Relative Clause Acquisition in a Cross-linguistic Perspective. Paper presented at the annual meeting of the Linguistic Society of America, New York City.
- Tekin, Talat. 1967. *A Grammar of Orkhon Turkic*; Uralic and Altaic Series, Vol. 69. Bloomington: Indiana University Publications.
- Underhill, R. 1972. "Turkish Particples". *Linguistic Inquiry* 3.1: 87-99.
- Zimmer, Karl. 1987. Turkish Relativization Revisited. *Studies on Modern Turkish*, ed. by Hendrik Boeschoten and Ludo Verhoeven, 57-61. Tilburg: Tilburg University Press.
- Zimmer, Karl. 1996. Overlapping Strategies in Turkish Relativization. *Current Issues in Turkish Linguistics*, ed. by Bengisu Rona, 159-164. Ankara: Hitit Yayınevi (Copyright London: SOAS).

The Case of the Noun Phrase in the Finnish Adpositional Phrase: A Discourse-pragmatic Account

Ritva Laury
California State University, Fresno

It is commonly assumed that adpositions are heads of adpositional phrases and therefore syntactically govern the form of the noun phrase within the adpositional phrase (Nichols 1986; Zwicky 1993; Comrie 1989; and many others). In this paper, I will argue that such an assumption is challenged by examination of adpositional phrases in naturally occurring language. I will present data from spoken Finnish which show that sometimes it is not plausible to claim that the case of the noun phrase is determined by an adposition with which it can nevertheless be said to form, or be intended to form, an adpositional phrase. I will argue that the case marking of noun phrases within adpositional phrases is motivated by discourse-pragmatic and semantic factors, while purely syntactic factors are inadequate to account for the form of the noun phrases in Finnish adpositional phrases.

1. DATA. My data consist of twenty spoken narratives, some elicited, some spontaneously told, and eight ordinary conversations between friends, co-workers and family members (for further details, see Laury 1997). In addition, I have taken several examples from the Finnish Pear Stories (Chafe 1980), recorded in Finland in 1984.

2. ABOUT FINNISH ADPOSITIONS. In addition to its rich system of local cases, Finnish also has both prepositions and postpositions. Postpositions dominate in the data; 89% (84 out of the 94 total) of the adpositions in the data are postpositions (cf. Palola 1975 with 82.5% postpositions). Most postpositions co-occur with genitive NPs (77 out of 84), while prepositions appear mostly with partitive noun phrases (9 out of 10). Some adpositions occur with either genitive or partitive NPs, and some occur with NPs in local cases (Suova 1938; Airila 1938; Hakulinen and Karlsson 1979; Vilkuna 1996).

3. NOUN PHRASES WHOSE CASE CANNOT BE GOVERNED BY AN ADPOSITION. In this section, I will present and discuss data which show that the case of an NP complement of an adposition cannot always be claimed to be governed by an adposition. I will first discuss co-constructed adpositional phrases; then I will discuss what I call independent landmarks, noun phrases which appear to be intended as, or have potential to become complements of adpositions but are not accompanied by adpositions; and finally I will discuss adpositions which appear with noun phrases in more than one case.

It is very common in ordinary conversational language that two or more speakers are responsible for a particular grammatical construction; this happens, for example, when one speaker starts a sentence or phrase and another speaker finishes it. It is also possible for one speaker to 'abduct' part of a construction produced by one speaker and use it for a new construction different from the construction it was used for by the first speaker. In example (1) below, a group of friends gathered together for coffee are discussing where the hostess, speaker EK, has purchased some of her dishes. This is achieved through a considerable amount of overlap and other types of cooperative construction.

(1)

1 LP Onk 'nämät niitä ku sä olet sielt,
be-Q this-PL 3PL-PRT COMP 2SG be-2SG DEM.LOC-ABL

Are these the ones that you have,

2 EK ^Roomasta [tuonu].¹

Rome-ELA bring-P.PPLE

brought from Rome.

3 LP [Nii] sie[ltä,
PTCL DEM.LOC-ABL

From the,

4 EK [Vatikaanin] torilta.

Vatican-GEN market.place-ABL

Vatican market place.

5 LP sieltä] torilta.

DEM.LOC-ABL market.place-ABL

the market place.

6 EK [Sielt] nii,

DEM.LOC-ABL PTCL

From the,

7 LP [Nii.]

PTCL

Right.

8 EK Vatikaanin kesä--

Vatican-GEN summer

Vatican summer—

- 9 Siis se,
PTCL DET/DEM
I mean the,
- 10 .. eiku=,
NEG-CLTC
no,
- 11 ... san nyt,
say-2SGIMP now
what is it,
- 12 'paavin ^kesäasunnon.
pope-GEN summer-residence-GEN
of the pope's summer residence.
- 13 LP Niin 'paavin ^kesäasunnon.=
PTCL pope-GEN summer-residence-GEN
Yeah, (of) the pope's summer residence.
- 14 EK =Vieres siel oli se,
next.to-INE DEM.LOC-ADE be-PST DET/DEM
Next to [the pope's summer residence] there was the,

In line 2, in response to the question begun by LP in line 1, EK completes LP's clause by saying that the dishes were brought from Rome. In line 4 she specifies further that they come from the Vatican marketplace. In 6-11 she initiates further specification, and the NP *paavin kesäasunnon* 'the pope's summer residence-GEN' in line 12 appears strongly to be intended as a repair of the NP *Vatikaanin* in line 4, which functions as a genitive modifier of *torilta* 'from the market place'. After LP repeats the NP in line 13, my interpretation is that EK abducts this NP and uses it as a complement of the postposition *vieres* which LP produces in line 14. Since the genitive case of this noun phrase is motivated by the fact that it is functioning as (a repair of) a genitive modifier, it does not seem possible to claim that the postposition *vieres* in any way governs or determines the case of the NP with which it can nevertheless be said to form an adpositional phrase.

Furthermore, there are also occasions when speakers produce a genitive NP without any accompanying adposition or noun phrase. In these cases, the independent genitive ordinarily functions as a landmark (O'Dowd 1994), while the relational element, the pathway, which the adposition would have provided, is missing. It is not always easy to know how to interpret the noun phrase in terms of

providing the 'missing' element. It is also not clear whether listeners experience any difficulty in understanding what was meant. Consider the following example, where several speakers are discussing a particular location in the archipelago outside the town of Turku.

(2)

1 IW Ja sit vast [se ^Naan]talin aukko on **sen** [toisen],
and then only DET N.-GEN gap be-PRES DET-GEN other-GEN
And then the Naantali gap is (*) the other,

2 LP [Nii.]
PTCL
Right.

3 JS [<X Nii eikä seX>]
PTCL NEG-CLTC DET/DEM
Right, and it doesn't

4 LP Nii.
PTCL
Right.

5 IW .. **saaren**,
island-GEN
island's,

6 siin on semmonen [^toinen] saari sit [siin],
DEM-INE be-PRES such other island then DEM-INE
then there's another island there,

7 LP [Joo]
PTCL
Yeah.

8 [Joo.]
PTCL
Yeah.

9 IW .. koht **siin** ^Särkän saaren,
soon DEM-INE S.-gen island-GEN
soon (*) Särkkä island,

10 ... **ettippäi siitä** [sit].

ahead DEM-ELA then
ahead from there.

11 LP [Ni=i.]

PTCL

Right.

In this excerpt, speaker IW produces two genitive noun phrases without a following noun or adposition, one in lines 1 and 5 and the other in line 9, but although it could be argued that something was omitted, it is difficult to say what. Either a postposition or a local case noun phrase would be syntactically possible here, although it is easier to imagine postpositions filling in the slot. However, the other speakers seem to be able to comprehend her message and do not ask for clarification; the backchannel responses *joo* and *nii* can be taken as an indication that the addressee(s) are comprehending the message (Sorjonen 1997).

It is my view that what the speaker is doing here is providing a landmark; that is, she is indicating that the Naantali gap can be located by locating *se toinen saari* 'the other island' and *Särkän saari* 'Särkkä island', which she presents as identifiable to her addressees by using the determiner *se* (Laury 1997). However, when she produces these noun phrases, she does not provide a pathway; in other words, she does not indicate, by supplying an adposition, where the Naantali gap is located with respect to these landmarks. In my view, in cases like example (2), it would not be reasonable to argue that the case of the NP was governed by an adposition which was never produced. Instead, the genitive case here functions to simply mark the NP as a landmark. We could, of course, assume that the speaker did have a particular adposition in mind but failed to produce it for some reason. However, I see no reason to assume so; one piece of evidence that she had not planned the adposition yet when she produced the noun phrases in lines 1 and 5 and in line 9 is the adpositional phrase in line 10.² The adposition she finally does produce, *ettippäi* 'ahead (from)' is a preposition, not a postposition, and occurs with noun phrases in the elative case, not with genitive noun phrases. Thus, I take data like example (2) as evidence that at the time a speaker produces a noun phrase intended as a landmark, which has the potential to be followed by a postposition, the speaker may not yet have planned the identity of a postposition which may (or may not) follow. This in turn indicates that the case of a noun phrase within a postpositional phrase may have been determined by other factors; for example, as in this case, the function of the NPs as identifiable landmarks.

Also look at example (3). This is an excerpt from the Finnish Pear stories. The speaker has just heard a Pear story and is instructing the teller as to where to turn in a questionnaire he has been given.

(3)

1 ... jos 'täytät ^tän,
 if fill-2SG this-ACC
 if you'll fill this out,

2 ja 'annat **tuolla** ^**odottavan** '**henkilön**,
 and give-2SG DEM.LOC-ADE wait-PRES.PPLE-GEN person-GEN
 and give [it to?]/let the person waiting,

3 ... tota,
 PTCL
 um,

4 ... 'tytölle.
 girl-ALL
 to the girl.

In this example, *tuolla odottavan henkilön* 'the person waiting there' is in the genitive case. At the point it is produced, the speaker has several options. Here, he could have followed this with a postposition, a noun phrase, or an infinitive construction. I think this again shows that all genitive noun phrases which occur with postpositions are not necessarily genitive because of the valency of the following postposition. Both example (2) and (3) show that at the time the speaker produces a genitive NP, it is possible that he or she may not have yet planned the construction the NP will be a part of. In example (2), the speaker provided two genitive NPs which appeared to be sufficient landmarks for the addressees even without a path-specifying adposition; when she did produce an adposition, it was one which appears with noun phrases in a case different than the case of the earlier landmark NPS. And in (3), the speaker produced a genitive NP but self-corrected and produced an oblique NP instead.

We have seen two types of constructions which make it difficult to maintain that the case of an NP complement of an adposition is always governed by an adposition, constructions where the adpositional phrase is cooperatively constructed and constructions where the adposition is never produced. Thirdly, although some Finnish adpositions only occur with genitive NPs and others only with partitive NPs, some adpositions, such as *ympäri* 'around' and *kesken* 'in the middle of' take both genitive and partitive complements. The case marking of the complements of this third type of adposition also cannot be governed by the adposition, but must be determined by some other factors. In the next section, I will discuss semantic and discourse factors which motivate case marking of NPs within Finnish adpositional phrases.

4. DISCOURSE-PRAGMATIC AND SEMANTIC MOTIVATIONS FOR CASE MARKING OF NPS WITHIN ADPOSITIONAL PHRASES. In this section, I am going to propose that the case marking of noun phrases in adpositional phrases is motivated by the role played by the referent of the NP in the current discourse and by semantic factors. I am first going to review research concerning semantic and pragmatic motivations for case marking in Finnish, and then I will show how the case marking of NPs in my database can be argued to be motivated by semantics and pragmatics.

Quite a bit of research on the discourse functions of the Finnish cases has shown that referents of partitive and genitive noun phrases play quite different roles in discourse. The partitive case was originally a local case, and even in their grammatical uses, partitive NPs in discourse still manifest features which are more characteristic of obliques than grammatical cases. Partitive NPs have been found to generally occur in contexts of irresultativity, irrealis, negation, partial affectedness, and low transitivity; they are characterized by a low degree of individuation and nonreferentiality, and their referents are likely to be non-human (Helasvuo 1996).

In contrast, one of the prototypical functions of the genitive case is the expression of ownership and part-whole relationships, and more than a third of all adpositions which occur with genitive noun phrases (and more than half of such constructions which express location) are thought to originate from constructions where the genitive noun phrase expressed the whole and the emerging postposition expressed the part (Jaakola 1997). In addition, Huumo and Inaba (1997) argue that the historical origin of the genitive case in Finnish is not adnominal, but rather a lative case which later developed into a dative. In this function, Huumo and Inaba suggest that its function was one of possession, and that the referents of the genitive noun phrases in these constructions were exclusively animate and predominantly human.

If we compare the partitive NPs in the adpositional phrases in my data to the genitive NPs, quite a few differences emerge. The referents of the partitive NPs are more likely to be new, less likely to be identifiable, less likely to be subsequently mentioned, and much less likely to have human referents, than the genitive NPs within adpositional phrases, which in turn are much more likely than partitive NPs to have been already mentioned, to be mentioned again, to be identifiable, and to have human referents. This can be seen in the table below.

Case	Total NPs	New	Identifiable	Subs. mentioned	Human
PRT	16	9 60%	6 40%	3 13%	1 7%
GEN	78	25 32%	56 72%	34 44%	24 31%

Table 1. Discourse profiles of partitive and genitive noun phrases in adpositional phrases

In other words, it can (and should!) also be argued that the partitive and genitive NPs perform different functions in discourse, and have become grammaticized (or lexicalized) into use in adpositional phrases which reflect these discourse functions. For example, as we have seen above, genitive NPs are often used as landmarks; they also manifest characteristics suitable for such a function, such as identifiability. Recall that the landmark NPs in example (2) were explicitly marked for identifiability; the genitive NP in example (1) was also used as a landmark, and constituted the second mention of that NP in the discourse. On the other hand, as is typical of nonreferentials and other types of nominals with a low degree of individuation, nominals in the scope of negation and other types of irrealis, partitive NPs in my data are likely to be new, not likely to be subsequently mentioned, and much less likely to be identifiable to the addressee than are genitive NPs.

Further, if we examine pairs of adpositions which appear with genitives vs. partitives, it turns out the semantic differences between them are reflective of the kinds of differences which showed up in the comparison of partitive and genitive noun phrases above. Thus the adposition *kanssa* 'with', which implies existence (realis), occurs with genitive noun phrases. Eleven out of the 21 uses of *kanssa* had human noun phrases as objects, and two others which were animate. In contrast, *ilman* 'without', which implies nonexistence (irrealis), occurs with partitive noun phrases; there were only two uses of *ilman* in my database, but both had nonhuman referents. *Jälkeen* 'after' implies existence and occurs with genitive noun phrases, while *ennen* 'before' which implies nonexistence, occurs with partitive noun phrases. Adpositions which imply contiguity and containment (a high degree of affectedness), such as *halki* 'across', *läpi* 'through', *sisällä* 'inside', and *kohdalla* 'at' occur with genitives, while adpositions which imply noncontiguity, noncontainment or opposition (and thus a lower degree of affectedness), such as *kohti* 'towards' *pitkin* 'along' and *vasten* 'against', occur with partitive noun phrases.

With those adpositions which occur with both genitive and partitive noun phrases, the variation between the cases manifests predictable semantic differences. Thus *ympäri* 'around' occurs with both genitive and partitive noun phrases. With a genitive noun phrase, as in *kaupungin ympäri* 'around the town', a complete circle, or containment (total affectedness) is implied; however, in *ympäri kaupunkia* 'around town' with a partitive noun phrase, the implication is a scattered location (partial affectedness) in various parts of the town (Jaakola 1997).

Similarly, in my data, two different forms of the adposition *päässä/stä* 'at/from the end (of)' are used by the same speaker twice within one narrative, once with a genitive noun phrase and another time with a partitive noun phrase. The speaker is foxhunting on a mountain, and is disturbed by woodcutters:

(4)

..Sit tul pari ^hevosmiestä sit,
Then come.PST couple horse-man-PL-PRT then
Then a couple of men came with horses,

...'rantaan vielä ja,
shore-ILL still and
to the shore and,

...tulivat mettäh,
come-PST-3PL forest-ILL
and they came into the forest,

..halkoi hakemah sielt.
firewood-PL.PRT fetch-3INF-ILL DEM.LOC-ABL
to get firewood from there.

..**Vuoren päästä.**
mountain-GEN end-ELA
From the end of the mountain.

In this example, the location referred to by the adpositional phrase is viewed as contiguous; it is near to the speaker and accessible to him in that he can hear the woodcutters. The hunter is disturbed about the noise the woodcutters are making, and starts walking away. It then occurs to him that he should go to the other side of the mountain where he knows of a foxgallow.

(5)

ja muistin,
and remember-PST-1SG
and I remembered (that)

heill on siel **toises pääs vuorta**,
3SG-ADE is DEM.LOC.ABL other-INE end-INE mountain-PRT
they have at the other end of the mountain

sellai kivi,
such rock
this rock,

This time, the speaker uses a partitive noun phrase *vuorta* with the adposition *päässä* (which is modified by the adjective *toises*). The other end of the mountain

is not near the speaker or accessible to him in the way that the end of the mountain referred to in example (4) is; the speaker is currently at the end mentioned there, but not at the end mentioned in example (5) above. It is interesting that while only the genitive is possible with *päässä*, both genitive and partitive are possible with *toises(sa) pääs(sä)*. Thus the choice of cases used with adpositions is sensitive to both semantic features (the other side, vs. just the side) and pragmatic features (how a particular referent is being viewed, in terms of its current accessibility to the speaker).

CONCLUSION. I have presented evidence which indicates that the form of an NP within an adpositional phrase is not just syntactically governed by the adposition, but is instead strongly motivated by discourse-pragmatic and semantic factors. I did this by first presenting data which showed that in naturally occurring language it is not always possible to claim that the form of an NP is governed by an adposition with which it can be said to form an adpositional phrase, since speakers abduct NPs which were originally functioning in some other type of grammatical function. They also sometimes produce NPs which appear to have been planned as complements of adpositions before the identity of the adposition is planned. Further, some adpositions take complements in more than one case.

I then discussed earlier studies which have investigated the discourse functions of the partitive and genitive case, and showed that the discourse profiles of NPs within adpositional phrases in my data are consistent with the findings of the earlier studies. I also showed that the semantics of the adpositions which occur with partitive vs. genitive NPs are consistent with the kinds of meanings partitives and genitives tend to express. I conclude that partitive and genitive noun phrases perform distinctly different functions in discourse, and to the degree that they have become grammaticized into use with particular adpositions, the case assignment is a result of, and still transparently reflects, these discourse functions.

NOTES

1. ^ indicates the word with the primary stress in that intonation unit.
2. It is not entirely clear whether *ettippäi siitä* is an adpositional phrase; it could also be considered an adverbial phrase. The distinction between adverbs and adpositions in Finnish is a continuum, not a strict dividing line (Airila 1938; Hakulinen and Karlsson 1979:154; Vilku 1996:46).

BIBLIOGRAPHY

- Airila, Martti. 1938. Suomen pre- ja postpositioiden rektiosta. (About the valency of Finnish pre- and postpositions). *Virittäjä* 42:224-226
- Chafe, Wallace (ed.) 1980. *The Pear Stories*. Norwood: Ablex.
- Comrie, Bernard. 1989 (1981). *Language Universals and Linguistic Typology*. Chicago: University of Chicago Press.
- Hakulinen, Auli and Fred Karlsson. 1979. *Nykysuomen lauseoppi*. (Modern Finnish syntax). Helsinki: Suomalaisen Kirjallisuuden Seura.

- Helasvuo, Marja-Liisa. 1996. A Discourse Perspective on the Grammaticization of the Partitive Case in Finnish. SKY 1996, ed. by Timo Haukioja, Marja-Liisa Helasvuo and Elise Kärkkäinen Helsinki: Suomen kielitieteellinen yhdistys.
- Huomo, Tuomas and Nobufumi Inaba. 1997. Irrallinen genetiivi ja omistusrakenteen ongelma. *Virittäjä* 101:27-48
- Jaakola, Minna. 1997. Genetiivin kanssa esiintyvien adpositioiden kieliopillistumisesta. Kieliopillistuminen: Tapaustutkimuksia suomesta (Grammaticization: Case studies from Finnish), ed. by Tapani Lehtinen and Lea Laitinen. Helsinki: Department of Finnish, University of Helsinki.
- Laury, Ritva. 1997. Demonstratives in Interaction: The mergence of a definite article in Finnish. Amsterdam: Benjamins.
- Leino, Pentti. 1991. Lauseet ja tilanteet. (Sentences and situations). Helsinki: Suomalaisen Kirjallisuuden Seura.
- Nichols, Johanna. 1986. Head-marking and Dependent-marking Grammar. *Language* 62:56-119.
- O'Dowd, Elizabeth. 1994. Prepositions and Particles in English: A Discourse-Based, Unifying Account. Ph.D. dissertation, Department of Linguistics, University of Colorado at Boulder. In press as Prepositions and Particles in English: A Discourse-Based Account. Oxford: Oxford University Press.
- Palola, Helena. 1975. Adverbit, postpositiot, prepositiot ja interjektiot 1960- luvun suomen yleiskielessä. Oulun yliopiston suomen ja saamen kielen laitoksen tutkimusraportteja 2. Oulu: University of Oulu, Dept. of Finnish and Sami.
- Sorjonen, Marja-Leena. 1997. Recipient Activities: Particles *nii(n)* and *joo* as Responses in Finnish Conversations. Ph.D. dissertation. Department of Applied Linguistics, UCLA.
- Suova, Eino. 1938. Post- ja prepositioiden pääsanan sijoista (About the cases of the head word of post- and prepositions). *Virittäjä* 42:114-118.
- Vilkuna, Maria. 1996. Suomen lauseopin perusteet. (Foundations of Finnish syntax.) Helsinki: Edita.
- Zwicky, Arnold M. 1993. Heads, bases and functors. Heads in Grammatical Theory, ed. by Greville G. Corbett, Normal M. Fraser and Scott McGlashan. Cambridge: Cambridge University Press.

*Anar 'Go', Donar 'Give', and Posar 'put': Shift Verbs in Catalan*¹

Roser Morante
Universitat de Barcelona

and Glòria Vázquez
Universitat de Lleida

1. Introduction

The main aim of this article is to prove the consistency of the semantic class of *shift verbs*. It has been developed as part of a wider project of semantic classification of Catalan verbs. The project is framed in the lexical semantic theoretical approach and it takes as a starting point the work by Beth Levin *English Verb Classes and Alternations* (1993). Levin's work proposes a large-scale classification of English verbs. Except for a few examples (Martí *et al.* 1997 and Vázquez and Martí 1998), Catalan verbs have not been examined from this perspective.

The classification of English verbs carried out by Levin is founded on the hypothesis that syntax and semantics are closely related, so that verbal classes not only share a basic meaning, but also show the same syntactic behavior. However, her initial assumptions are not always maintained: some classes are delimited following a semantic field criterion, only and do not always present an accurate syntactic-semantic uniformity (Castellón *et al.* 1998). As a matter of fact, Levin does not prove that verbs belonging to the same verbal class share the same meaning components, since no meaning components are defined for each verbal class. It is also not indicated which syntactic alternations are accepted by all the members of the larger classes.

We attempt to resolve these inconsistencies in the following way. Firstly, we start with a higher level of semantic abstraction. We establish a set of meaning components, capable of defining the classes. Secondly, we do not consider that verbal classes, even sharing the same meaning components, must necessarily show a fine-grained syntactic homogeneity. Obviously, the results we obtain under these assumptions differ from Levin's, although her hypothesis will be maintained: the analysis of the shift verb class will prove that verbs sharing the same basic meaning components show a common syntactic behavior.

In the following section the shift class will be outlined and its meaning components will be defined. Later, meaning components (section 3), realization of meaning components (section 4) and syntactic behavior (section 5) will be analyzed in detail.

2. The shift verb class

The shift class is defined as the group of verbs whose basic meaning expresses an event in which an entity shifts from a source to a goal through a path. In general terms, the verbs of this class, whose representative members would be *donar* 'to give', *posar* 'to put', and *anar* 'to go', denote change of possession, change of location, and movement. In the following schemes A refers to the *originator*, B to the *entity*, and C to the *path*:

- (1)
 - a. someone (A) gives something (B) to someone (C)
 - b. someone (A) puts something (B) somewhere (C)
 - c. someone (A/B) goes from one place to another (C)

The concept of shift adopted here is similar to Ikegami's *motion* (1973). This author considers that movement can occur with an abstract entity. However, *shift* differs from Ikegami's *motion* in two respects: firstly, Ikegami's *motion* includes verbs of change of state among the verbs of motion, while *shift* does not. Secondly, Ikegami's *motion* does not include either verbs expressing movement of a part of an entity, or verbs expressing movement of an entity when the entity remains in the same place (*estirar* 'to stretch' or *sacsejar* 'to shake').

In addition to Ikegami's, we have considered Talmy's conception of motion. We differ from Talmy in the semantic elements he defines for motion predicates, which are: the abstract predicate of motion, the moving entity (figure), the reference point (ground), and the path of the motion with respect to the ground. For us the abstract predicate of motion is a cognitive image. Motion is conceived at an abstract level as the product of a combination of semantic components, so that it is not conceived as simply one element of the event.

Motion verbs have also been analyzed by Jackendoff (1976, 1983, 1990). He establishes two groups of verbs as a result of the application of the functions MOVE and GO. We unify both groups of verbs, although, at the same time, we exclude some members of each group. Regarding the function MOVE, we put apart verbs like *to laugh* and *to snooze*. From our point of view, for a predicate to belong to the shift class, it is necessary that it expresses that an entity moves, not as a secondary effect, but as the main event activity. Concerning the function GO, Jackendoff uses it to define prototypical verbs of motion (*anar* 'to go'), of change of possession (*donar* 'to give'), and of change of state (*barrejar* 'to mix'). In our classification, verbs of change of state do not belong to the shift class.

A list of Levin's classes that have been included in the shift verbs is to be found in the appendix. We have taken these classes as a reference, but not all the verbs belonging to them have been classified under the shift class. Most of the verbs excluded belong to the change of state class. An example are groups (9.7), (9.8), (9.9), (10.6), and (10.7), which contain verbs of the type *emblanquinar* 'to paint white' or *omplir* 'to fill':

- (2)
 - a. El pintor emblanquina la paret (the painter paints white the wall)
 - b. El ciclista omple una ampolla d'aigua (the cyclist fills a bottle with water)
 - c. El ciclista posa aigua a l'ampolla (the cyclist puts water into the bottle)

Levin includes these verbs (2a, 2b) in the same classes as the verb *posar* 'to put' (2c). She applies criteria which differ from Jackendoff's or Ikegami's. Levin claims that these verbs express the physical movement of an entity A to a place B, (in the case of *paint white*, the white paint A moves to the wall B). As we see it, these verbs do not focus on the change of location of entity A, but on the change of

state of the final location of entity B (the wall becomes white)². Thus, the meaning of change of state verbs always focuses on the final state of the entity (Fernández *et al.* 1998), whereas in shift verbs it focuses exclusively on the change of location. Our proposal is that (2c) is an example of a shift verb, but (2a) and (2b) are not.

3. Meaning components of shift predicates

Meaning components are the semantic units on the basis of which a verbal class is defined. Two kinds of components are differentiated: basic and secondary. The combination of basic components defines a semantic class. Secondary components may complete the meaning of a semantic class, but do not define it. Each of the components has its own way of being expressed (section 4).

3.1. Basic meaning components

Verbs belonging to the shift class share three meaning components: *entity*, *path*, and *originator*. They are related in the following way: the entity moves along a path, its movement being originated either by another entity or by the shifting entity itself (in this case originator and shifting entity do coincide).

- *Entity*: the element that shifts. In change of possession verbs it corresponds to the possessed object (3a), in motion verbs to the entity which moves (3b, 3d), and in verbs of change of position, to the object which changes its location (3c):

- (3)
- a. El Robert va vendre *un compact* al seu amic
(art. Robert aux. sell_inf a CD to his friend)
 - b. *El tren* va de Madrid a Barcelona via Saragossa
(The train goes from Madrid to Barcelona through Saragossa)
 - c. La grua va portar *el cotxe* fins al taller
(The tow_truck aux. bring_inf the car to the garage)
 - d. *L'aigua* va arrossegar els cotxes dos metres
(The water aux. drag_inf the cars two meters)

- *Originator*: the entity that initiates the movement. Although it is usually represented by a human entity, there are also cases in which the originator is an object (*la grua* in 3c) or a natural cause (*l'aigua* in 3d).

- *Path*: the distance, abstract or real, traveled by the shifting entity. In change of possession verbs the distance covered is from one possessor to another (3a), in verbs of change of position, from one position to another (3c), and in motion verbs, from the point where the movement begins to the point where it ends (3b, 3d).

The path component is divided into three subcomponents: *source* (path_source) -the initial point of the distance (*de Madrid* in 3b), *goal* (path_goal) - the final point (*a Barcelona* in 3b), and *intermediate* (path_intermediate) - any point or distance between the source and the goal (*via Saragossa* in 3b). The realization of one of the subcomponents suffices for the realization of the path component as a whole.

3.2. Secondary components

Time, place, manner, and instrument are the secondary components. Since they do not define the shift verb class, they will not necessarily be expressed, either lexically or syntactically. One or the other will appear according to the type of predicate. They can be found in most other verb classes as time, place, and manner are concepts that underlie any verbal expression: all action happens in a concrete moment, place, and manner, which can be more or less specified in the predicate (Castellón *et al.* 1998).

• *Time*: either the moment when the entity travels over the path (4a) or the duration of the movement (4b). It is especially related to some motion verbs, perhaps for cultural reasons. Verbs like *sortir* ‘go out’, *arribar* ‘arrive’ have a closer relation to time than other verbs:³

- (4) a. La Maria arribà a les 9 en punt (art. Maria arrived at art. 9 o'clock on point)
 b. La Maria caminà durant dues hores (art. Maria walked for two hours)

• *Place*: location where the movement happens. In the case of shift verbs it may appear conflated in the verb (5a) expressing a physical environment (air, sea) or it may be incorporated in the form of a syntactic constituent (5b):

- (5) a. El president va volar al Japó (The president aux. fly_aux. to Japan)
 b. El petroli es transporta per mar (art. oil pron. transports by sea)

• *Manner*: it can refer to two concepts: (i) Manner in which the displacement of the entity is carried out. This component is conflated in certain verbs of change of position (6a), motion (6b), and change of possession (6c):

- (6) a. L'atleta va llançar la javelina dos metres més enllà
 (The athlete aux. throw_inf the javelin two meters more far)
 b. La gent es va precipitar cap a la sortida
 (The people pron. aux. rush_inf to to the exit)
 c. L'estat va confiscar les propietats al polític
 (The state aux. confiscate_inf the properties from_the politician)

(ii) Form which the shifting entity acquires after the displacement. Some verbs of change of position conflate it:⁴

- (7) a. El vent va escampar els fulls (The wind aux. scatter_inf the leaves)
 b. El bibliotecari va apilar els llibres (The librarian aux. pile_inf the books)

Both types of manner can also be expressed syntactically or be left unspecified.

• *Instrument*: means by which the entity moves. It is usually expressed in predicates of motion, lexically (8a) or syntactically (8b):

- (8) a. El Jordi va esquiar fins al poble (art. Jordi aux. ski_inf to the town)
 b. El Jordi sempre viatja amb autobús (art. Jordi always travels by bus)

4. Realization of meaning components

The basic meaning components (entity, originator, and path) can be realized in the predicate by several means: they might appear as verbal arguments, lexicalized or incorporated (Talmy 1985), focused (Ikegami 1988), or understood (Fillmore 1986). Each of these cases will be treated separately for each component.

4.1. Argument realization

This section deals with the syntactic realization of meaning components. There are two types: simple and complex. In the simple type, each argument is related to a meaning component (3c). In the complex, three cases are distinguished: (i) An argument can be related to two components. In (3a) *El Robert* is both originator and path_source -in which case we will say that the components are coindexed. (ii) Two or more syntactic constituents might belong to the same meaning component. In (3b) each of the two PPs expresses one of the subcomponents of the same component. (iii) Two syntactic constituents might belong to the same meaning subcomponent. In (15) both PPs refer to the same semantic subcomponent⁵.

- *Entity*: the entity may appear coindexed with the originator component in verbs of position (9a) and verbs of motion (9b), when the entity that shifts moves autonomously. When the entity and the originator are coindexed, the entity always has the subject function (9a, 9b). If the originator and the entity are realized in separate syntactic constituents (9c), the originator has the subject function and the entity the object function:

- | | | |
|-----|---|--|
| (9) | a. <i>La nena s'ha assegut a la cadira</i>
b. <i>El grup va sortir a l'escenari</i>
c. <i>El jurat donà un premi a l'escriptora</i> | (The girl pron.aux. sit_part on the chair)
(The band aux. come_out on the stage)
(The jury gave a prize to the writer) |
|-----|---|--|

- *Originator*: has the subject function in transitive constructions. The originator and the shifting entity may be coindexed in verbs of autonomous motion. In verbs of change of position and nonautonomous movement the originator is never coindexed with the entity:

- | | |
|------|--|
| (10) | <i>El botiguer va posar les llaunes a la lleixa</i>
(The shop_assistant aux. put_inf the cans on the shelf) |
|------|--|

In change of possession verbs the originator is frequently coindexed with a subcomponent of the path. Thus, in some verbs of reception of a possession (11a) the originator is coindexed with the path_goal, since it causes the entity to go to the recipient and in verbs of releasing of a possession (11b), the originator is coindexed with the path_source, since the originator causes the entity to go from a source to a goal:

- | | |
|------|---|
| (11) | a. <i>El lladre ha robat el quadre del museu</i>
(The thief aux. steal_part the picture from_the museum) |
|------|---|

- b. *L'empresa exporta productes al Brasil*
(The firm exports products to_the Brazil)

• *Path*: in motion verbs the path can appear expressed in only one constituent. It will never have the subject function and its content might be expressed in different ways: with a noun phrase whose head expresses a physical entity referring to the total distance (12a), or with a noun phrase expressing the displacement in terms of surface measures (12b):

- (12) a. El nedador creuà *la piscina* (The swimmer crossed the swimming pool)
b. El corredor va córrer *dos quilòmetres* (The runner aux. run_inf two kilometers)

Furthermore, each path subcomponent can be realized in separate phrases in motion verbs (13a), verbs of change of position (13b), and verbs of change of possession (13c):

- (13) a. Els nens van anar *del campament* (source) *fins al cim* (goal) *pel refugi* (int.)
(The boys aux. go_inf from_the camp to_the top through_the cabin)
b. El professor va portar l'ordinador *de casa seva* (source) *al despatx* (goal)
(The teacher aux. take_inf the computer from home his to_the office)
c. *El propietari* (source) va llogar el pis *a uns estrangers* (goal)
(The landlord ux. rent_inf the apartment to art. foreigners)

A path subcomponent has the subject function only when it coincides with the originator. This happens in verbs of change of possession (*El propietari* in 13c). When it does not coincide with the originator, it is frequently expressed by a complement introduced by a preposition (13a, 13b and *a uns estrangers* in 13c). In verbs of change of possession this PP is optional (13c). In motion verbs and verbs of change of location the PP must sometimes be present (the absence of the path goal in (14a) and (14b) produces an ungrammatical sentence).

As for the sentences where not all the path subcomponents are expressed, while most verbs of change of position do not accept the specification of the subcomponent not expressed (path_source in 14a), some verbs of motion do (path_source in 14b):

- (14) a. El Sergi ha ficat l'abric (*de la cadira**) (*a l'armari**)
(art. Sergi aux. put_part the coat (from the chair*) (in the wardrobe)*)
b. El Sergi ha anat (*de Lyon*) (*a París**)
(art. Sergi aux. go_part (from Lyon) (to Paris)*)

A last case of realization of the path component is that in which a subcomponent is expressed in two constituents, so that it is underspecified. In (15) *a la Maria a Barcelona* expresses the path_goal subcomponent:

- (15) El Joan va enviar una carta *a la Maria a Barcelona*
(art. Joan aux. send_inf a letter to art. Maria to Barcelona)

4.2. Lexicalization and focalization

Lexicalization is a phenomenon by which a lexical item incorporates a meaning component (incorporation). This phenomenon might frequently correspond to a morphological process that consists of the derivation of a verb from the noun being lexicalized (the verb *ensacar* 'to put in a bag' (16a) comes from adding two affixes to the noun *sac* 'bag').

Neither the shifting entity nor the originator is lexicalized in any of the shift verbs. The path can be lexicalized in many change of position verbs (16a) and in verbs of assuming a position (16b). In all of these cases what is lexicalized is the *path_goal*:

- (16) a. El pagès *ensacà* les patates (The farmer *put_into_bags* the potatoes)
 b. El pacient *va allitar-se* (The patient *aux. took_to_his_bed_inf*)

Focalization is what affects verbs like *anar* 'to go' or *deixar* 'to leave'. In these cases a subcomponent of the path is semantically profiled (*path_goal* and *path_source*, respectively). This kind of verb includes information about the boundaries of the path, which is not present in other shift verbs, like *caminar* 'to walk', where a path is simply understood.

The application of a different treatment in the cases of focalization and lexicalization is justified not only from the morphological point of view, but also by the syntax: those verbs lexicalizing a component can not express it syntactically, unless it is underspecified (17a) (Taulé 1995), while in the cases of focalization, the focused element is often compulsory (17b), except for the examples of deixis, in which it can be ignored (see section 4.3):

- (17) a. El pagès *va ensacar* les patates *a un sac (gran)**
 (The farmer *aux. put_into_bag_inf* the potatoes *into a bag (big)**)
 b. El pagès *va treure* les patates *(del sac)**
 (The farmer *aux. take_out_inf* the potatoes *(from_the bag)**)

4.3. Underspecification

This section treats those cases in which the meaning components are not specified at the surface level. Three main circumstances are observed in which this happens: (A) the interpretation of that component is generic. (B) The context allows the recovery of the unexpressed content, be it (B.1) either because the constituent has been cited in the previous discourse (linguistic context) the information is known to the speaker (social and particular context), or (B.2) because of deixis (spatio-temporal context). (C) In the case of subcomponents, however, it happens that some of them have to be necessarily understood (in A and B the underspecification is optional).

- **Entity**: only a very reduced number of verbs admit its underspecification. They are change of possession verbs with cognate object (type A):

(18) La gent jove compra a botigues barates (The people young buy in stores cheap)

- **Originator**: it can only be understood in intransitive verbs (in anticausative and passive alternations) (19b) and in some verbs of reception of possession (20):

(19) a. *La ventada* ha desplaçat el cotxe (The gale aux. displace_part the car)
 b. Els cotxes s'han desplaçat (The cars pron. aux. displace_part)

(20) L'inversor va recuperar les accions (The investor aux. recover_inf the shares)

In (19b), the meaning component which is not syntactically realized can either be recovered by the context (type B1) or can be understood in a generic way (type A). In (20) the reading is ambiguous. In one interpretation, the phrase occupying the subject position refers to the originator of the action (argumental realization), while in the other, the action can be originated by another agent (type A or B1).

- **Path**: many shift verbs admit underspecification of the path, although there are exceptions. Some verbs can leave all but one of the path subcomponents unspecified. Among these are change of possession verbs (type A or B1), where either the source_path (21a) or the goal_path (21b) usually coincides with the originator in subject position:

(21) a. *L'agència* vengué el pis (*a la Laura*) (The agency sold the flat to art. Laura)
 b. *La Laura* comprà el cotxe (*al Joan*) (art. Laura bought the car from_art. Joan)

As previously mentioned (section 4.1), some verbs from the change of location and motion groups require the presence of one of the subcomponents and do not accept the expression of the others (14). In other cases the path component might be left unspecified, both when it is expressed in a single constituent (22a, 22c) and when expressed in more than one (22b) (type A or B1):

(22) a. El nen va creuar (*el carrer*)
 (The boy aux. cross_inf the street)
 b. El camió arrossega el cotxe (*des del mig fins a la vorera*)
 (The truck drags the car (from the middle to the sidewalk))
 c. El vaixell s'enfonsà (*uns metres*)
 (The ship pron. sank some meters)

Some verbs requiring the argumental expression of one of the path subcomponents might not express it in cases of deixis (B2). For example, in *arribar* 'to arrive' the path_goal is not expressed when it coincides with the location of the speaker:

- (23) a. El convidat ha arribat (aquí) (The guest aux. arrive_part (here))
 b. El convidat ha arribat (*a Tòquio*)* (The guest aux. arrive_part (to Tokyo)*)

Arribar usually requires the expression of the path_goal, although it might also accept the syntactic expression of the path_source. However, because of the deixis phenomenon, sometimes only the path_source is expressed, in spite of the fact that it is not the profiled subcomponent:

- (24) La Núria ha arribat *de París* (art. Núria aux. arrive_part from Paris)

More verbs with deixis are *deixar* 'to leave', *portar* 'to bring', *apropar* 'to bring/go nearer', *allunyar* 'to bring/go further', *baixar* 'to go/put down', *pujar* 'to go/put up', and *venir* 'to come'. The verb *venir* allows the omission of the path component in every usage, as it is a highly deictic verb:

- (25) a. Vine! (Come!)
 b. Vindràs demà? (Future_come_you tomorrow?)
 c. Ahir vas venir massa aviat (Yesterday aux. _you come_inf too early)

5. Syntactic behavior of shift verbs as a class

We have tried not to assign a set of verbal alternations to general classes in order to justify their existence. Instead, we have assumed that the acceptance of only one alternation suffices, if it is relevant for the meaning components.

Thus, although the realization of meaning components changes depending on the verb and the nature of each meaning component (see section 4), a common syntactic alternation is accepted by the shift verbs: the *path alternation*. It consists of the possibility of syntactic expression of either the path component or, at least, one of the path subcomponents, as a verbal complement. It is accepted by motion verbs (autonomous and nonautonomous) (26a, 26b), change of possession verbs (26c), and change of position verbs (26d):

- (26) a. L'empresa es traslladà *a Madrid*
 (The firm pron. moved to Madrid)
 b. L'empresa traslladà un empleat *a París*
 (The firm transferred an employee to Paris)
 c. L'estudiant va obtenir una beca *del govern català*
 (The student aux. get_inf a grant from_the government Catalan)
 d. La noia col·locà el llibre *a la taula*
 (The girl placed the book on the table)

Some special cases in the acceptance of the path alternation have to be mentioned. For a syntactic construction to participate in an alternation, it is necessary that it have a counterpart construction. Following this criterion, the path alternation should have two variants, one with, and another without, the phrase expressing the path component:

- (27) a. The company moved
b. The company moved to Morocco

A few verbs in the shift class do not accept the omission of the path component; they only accept the variant (27b). However, this fact does not invalidate our assumption that the path alternation gives syntactic support to the semantic class of shift verbs. We consider that arbitrary phenomena like focalization, lexicalization or conflation may cause some verbs not to show this alternation. Examples of this behavior are the verbs *anar* 'to go' (28a) and *posar* 'to put' (28b). They must syntactically express the path_goal:

- (28) a. Els estudiants anaren (*al congrés*)* (The students went (to the conference)*)
b. El jugador posà una carta (*a la taula*)* (The player put a card (on the table)*)

Thus, we will maintain the *path alternation* as characteristic of the shift class.

6. Conclusions

Starting from the work by Levin's *English Verb Classes and Alternations* (1993), we have presented a proposal of semantic classification for a group of verbs, namely *shift* verbs. This class is basically composed of motion verbs, change of possession verbs, and verbs of change of position. The grouping is based on the hypothesis that verbs belonging to the same class will share both the same meaning components and a characteristic syntactic behavior. As for the meaning components, we have defined three basic ones: entity, originator, and path. Regarding syntax, shift verbs admit the *path alternation*.

Thus, although our study confirms the hypothesis formulated by Levin, the results of our research differ from hers in two aspects: firstly, the class we propose includes verbs that in Levin's classification belong to different classes. Secondly, we have not associated a list of alternations with the class. We consider that the realization of meaning components is affected by phenomena like lexicalization or focalization, which prevent the verbal classes from showing a fine-grained common syntactic behavior. This is why only one characteristic alternation has been associated with the class, namely the path alternation, which is thought to be relevant because it affects one of the basic meaning components.

As previously mentioned, this work is framed in a wider project of study and classification of Catalan and Spanish verbs. As research goes on, verbal groups are being completed and new groups arise. In the case of the shift class, it seems possible to include communication verbs, since they share the same meaning components and the same relevant syntactic alternation.

Notes

¹ This work has been supported by the DGICYT project PB-940830, the DGICYT action APC-96 0125, the ITEM TIC-96 1243-C03-02 project, the Acalex (29.12.95) project of La Paeria and two postgraduate scholarships given by Comissionat per a Universitats i Recerca de la Generalitat de

Catalunya (ref. FI 96/6.008 PG and FI 97/00306 PG). We would like to stress that this work has been possible thanks to the discussion with the other members of the Pirapides project: M. Antònia Martí, Irene Castellón, and Ana Fernández. We are also grateful to those colleagues who revised the English version: Ana Fernández, Ana Ferrero, Abby Levinson, Gabriel Oreggioni and Matthew L. Juge. However, we are solely responsible for any possible errors.

² Besides, the exclusion of these verbs, which we consider change of state, is coherent with one of the general tendencies of the shift class: the shifting entity is never conflated.

³ We could consider that the specification of time acts as a subspecification of the path component, more than as a modification of the predicate.

⁴ The verbs in (7) belong also to change of state verbs, since there is a double focalization: the new state and the new place. The same happens to verbs of Levin's class "Assuming a position" (50) (agenollar-se 'to kneel', seure 'to sit'...).

⁵ When an exchange of entities is expressed, like in 'la Montse (A) intercanvia llibres amb la seva germana (B)' (art. Montse exchanges books with art. her sister), we consider that there are two paths. In one, A is the path_source and B the path_goal; in the other, it is the other way around. There are also two events: one in which A's books shift and the other in which the shift is of B's books. These are special cases of double realization. Verbs like *substituir* 'to substitute', *bescanviar* 'to change', *pagar* 'to pay' and *cobrar* 'to collect' are also double path verbs.

References

- Fernandez, A., M. A. Martí, and G. Vázquez. 1998. *Eventualities of change*. Forthcoming.
- Fernández, A., A. Martí, G. Vázquez, I. Castellón, & R. Morante. 1998. Criterios de una clasificación de las estructuras eventivas del cambio y la transferencia. *XVI Congreso de AESLA*. Logroño: Universidad de la Rioja.
- Fillmore, Charles. J. 1986. Pragmatically Controlled Zero Anaphora. *BLS* 12. 95-107.
- Ikegami, Y. 1973. A Set of Basic Patterns for the Semantic Structure of the Verb. *Linguistics* 117. 15-58.
- Ikegami, Y. 1988. Transitivity: Intransitivization vs. Causativization: Some Typological Considerations Concerning Verbs of Action. *On Language: Rethorica, Phonologica, Syntactica*, ed. by C. Duncan-Rose and T. Vennemann. 389-401. London: Routledge.
- Jackendoff, Ray. 1976. Toward an Explanatory Semantic Representation. *LI* 7. 89-150.
- Jackendoff, Ray. 1983. *Semantics and Cognition*. Cambridge, Mass.: MIT Press.
- Jackendoff, Ray. 1990. *Semantic Structures*. Cambridge, Mass.: MIT Press.
- Levin, Beth. 1993. *English Verbs Classes and Alternations. A Preliminary Investigation*. Chicago: The University of Chicago Press.
- Martí, M. A., R. Morante, J. Tió and G. Vázquez. 1997. Anàlisi del comportament verbal en el marc de la teoria dels esquemes semàntics. *Bulletí de l'ACIA*, 12, 7-14.
- Talmy, Leonard. 1985. Lexicalization Patterns: Semantic Structures in Lexical Forms. *Language Typology and Syntactic Description III: Grammatical Categories and the Lexicon*. Ed. by T. Shopen. 57-149. Cambridge, Mass.: MIT Press.
- Taulé, Mariona. 1995. *Representación de las entradas verbales en una base de conocimiento léxico: Diátesis y semántica léxica*. PhD Dissertation. Barcelona: Universitat de Barcelona.
- Vázquez, G. & A. Martí. 1998. *Els verbs de canvi d'estat*. Manuscript.

ANNEX: LIST OF SHIFT VERBS IN CATALAN (BASED ON LEVIN'S CLASSES)

(9) Verbs of putting: (9.1) col·locar, deixar, dipositar, disposar, embeinar, embossar, embotellar, emmagatzemar, empresonar, encabir, encaixar, encaixonar, encarcerar, encastar, enfilar, enfundar, engabiar, engarjolar, ensacar, entaforar, envasar, ficar, guardar, instal·lar, introduir, posicionar, posar, situar, tancar, (9.2) enfondir, emperxar, encimbellar, enfonsar, ensorrar, enterrar, juxtaposar, organitzar, penjar, reclinar, recolzar, seure, submergir, superposar, (9.3) abocar, apinyar, descarregar, empenyer, encaixar, escampar, escorrer, rampinar, rascлар, reunir, submergir, vessar, (9.4) abaixar, aixecar, alçar, arriar, baixar, calar, despenjar, elevar, hissar, llançar, llençar, (9.5) arremolinar, enroscar, enrotllar, girar, rodar, trenar, (9.7) acumular, agarberar, amuntegar, apartar, apilar, apilonar, apilotar, aplicar, apujar, arraconar, atapeir, carregar, encabir, entaforar, espargir, estendre, estirar, ficar, injectar, introduir, muntar, pujar, replantar, (9.8) encastar, escampar, incrustar, intercalar. **(10) Verbs of removing:** (10.1) allunyar, desallotjar, desarrelar, descarregar, distanciar, expel·lir, expulsar, extirpar, extre/aure, llançar, llençar, llevar, recollir, retirar, sostre/aure, tirar, treure, (10.2) deportar, enretirar, evacuar, extradir, (10.3) enretirar, escorrer, evacuar, (10.4) absorbir, arre/ancar, arxivar, aspirar, gratar, rascar, raspar, succiónar, xuclar, (10.5) aconseguir, agafar, apoderar-se, apropiat-se, arrabassar, arre/ancar, comissar, confiscar, decomissar, embargar, emportar-se, endur-se, extre/aure, furtar, pispar, prendre, rapir, raptar, recobrar, recuperar, reemborsar, reembossar, rescatar, retirar, retornar, robar, tornar, treure, usurpar. **(11) Verbs of sending and carrying:** acostar, apropar, arrossegar, atansar, carregar, carregar, carretejar, deportar, despatxar, dur, embarcar, empentar, empènyer, entregar, enviar, esmunyir, estirar, expedir, importar, llançar, llençar, lliurar, moure, mudar, percebre, portar, rebre, remetre, remolcar, retornar, tirar, tornar, tragar, trametre, transbordar, transferir, transportar, traslladar. **(12) Verbs of exerting force:** arrossegar, empentar, empentejar, empènyer, impel·lir, pitjar, polsar, premer, pressionar, propel·lir, propulsar, tirar. **(13) Verbs of change of possession:** abastar, abonar, aconseguir, adjudicar, administrar, adquirir, agafar, anticipar, aportar, arrencar, arrendar, atorgar, avançar, bescanviar, canviar, carregar, cedir, cobrar, comissar, comprar, concedir, concedir, conferir, confiar, confiscar, contribuir, decomissar, deixar, deixar, demanar, desembossar, distribuir, donar, dur, entregar, enviar, estendre, guanyar, intercanviar, llegar, lliurar, llogar, malvendre, obtenir, oferir, pagar, passar, permutar, portar, prendre, presentar, prestar, proporcionar, recobrar, recol·lectar, recuperar, reembossar, reemplaçar, regalar, remetre, repartir, retornar, retre, robar, servir, servir, subministrar, substituir, tornar, trametre, transferir, traspassar. **(17) Verbs of throwing:** bolcar, catapultar, empènyer, llançar, llençar, passar, tirar, tombar, xutar. **(50) Verbs of assuming a position:** abocar-se, agenollar-se, agotnar-se, aixecar-se, ajeure's, ajupir-se, allitar-se, arraulir-se, arrepapar-se, arronsar-se, asseure's, desplomar-se, doblegar-se, encamellar-se, encongir-se, enfondrar-se, enfonsar-se, escarxofar-se, estendre's, inclinar-se, incorporar-se, jeure, llevar-se, reclinar-se, recolzar-se, repenjar-se, seure's. **(51) Verbs of motion:** abalançar, abandonar, adreçar-se, agrupar, aixecar, alçar, allunyar, anar, anar-se'n, apartar, apinyar, apropar-se, aproximar-se, arribar, arrossegar, ascendir, avançar, baixar, balancejar, balandrejar, ballar, barrejar, belluguejar, bordejar, botar, bressar, bressolar, caminar, capalçar, caure, coixejar, conduir, correr, costejar, creuar, cuitar, dansar, davallar, deambular, deixar, descaminar, descendir, desertar, desfilar, desviar-se, dirigir-se, elevar, embocar, encabritar-se, endinsar, enfonsar-se, entrar, escalar, escapar, escapollir-se, esfondrar, esquiar, esquitllar-se, evadir-se, fer cap, fluir, fugir, galopar, gambar, gatejar, girar, giravoltar, grimpar, gronxar, internar, introduir, llançar, lliscar, marxar, muntar, navegar, nedar, oscil·lar, partir, passejar, patinar, pedalar, pedalejar, planar, precipitar, pujar, ranquejar, rebotar, recorre, regirar, remar, remenar, retirar, retornar, retrocedir, rodolar, rondar, rotar, sabatejar, saltar, saltironejar, serpentejar, sortir, submergir-se, titubar, tombar, tornar, travessar, trontollar, vacil·lar, vagabundejar, vagar, vagarejar, venir, viatjar, vogar, volar, voltar, voltejar, vorejar, zigzaguejar.

Linguistics and 'The Linguistic Turn': Language, Reality, and Knowledge

Eve Ng

State University of New York at Buffalo

1. INTRODUCTION. The majority of linguists remain unfamiliar with what has been called 'the linguistic turn', a theoretical development integral to postmodern social theory, which has been prominent in many of the humanities. This is partly because linguistics as a discipline resembles a 'hard' science in various respects, and hence has maintained a certain theoretical distance from the other social sciences. Nevertheless, language is a central theme for both linguists and postmodern scholars, and significant claims about it are centrally associated with the linguistic turn. We can, therefore, ask whether these theoretical implications should be of concern to linguists, and what sort of input linguists can make to this discussion.

In a nutshell, the linguistic turn refers to a theoretical shift on how the relationship between reality and representation (or 'signification') is viewed. Whereas the familiar modernist paradigm regarding scientific endeavor delineates a clear boundary between the (human) subject conducting the study, the object of study, and the language used to describe it, postmodernists have put into question the separability of all three of these. In particular, representation is not thought of as simply reference to or description of reality, but rather, signification acts are argued to play an integral part in constituting the perception of the object, or perhaps the object of representation itself.

This paper, then, seeks to address two questions: 1. How do linguist notions of linguistic determinism compare with postmodern positions about language and reality? 2. What perspectives can linguists take on the postmodern claim that language is an inescapable component of the foundations for scientific knowledge?

2. 'LANGUAGE'. It is useful to start by clarifying how the term 'language' is understood by linguists and postmodernists.

In contemporary linguistics, much of the focus has been on finding regular patterns in the **structure** of natural human languages — sound systems, words, phrases and sentences¹. These generalizations are sought in linguistic data, i.e. occurring or occurrable bits of human languages, making linguistics a fundamentally empirical discipline.

By comparison, postmodernists are almost exclusively concerned with the **function** of language rather than its structure, with the exception of its semantic or semiotic structure. Whereas linguists typically characterize the function of language as a tool for interpersonal communication, postmodernists have a broader conception of communicative function, seeing language as necessarily involved in important ways in the socialization of individuals with respect to how social norms, political values, identity roles, and so on, become part of who people are.

Furthermore, for linguists, instances of **language use** are more often than not thought of as the actual utterances that are produced, while postmodernists tend to focus on the fact of the act or process of signification. As a point of similarity, however, both postmodernists and linguists have included as part of their concern

with language the names or labels that are given to phenomena and the influence these have on how we understand the phenomena.

To summarize the main notions from both disciplinary approaches, we have 'language' considered as:

a. Structure.

b. Instances of use, including —

(i) acts of speech (utterances)

(ii) acts of labeling (entities, phenomena, etc.)

(iii) acts/processes of representation or signification.

This paper will focus on 'language' in the senses of *b*, looking in particular at claims made about labeling (in Section 3) and about representation/signification (Section 4). In Section 3, a comparison of the discussions in linguistics and postmodern theory will shed light on the fact that linguistic determinism does not mean the same in each approach. Furthermore, in Section 4, a consideration of why languages have the characteristics they do will suggest that postmodernism's attack on modernist epistemology and the role it accords to systems of signification are fundamentally problematic for linguistics.

3. LANGUAGE, THOUGHT, AND REALITY. This section sketches out ways that the relationships between language, thought, and reality have been conceived of, a sketch that will involve a certain degree of simplification for the sake of expository simplicity.

To begin, consider a "naive" or "commonsensical" view of these relationships, namely that there are things or phenomena out in the world, we perceive them, and we talk about them through language. This is the familiar modernist episteme, sometimes called objectivist metaphysics, and is diagrammed in Fig. 1:

[1] Modern episteme (objectivist metaphysics)

perceived by coded by
reality -----> **cognition** -----> **language**

A key characteristic of postmodernism has been its challenges to modernist epistemology and the validity of scientific objectivism on various accounts. In this vein, it is argued that language does not merely code a reality that exists independent of it, but rather, language is one of the ways by which that 'reality' comes to be seen as real. With respect to the relationship between language and reality, then, the directionality is the opposite of what the modernist episteme assumes, and language reflects, if anything, the workings of *power* in society [Fig. 2]:

[2] Postmodern position on language and reality

work through constructs
sociopolitical forces -----> **language** -----> **(perception of)**
reality

A number of linguistic anthropologists have applied aspects of postmodern social theory to their research. For example, Susan Gal argues in [3] that **linguistic behavior** does not merely reflect a speaker's social identities — categories such as class, gender, and sexuality — but helps construct (i.e. create) them.

[3] ... the categories of *women's speech*, *men's speech*, and *prestigious* or powerful speech are not just indexically derived from the identities of its speakers. Indeed, sometimes a speaker's utterances create her or his identity. These categories, along with broader ones such as *masculine* or *feminine*, are created within social groups... (Gal 1995, italics in original)

A related claim, as, for example, Eckert & McConnell-Ginet (1995) assert, is that the **labels** that languages have for social categories create those categories as realities for speakers, rather than merely naming preexisting concepts:

[4] Language is a primary tool people use in constituting themselves and others as 'kinds' of people... Social categories and characterizations are human creations; the concepts associated with them are not preformed, waiting for labels to be attached, but are created, sustained, and transformed by social processes that importantly include labeling itself. (Eckert & McConnell-Ginet 1995)

These quotes recall the linguistic determinism that Edward Sapir and Benjamin Lee Whorf explicated earlier this century (and there have been a number of recent studies and reexaminations of the Sapir-Whorf hypothesis within linguistics, e.g. Lucy 1992; papers in Gumperz & Levinson 1996). The passages in [5] and [6] below express the now-familiar claim that how we understand a "reality" outside of our minds is influenced by the language that we speak — the vocabulary, grammatical categories, and syntactic structures that it has.

[5] To pass from one language to another is psychologically parallel to passing from one geometrical system of reference to another. The environing world is the same for either language; the world of points is the same in either frame of reference. But the formal method of approach to the expressed item of experience, as to the given point of space, is so different that the resulting feeling of orientation can be the same neither in the two languages nor in the two frames of reference. (Sapir 1924)

[6] We dissect nature along lines laid down by our native languages. The categories and types that we isolate from the world of phenomena we do not find there because they stare every observer in the face; on the contrary, the world is presented in a kaleidoscopic flux of impressions which has to be organized by our minds — and this means largely by the linguistic systems of our minds. (Whorf, 1956: 213)

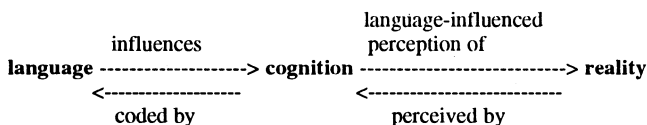
Sapir and Whorf differed on whether there is a fixed reality independent of human perception, but expressed similar ideas about how different languages can by their influence on cognition lead to different experiences of the external world, a position that is represented in Fig. [7].

[7] Sapir-Whorf linguistic determinism position



Since, actually, the majority of linguists sympathetic to some version of the Sapir-Whorf hypothesis also subscribe to the validity of the modernist conception, Fig. [7'], combining [1] and [7], is probably a more representative summary of the position of these linguists.

[7'] Combined modernist/Sapir-Whorfian position



Comparing this diagram with the postmodern position summarized in Fig. [2], three differences can be noted. Firstly, the role of the mind or cognition tends to be backgrounded in postmodern views (hence the parentheses around 'perception of' in Fig. 2), often with no clear distinction drawn between claims about language determining the perception or understanding of the world, and language constructing reality itself. This point will be discussed further in Section 4.

Secondly, the quotes of Gal in [3] and Eckert & McConnell-Ginet in [4] illustrate a somewhat more dynamic version of linguistic determinism, in that it is language use, and language use as social process, that is posited to be constructing what is perceived as reality. Traditional interpretations of the Sapir-Whorf hypothesis, in contrast, focus on the linguistic items and structures in a language.

Thirdly, in looking at semantic fields, both linguists and postmodernists have offered explanations about how **culture** contributes to what items are found in language, along the lines that languages code what is 'culturally relevant', but there is a clear difference in how 'dynamic' a perspective of culture is taken. Linguists, at least traditionally, have had a fairly apolitical, non-dynamic conception of society and culture. Consider the following passage from Sapir, listing the lexicalized terms that Paiute, a Native American language of the Southwestern U.S., has for geological features:

[8] ... divide, ledge, sand flat, semi-circular valley, circular valley or hollow, spot of level ground in mountains surrounded by ledges, plain valley surrounded by mountains, plain, desert, knoll, plateau, canyon without water, canyon with creek, wash or gutter, gulch, slope of mountain or canyon wall receiving sunlight, shaded slope of mountain or canyon wall, rolling country intersected by several small hill-ridges. (Sapir 1949a: 91, in Bonvillian 1993: 54)

A linguist is apt to cite data like this to argue that Paiute speakers happened to live a particular lifestyle in a particular habitat which made distinguishing those formations culturally important; hence, we can say that language, reflecting culturally relevant aspects of the natural environment, comes to have labels for them, and these labels then influence how such 'realities' are perceived by speakers, with lexicalization helping to crystallize aspects of the perceptual world as entities in the mind.

In comparison, consider the passage in [9] from the French social theorist Michel Foucault, where he argues that what counts as 'truth' is subject to the dynamics of power in society:

[9] Truth is a thing of this world: it is produced only under multiple forms of constraint. And it induces regular effects of power. Each society has its regime of truth, its 'general politics' of truth: that is, the types of discourse which it accepts and makes function as true; the mechanisms and instances which enable one to distinguish true and false statements; the means by which each is sanctioned; the techniques and procedures accorded value in the acquisition of truth; the status of those who are charged with saying what counts as true. (Foucault 1972: 131)

Hence, for example, in his (1978) consideration of why the West has the concepts of sexuality ('homosexual', 'heterosexual', etc.) that it does, Foucault's position is not that such types of categorization merely happen to reflect culture; rather, cultural forces of sociopolitical power are crucially implicated in the production and maintenance of what is taken as knowledge, truth, or reality. Hence, culture, working partly through language, has a productive power that enforces the categories that language contains.

To summarize, although we see basic similarities between the classic Sapir-Whorf hypothesis on linguistic determinism and postmodern positions about language and reality, there are also a number of important differences. Linguists, unlike postmodernists, do not reject the modernist episteme, while postmodernists, unlike linguists, do not always clearly figure in the place of perception or the mind as separable from either language or reality. Also, the postmodern position attributes a more dynamic role to both language, in articulating how it shapes reality through use rather than just by virtue of vocabulary and grammar, and to culture, as something which is not just passively reflected by language but which intertwines with power relations that are integral to the determination of 'reality'.

4. LANGUAGE AND EPISTEMOLOGICAL FOUNDATIONS. Another, perhaps more fundamental, difference in approaches is that linguists have not looked just for purely culture-dependent explanations to account for why languages have the items and structures they do, since linguists (from both generative and functional perspectives) are typically interested in generalizing across languages. The main types of explanations for cross-linguistic similarities that have been offered are summarized below:

1. The innateness hypothesis, which underlies generative linguistics, proposes that there is a biologically programmed knowledge of grammar (linguistic structure) common to speakers of all languages.

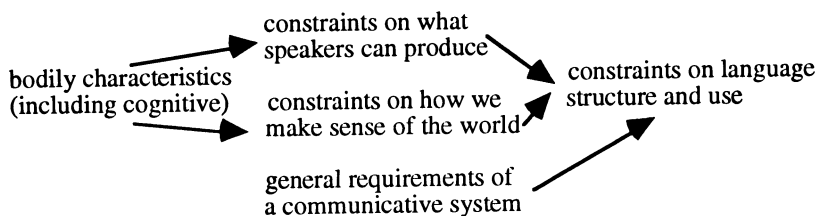
2. The common communicative functions of language, which, being sensitive to constraints on the human perceptual, articulatory, and cognitive apparatuses, are proposed to explain why certain kinds of linguistic patterns — morphosyntactic,

phonological, and phonetic — are more frequent in language data (e.g. Croft 1990; Mayerthaler 1988; Maddieson 1984).

3. Certain characteristics of the human body are argued to shape how we perceive and understand the world, and hence what we code in language. The classic studies are those by Brent Berlin, Paul Kay and their colleagues (Berlin & Kay 1969; Kay & McDaniel 1978) in establishing near-universals of lexicalization patterns for color terms that can be traced to shared aspects of human neurophysiology. Also, cognitive linguists such as G. Lakoff (1987), Talmy (1988), and others have argued that characteristics of our bodies and minds crucially influence how we experience the world, reflected not just in patterns of lexicalization but also those of metaphorical extension.

These explanations are diagrammed in Fig. [10].

[10] Linguistic universalist explanations for language characteristics



Thus, for various areas of linguistics, the human body serves as an epistemological ground, i.e. a basis for explanation. Hence, with respect to the tension between linguistic relativism and universalism, many linguists would argue that linguistic determinism can be taken only so far, since surely there is not a unidirectional causality of language on perception of reality; rather, the non-linguistic world, for example in the form of our bodies, comes to shape language as well. However, the tenability of this position is challenged by two prominent postmodern claims:

1. There is no natural human body which is not already 'inscribed' or marked through with the cultural.

2. There can be no grounds for human knowledge outside systems of signification.

How these discussions relate to relevant areas within linguistic theory will be explored in the following two subsections.

4.1 THE HUMAN BODY. The concept of the 'natural', or more specifically, the nature-culture dualism, is part of a broader critique on binary oppositions that postmodern theory has identified and attacked. This section focuses on one strand of this discussion within postmodern feminist theory with regard to the human body.

Scholars such as Judith Butler and Elizabeth Grosz have sought to reconcile, on the one hand, the physicality or materiality of the body with, on the other hand, the rejection of the nature-culture and mind-body dualisms as well as the postmodern tenet that nothing is comprehensible outside of language. Sophisticated postmodern discussions of the body are careful not to claim that everything can be reduced to language or to effects of language. However, at the

same time, Butler and Grosz argue that the physical body, its materiality, cannot be thought of or understood outside of signifiatory systems:

[11] ... the body is neither brute nor passive but interwoven with and constitutive of systems of meaning, signification, and representation. [...] [The human body is not] a precultural, presocial, or prelinguistic pure body, but a body as social and discursive object, a body bound up in the order of desire, signification, and power. (Grosz 1994: 18-19)

[12] Language and materiality are fully embedded in each other, [...] but never fully collapsed into one another, i.e. reduced to one another, and yet never fully exceeds the other. Always already implicated in each other, always already exceeding one another, language and materiality are never fully identical or fully different. (Butler 1993: 69)

Grosz (1994) argues that the body has been equated with nature, and the mind with language, in ways that fail to recognize how these supposed dualistic opposites interact. She appeals for new terminology in talking about human bodies and subjectivity that goes beyond the mind-body dualism, suggesting the expressions 'embodied subjectivity' and 'psychic corporeality'. These terms are not unlike similar expressions that have arisen within cognitive semantics, which in certain respects is quite consonant with some parts of postmodernism. For example, G. Lakoff (1987) has argued for rejecting both objectivist metaphysics and the mind-body dualism, criticizing formal model-theoretic semantics for relying on an objectivist philosophy that there is a disembodied, 'God's-eye' point of view from which to attain the truth. Instead, in explicating the thesis of 'functional embodiment', Lakoff argues that what humans perceive and how we interact with the world depend crucially on our very humanness — the way our bodies are, the way our cognitive structures and processes are.

On the other hand, Lakoff explicitly denies that rejecting objectivist metaphysics leads inevitably to rampant relativism, because human beings, despite their differences, share certain commonalities of physicality and cognition. Hence, cognitive semantics is still grounded epistemologically in the specifics of the body, and for this reason remains fundamentally distinct from postmodern theory, which questions the possibility of a human body that is independent of 'language' (referring to systems of representation). This position derives from a more general argument about the role of the signifier in originating meaning, which we explore in the next sub-section.

4.2 THE DERRIDEAN SIGNIFIER. The writings of French philosopher Jacques Derrida, attacking both the content and method of Western philosophy and science, have become key works of postmodern theory (e.g. 1978, 1976, 1973). In his considerations about language, meaning, and reality, he seeks to subvert the major underpinnings of both phenomenology, which locates meaning in interior consciousness, and structuralism, for which meaning arises in the relations between the units of a signifiatory system. The following paragraphs give the briefest of summaries of some of the relevant ideas.

One central Derridean critique concerns what he terms "the metaphysics of presence" in Western philosophy, which assumes that we are able to perceive reality in an immediate way (i.e. without mediation). Instead, Derrida argues that perception happens in the unconscious, which is inaccessible, so that we can never derive meaning from perceptual processes; rather, all reflective or conscious

encounters with the world occur via the mediation of re-presentations, and hence, all knowledge is necessarily linguistic.

Derrida also draws on Ferdinand de Saussure's (1989 [1916]) discussion about the linguistic sign. Saussure argued that meaning is produced in both the formation of signs as entities consisting of both the signifier (e.g. the word) and the signified (the concept), and in the oppositions between elements in a signifiatory system. Whereas linguists have tended to unproblematically accept these Saussurean concepts, Derrida pushes Saussure's point that signifier and signified are indissolubly bound, and hence rejects the possibility of meaning or concepts existing independent of language. Derrida also develops the Saussurean notion of oppositional difference into his theory of the 'trace', making a complex argument that the meaning or significance of a signifier resides in what is absent as well as what is present with respect to its relationship with other signifiers in the signifiatory system. The crucial point here is that Derrida denies the possibility of meaning existing independent of signification.

Butler accepts Derrida's position about the impossibility of human conceptualization outside of language, and goes on to make the argument in [13] that positing something, such as the human body, to be extralinguistic (i.e. independent of signifiatory systems) is an impossible paradox, since any such act of positing is always an act of signification itself, and such an act will then become part of ('constitutive of') the concept. Therefore, nothing is material without being linguistic, at least as focuses of human contemplation.

[13] The body posited as prior to the sign, is always *posited* or *signified* as *prior*. This signification produces as an effect of its own procedure the very body that it nevertheless and simultaneously claims to discover as that which *precedes* its own action... To posit by way of language a materiality outside of language is still to posit that materiality, and the materiality so posited will retain that positing as its constitutive condition. (Butler 1993: 30, italics in original)

There are two different types of objections that can be made to this argument. The first is concerned with having a metalanguage to talk about phenomena, which Butler appears to be refusing to allow. If every propositional speech act about some phenomenon then makes that phenomenon 'linguistic', then it is unclear how much content this sense of 'linguistic' retains or whether it allows Butler to make a theoretically meaningful point about the presence of language in the material body. For example, Butler's assertion in [13] presumably means that the act of making the proposition, 'Unconscious reflexes such as the heat recoil reaction are non-linguistic' would by virtue of its signifiatory character make those unconscious reflexes which the proposition mentions 'linguistic'.

There are also logical problems associated with the strong anti-naturalist position that Butler takes. Language (meaning here any system of signification) evolved within the context of the particular bodies that humans (and proto-humans) had, so that we may surely be able to identify bodily characteristics which preceded the development of language as extralinguistic.^{2,3} In a related vein, Butler has also been criticized for being anthropocentric in not taking into consideration that non-human animate beings also have bodies, and it is unclear how language or systems of signification would construct or interact with the materiality of their bodies (Cheah 1996).

What many would argue to be a central problem in Butler's position, and Derrida's, comes back to their view of perception. Recall from Section 3 that the place of perception in the postmodern conception of linguistic determinism was somewhat unclear, and Derrida, for one, explicitly claims that 'there never was any "perception"' (1973: 103); what postmodernists deny is not the existence of some real world, but our ability to (consciously) experience it non-linguistically.

Arguments against this position have come from various directions, but contemporary linguists are typically not philosophers of language, and this paper will not try to summarize this discussion. However, let us just mention one critique which seems to fit in well with what linguists would find theoretically coherent, offered by Dillon (1995), a phenomenological philosopher who faults Derrida for 'semiological reductionism' (reducing everything to signs and their effects). Dillon makes two arguments which are summarized briefly here. Firstly, he argues that language which does not refer to anything outside of language would, from a logical standpoint, ultimately fail to be meaningful; however, we don't find language to be meaningless, and the reason that words can derive meaning from other words is because words do ultimately refer to the world (i.e. something outside the system of signification). Secondly, given that non-human animals are able to perceive and in some way understand the world, it is anthropocentric to argue that meaning derives only from signifiers, an argument related to one we have already mentioned by Cheah (1996) in critiquing Butler's discussion about how language constructs the body.

Linguists, then, may accept Derrida's claims that neither classical phenomenology nor structuralism are adequate to account for how language means, but are apt to reject the argument that perception cannot be discussed as part of the equation of deriving meaning. As we have seen in earlier sections, the human body plays a foundational, epistemological role in linguistic theory, and part of this role includes its mediation between language and the external world through perceptual processes. Hence, it should be clear that a Derridean position would be untenable to linguists, even those sympathetic to some version of the Sapir-Whorf hypothesis or to a Lakovian explication of cognitive semantics.

5. SUMMARY. These theoretical incompatibilities notwithstanding, we should not conclude without noting that there are aspects of postmodernism which many linguists would find appealing. One, as discussed in Section 3, is the rejection of a strict classical objectivist metaphysics by cognitive linguists such as Lakoff (1987). Another is that, in their attack on privileged discourses of knowledge, postmodernists often call for a plurality of perspectives to replace the application of one dominant worldview to everything, something which Rajagopalan (1993) notes has certain parallels with disagreements in various disciplines about the degree to which data should be treated on their own terms or worked into an existing general theory. Hence, in linguistics, more descriptivist-oriented scholars may welcome this general philosophy against theoretical chauvinism.

What we have seen here, however, is that the postmodern position on the relationships between language, reality, and knowledge diverges on several accounts from one which most linguists generally assume. In addition, postmodernists have put into question a concept of the natural that is separable from culture and language, something which linguistics relies on in fundamental ways. This paper has highlighted such differences between the two approaches, showing that sharing the common theme of 'language' has not meant the presence of basic theoretical commonalities. More recently, a number of linguists, particularly

within sociolinguistics and linguistic anthropology, have applied theoretical and political insights from postmodern theory in their work (e.g. see papers in Hall & Bucholtz 1995); how far and fruitfully such integrations can be taken must, however, be a question for another paper.

NOTES

¹ In social theory, 'language' often refers not just to natural human languages, but also to images and various other forms of representation. Also, the term 'discourse' in postmodern discussion is sometimes used interchangeably with 'language', but may encompass an even broader range of phenomena, including bodies of texts associated with particular recognized fields, various characteristics of social institutions such as the law, and other meaning-carrying aspects that characterize social organization (e.g. the use of physical space). These more extended uses of the term 'language' will not be explicitly discussed in this paper, since they have little or no intersection with uses of the term by linguists.

² While Butler concedes that materiality is implicated in language (see [12]), her reasoning is that this is because the sign is itself material in terms of its realization (e.g. as sound); however, as Eagleton (1996) notes, if even the sign is considered to be material, it is unclear how useful this sense of 'material' is.

³ At a recent (1997) talk entitled 'Naked', about the relationship between the evolution of vision and the development of art in human culture, Elizabeth Grosz, if I understood her correctly, made a similar point in criticizing some feminists for refusing to acknowledge that before there was culture, there was biology (i.e. vision preceded art).

REFERENCES

- Berlin, Brent & Paul Kay. 1969. *Basic color terms: their universality and evolution*. Berkeley: University of California Press.
- Bonvillian, Nancy. 1993. *Language, culture, and communication: the meaning of messages*. Englewood Cliffs, NJ: Prentice Hall.
- Butler, Judith. *Bodies that matter: the discursive limits of sex*. New York: Routledge.
- Cheah, Pheng. 1996. Mattering. *Diacritics* 26.1: 108-139.
- Croft, William. 1990. *Typology and universals*. New York: Cambridge University Press.
- Derrida, Jacques. 1978 [1967]. *Writing and difference*, tr. from French by Alan Bass. Chicago: University of Chicago Press.
- Derrida, Jacques. 1976 [1967]. *Of grammatology*, tr. from French by Gayatri Spivak. Baltimore: Johns Hopkins University Press.
- Derrida, Jacques. 1973 [1967]. *Speech and phenomena*, tr. from French by David B. Allison. Evanston, IL: Northwestern University Press.
- Dillon, M.C. 1995. *Semiological reductionism: a critique of the deconstructionist movement in postmodern thought*. Albany: State University of New York Press.
- Eagleton, Terry. 1996. *The illusions of postmodernism*. Cambridge, MA: Blackwell.
- Eckert, Penelope & Sally McConnell-Ginet. 1995. Constructing meaning, constructing selves: snapshots of language, gender and class from Belten High. *Gender articulated: language and the socially constructed self*, ed. by Kira Hall & Mary Bucholtz. New York: Routledge.
- Foucault, Michel. 1977 [1976]. *The history of sexuality*, tr. from French by Robert Hurley. New York: Pantheon Books.

- Foucault, Michel. 1972. Truth and power. Interview by Alessandro Duranti & Pasquale Pasquino, tr. from French by Colin Gordon. *Power/knowledge: selected interviews and other writings, 1972-1977.*, ed. by Colin Gordon. New York: Pantheon Books.
- Gal, Susan. 1995. Language, gender, and power: an anthropological review. *Gender articulated: language and the socially constructed self*, ed. by Kira Hall & Mary Bucholtz. New York: Routledge.
- Grosz, Elizabeth. 1994. *Volatile bodies: toward a corporeal feminism*. Bloomington: Indiana University Press.
- Gumperz, John & Stephen Levinson (eds). 1996. *Rethinking linguistic relativity*. New York: Cambridge University Press.
- Hall, Kira & Mary Bucholtz (eds.). 1995. *Gender articulated: language and the socially constructed self*. New York: Routledge.
- Kay, Paul & McDaniel, Chad K.. 1978. The linguistic significance of the meanings of basic color terms. *Language* 54.3: 610-646.
- Lakoff, George. 1987. *Women, fire, and dangerous things: what categories reveal about the mind*. Chicago: University of Chicago Press.
- Lucy, John. 1992. *Language diversity and thought: a reformulation of the linguistic relativity hypothesis*. New York: Cambridge University Press.
- Mayerthaler, Willi. 1988 [1981]. *Morphological naturalness*, tr. from German by Janice Seidler. Ann Arbor, MI: Karoma Publishers.
- Maddieson, Ian. 1984. *Patterns of sounds*. New York: Cambridge University Press.
- Rajagopalan, Kanavillil. 1993. 'Chameleon-linguist' and the revival of casuistry: afterthoughts on a controversial suggestion by Fernando Tarallo. *Revista de Documentacao de Estudos em Linguistica Teorica e Aplicada (D.E.L.T.A.)* 9 (special issue): 339-351.
- Sapir, Edward. 1931. Conceptual categories in primitive languages. *Language in culture and society: a reader in linguistics and anthropology*, ed. by Dell Hymes. New York: Harper & Row.
- Saussure, Ferdinand de. 1989 [1916]. *Course in general linguistics*, ed. by Charles Bally & Albert Sechehaye, with the collaboration of Albert Riedlinger, tr. from French and annotated by Roy Harris. LaSalle, IL: Open Court.
- Talmy, Leonard. 1988. Force dynamics in language and cognition. *Cognitive Linguistics* 12: 49-100.
- Whorf, Benjamin Lee. 1956. *Language, thought, and reality: selected writings of Benjamin Lee Whorf*, ed. by John B. Carroll. Cambridge, MA: MIT Press.

On Mohawk Ghost Vowels: Audibility vs. Visibility

Grażyna Rowicka
HIL/University of Leiden

Mohawk GHOST VOWEL *e* materialises to break up consonantal clusters. It has generally been analysed as epenthetic. The aspect of Mohawk phonology which has probably received most attention in the literature involves the intricate interaction between epenthesis and stress. Some ghost vowels are taken into account by stress, while others are not. The phonetic presence, i.e. AUDIBILITY, of ghost vowels is then not always accompanied by their VISIBILITY with respect to stress.

In this paper I propose an account of the apparently inconsistent stress behaviour of Mohawk ghost vowels. I will argue that ghost vowels materialise as a result of two independent sets of phonotactic well-formedness conditions: INTERNUCLEAR ones and INTERSEGMENTAL ones. Only the former type of conditions, being prosodic in nature, trigger the materialisation nuclei which are phonetically audible as well as stress-visible. On the other hand, intersegmental conditions can only determine the phonetic presence, i.e. the audibility of nuclei, but not their visibility to stress. Both types of well-formedness conditions which I will be talking about have been recognised in the literature on Government Phonology.

1. DATA

1.1 GHOST VOWELS IN MOHAWK

Mohawk is a Northern Iroquoian language, still spoken on reserves in Quebec, Ontario and New York State. The basic sources on Mohawk phonology are studies by Michelson (1981, 1983, 1988 and 1989) and that is where all the examples below come from. Other sources and studies on Mohawk include Alderete (1995), Beatty (1974), Bonvillain (1973), Chafe (1977), Hagstrom (1997), Mithun (1979), Piggott (1995), Postal (1968) and Potter (1994). (For a discussion of some of the analyses see Rowicka in prep.).

There are three vowels in Mohawk traditionally analysed as epenthetic which are phonetically identical to their non-epenthetic counterparts. The vowel *i* is added initially in subminimal words to satisfy a prosodic minimality condition on Mohawk verb forms. A well-formed verb form must minimally contain one branching foot (for a discussion see Piggott 1995). The vowel *a* appears between specific morphemes (as a 'stem joiner'). These vowels will not be considered in this paper. (For an analysis of the stem joiner vowel see Rowicka, in prep.) I will focus on the epenthetic vowel *e* since only this vowel exhibits the special prosodic behaviour to be discussed below. I will refer to it as the 'ghost vowel'.

The ghost vowel *e* materialises after the first member of a triconsonantal cluster or to split up a biconsonantal cluster if its second member is a sonorant /n, r, w/ or a word-final glottal stop /ʔ/. Ghost vowels fail to break up clusters containing /s, ʔ, h/ in certain positions.

Consider the examples in (1) where ghost vowels are capitalised.

- (1)
- | | | |
|---------------|-------------------------------|----------------------------------|
| a. | BETWEEN C AND WORD-FINAL /ʔ/: | |
| /ʌ-k-arat-ʔ/ | ʌká:ratEʔ | 'I lay myself down' |
| /ro-kut-ot-ʔ/ | rokú:totEʔ | 'he has a bump on his nose' |
| b. | BETWEEN C AND SONORANT: | |
| /ʌ-k-r-ʌʔ/ | ʌkErʌʔ | 'I will put it into a container' |
| /w-akra-s/ | wákEras | 'it smells' |
| c. | TO BREAK UP CLUSTERS: | |
| /s-rho-s/ | sÉrhos | 'you coat it with something' |
| /s-k-ahkt-s/ | skáhkEts | 'I got back' |

The acute accent indicates main stress. There is no secondary stress.

1.2 GHOST APPEARANCE VS. STRESS

In words with no ghost vowels main stress falls on the penultimate syllable. Stressed vowels in open syllables are subject to lengthening.

- (2)
- | | | |
|----------------|------------|---------------------|
| /s-atorat/ | sató:rat | 'hunt'-IMPERATIVE |
| /wak-ashet-u/ | wakashé:tu | 'I have counted it' |
| /k-atirut-haʔ/ | katirúthaʔ | 'I pull it' |
| /s-ho-ahkt-u/ | shóhktu | 'he went back' |

Apart from closed syllables, tonic lengthening is also blocked before some ghost vowels in the next syllable (see § 1.3 below). The ghost vowel is sometimes visible to stress and sometimes it is not.

- (3)
- | | | |
|---------------|------------------------------------|------------------------------------|
| a. | GHOST BEFORE A SINGLE CONSONANT | |
| /ʌ-k-arat-ʔ/ | ʌká:ratEʔ | 'I lay myself down' |
| /t-ʌ-k-rik-ʔ/ | tʌkErikEʔ | 'I will put together side by side' |
| b. | GHOST BEFORE A CONSONANTAL CLUSTER | |
| /wak-nyak-s/ | wakÉnyaks | 'I get married' |
| /s-k-ahkt-s/ | skáhkEts | 'I got back' |
| | *ÍskahkEts | |

Roughly speaking, in a closed syllable it is VISIBLE to stress, whereas in an open syllable it is INVISIBLE. Syllables ending in word-final /ʔ/ do not count as closed and ghost vowels preceding such /ʔ/ are invisible. Compare, for instance, the pre-antipenultimate stress in *tʌkErikEʔ* where none of the ghosts contributes to the metrical structure, and the penultimate stress in *skáhkEts* where the ghost vowel is included in the foot. If it were not, *i*-prothesis would apply (in order to satisfy the above-mentioned prosodic minimality condition) and stress would fall on the prothetic vowel. (An example where prothesis does actually take place in Mohawk is given in (4a) below.)

1.3 ADJACENT GHOSTS

Sequences of ghost vowels complicate the situation. Ghost vowels in non-neighbouring syllables are invisible for stress (cf. the pre-antipenultimate stress in (4a). On the other hand, in a sequence of two contiguous syllables with ghost vowels one is visible and the other is invisible for stress (cf. the antipenultimate stress in (4b).

- (4) a. GHOSTS IN NON-ADJECENT SYLLABLES:
- | | | |
|-------------|-----------|------------------|
| /t-n-ehr-ʔ/ | ítEnehrEʔ | ‘you and I want’ |
| /o-nraht-ʔ/ | ónErahtEʔ | ‘leaf’ |
- b. GHOSTS IN ADJACENT SYLLABLES:
- | | | |
|-------------------|------------------|---------------------|
| /te-wak-ahsutr-ʔ/ | tewakahsú:tErEʔ | ‘I have spliced it’ |
| | *tewakáhsutErEʔ | |
| | *tewakahsutÉ:rEʔ | |
| /yo-t-r-ʔ/ | yó:tErEʔ | ‘it’s in the dish’ |
| | *í:yotErEʔ | |
| | *yotÉ:rEʔ | |

In the first example in (4a) *i*-prothesis takes place. This indicates that none of the non-adjacent ghost vowels are visible to stress and may contribute to the minimally required word size. On the other hand, no prothesis takes place in the last example in (4b), which shows that the prosodic minimality condition is fulfilled by one of the adjacent ghost vowels.

Visible and invisible ghost vowels have a different effect on tonic lengthening in the preceding syllable. Tonic lengthening is blocked before invisible ghost vowels in the following syllable, again as in *ítEnehrEʔ* in (4a). On the other hand, visible ghosts, just like contentful vowels, do allow for tonic lengthening in the preceding syllable, for instance, as in *yó:tErEʔ* in (4b).

1.4 EVIDENCE FOR THE NON-UNDERLYING STATUS OF GHOSTS

To recapitulate, evidence for the non-underlying status of the ghost vowel involves vowel-zero alternations (not always available) and invisibility to stress and the prosodic minimality condition just mentioned. The examples in (5) illustrate:

- (a) the alternation between an audible visible vowel and phonetic zero within the root /ahkt/ ‘go back’,
- (b) the alternation within the root /ahsutr/ ‘splice’ where the ghost vowel alternates between stress-invisible and stress-visible, and,
- (c) the alternation between zero, an invisible ghost vowel and a visible one at the end of the 1-P-AGENT prefix /k/.

- (5) a. VISIBLE [E] VS. Ø
 /s-k-ahkt-s/ skáhkEts 'I got back'
 /s-ho-ahkt-u/ shóhktu 'he went back'
- b. VISIBLE [E] VS. INVISIBLE [E]
 /te-k-ahsutr-ha?/ tkahsutÉrha? 'I am splicing'
 /te-^k-ahsutr-^?/ t^kahsútEr^? 'I shall splice'
- c. VISIBLE [E] VS. INVISIBLE [E] VS. Ø
 /k-r-ha?/ kÉrha? 'I fill it in'
 /^k-r-^?/ ^kEr^? 'I will put it into a container'
 /k-hninu-s/ khní:nus 'I buy'

2 ANALYSIS

2.1 MOHAWK AS A STRICT CV LANGUAGE

For the analysis, I adopt a version of Government Phonology known as the Strict CV approach or the CVCV model (cf. Lowenstamm 1996, as well as Larsen 1995 and Scheer 1998). In this model, every consonant is followed in phonological representation by a nucleus, either filled by a lexically contentful vowel or empty. The distribution of empty nuclei is subject to strict well-formedness constraints. Vowel epenthesis does not involve the insertion of a position. It involves providing an existing empty nucleus with phonetic interpretation when it is required by well-formedness constraints.

I interpret Mohawk ghost vowels as empty nuclei. Given this, consider some words and their representations in 0:

- (6) a. wákEras /wakərasə/ 'it smells'
 b. skáhkEts /səkahəkətəsə/ 'I got back'
 c. kÉrha? /kərəha?ə/ 'I fill it in'
 d. ónErahtE? /onərəhətə?ə/ 'leaf'
- where 'ə' = empty nucleus

In the output some of the empty nuclei from the representations in (6) acquire phonetic interpretation, while others remain inaudible.

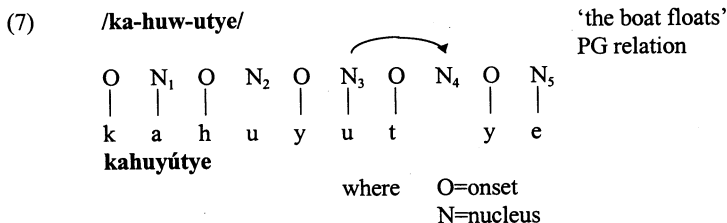
Cross-linguistically, empty nuclei are licensed to remain inaudible when preceded by specific consonants, in particular by /s/. This has been called Magic Licensing (cf. Kaye 1996). For the purposes of this paper I assume that, apart from /s/, in Mohawk also laryngeals /h, ʔ/ may have this Magic Licensing property. It affects, for instance, the final empty nucleus in (6a) and the first empty nucleus in (6b). Magic Licensing is responsible for most consonantal clusters in Mohawk. Almost all clusters of more than two consonants contain one (or more) of the Magically Licensing consonants (cf. Michelson 1988:12-3). For a more detailed discussion of the special properties of /s, h, ʔ/ in Mohawk see Rowicka (in prep.).

2.2 GHOST VOWELS AND INTERNUCLEAR RELATIONS

I argue that empty nuclei in Mohawk materialise as a result of two independent sets of phonotactic well-formedness conditions: INTERNUCLEAR ones and

INTERSEGMENTAL ones.

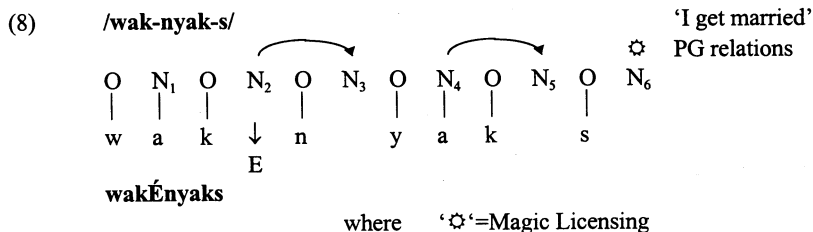
Internuclear conditions resemble well-formedness conditions on metrical feet. In rhythmic languages, sequences of unstressed syllables are banned. Analogically, sequences of empty nuclei are ruled out phonotactically. An empty nucleus requires the company of a contentful nucleus. Together they form a 'phonotactic foot'. Consider the representation of the word *kahuyútye* in (7). The simplification of the input sequence /uwu/ has been ignored as irrelevant.



The empty nucleus N₄ forms a 'phonotactic foot' with the preceding contentful nucleus N₃, N₄ being the head and N₃ being the dependent. An empty nucleus which is parsed into a foot like this is licensed to remain inaudible.

Relations between empty and contentful nuclei, called PROPER GOVERNMENT (PG), have been postulated in Government Phonology and shown to play a role in the phonology of many languages. In van der Hulst & Rowicka (1997) it has been argued that PG relations should be interpreted as lexicalised foot structure. Contrary to mainstream Government Phonology, I assume that PG feet are TROCHAIC: an empty nucleus is parsed together with the PRECEDING contentful vowel, and not the FOLLOWING (cf. Rowicka, in press).

Phonotactic lapses, i.e. sequences of empty nuclei, are resolved by providing one of the empty nuclei with phonetic interpretation. Consider the representation of the word *wakÉnyaks*.



The representation in (8) contains a few empty nuclei. The last one N₆ is magically licensed and does not participate in internuclear relations. The preceding N₅ occupies the dependent position in a PG foot the head of which is contentful N₄. In other words, it is PROPERLY GOVERNED by N₄. It is correctly predicted to remain inaudible. Empty N₂ and N₃ constitute a lapse. In order to solve the lapse, they form

together a trochaic PG foot. To be the head of this foot N_2 must acquire phonetic interpretation, as shown in (8).

The phonetic surfacing of heads in PG relations is required by the phonological Empty Category Principle (ECP) in the formulation given below (for the formulation of the ECP in standard Government Phonology, see Kaye, Lowenstamm & Vergnaud 1990: 219):

(9) THE REVISED ECP (Rowicka, in press)

An empty nucleus is phonetically realised if it properly governs another empty nucleus.

The Revised ECP boils down to a fairly uncontroversial requirement that phonological heads should be phonetically present.

Cross-linguistically, nuclei which are not heads in PG feet generally remain *INAUDIBLE*, just like N_3 , N_5 and N_6 in (8). However, the situation in Mohawk is a little different.

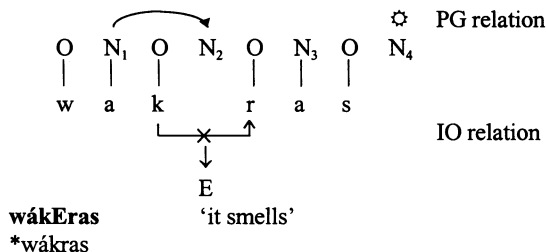
2.3 GHOSTS AND INTERSEGMENTAL RELATIONS

Proper Governing relations are formed between nuclei by and large irrespective of the nature of the surrounding consonants. By licensing the intervening empty nucleus to remain inaudible, they should be able to produce any type of consonantal clusters.

However, in Mohawk certain clusters are not tolerated at all. Clusters of a consonant and a sonorant or a glottal stop are always broken up by a ghost vowel. This indicates that not only internuclear relations are relevant in the determining the audibility of empty nuclei, but interconsonantal relations play a significant role as well. The existence of relations between consonants in separate onsets across an empty nucleus has also been recognised in Government Phonology. Two consonants separated only by an empty inaudible nucleus, and hence melodically adjacent, contract a relation, called *INTERONSET GOVERNMENT* (see, e.g. Guerssel & Lowenstamm 1988, Kaye 1990, Cyran 1996 as well as Cyran & Gussmann, in press). Sonorants can generally only be dependents in such relations. The headedness in Interonset (IO) relations is set parametrically in every language. It expresses the language's preference for 'branching onset-like' clusters or 'codan-onset-like' clusters.

In Mohawk 'branching onset-like' clusters are ruled out. They are always broken up by a ghost vowel. This indicates that no sonorant can be the dependent of the preceding consonant in an left-headed IO relation in Mohawk. Consider the representation in (10).

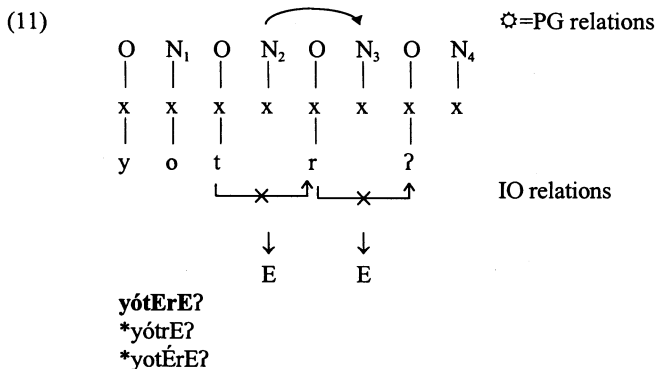
(10) NO LEFT-HEADED INTERONSET GOVERNMENT IN MOHAWK



The crossed out arrow in (10) indicates that there is no left-headed IO Government in Mohawk. That is why the melodic sequence /kr/ must be broken up by an audible nucleus.

One may conclude that IO in Mohawk is right-headed, rather than left-headed. Michelson (1988) does indeed report the occurrence of /rk/ clusters, but consonantal sequences do not generally exhibit falling sonority. I suggest that IO functions in Mohawk only in the form of NEGATIVE CONDITIONS on melodic sequencing.

The glottal stop patterns in behaviour with sonorants. Clusters of a consonant and /ʔ/ are impossible, not only in word-final position. Consider the word *yó:tErEʔ* 'it's in the dish':



The glottal stop cannot be in an IO relation with the consonant it stands next to. Therefore their melodic adjacency has to be broken up by a surfacing nucleus. Notice that N_2 in (11) must surface both due to the internuclear and due to the intersegmental requirements. On the other hand, the presence of N_3 is only required by Interonset.

2.4 PROPER GOVERNMENT RELATIONS AND STRESS

Compare how stress treats ghost vowels which emerge due to Proper Government, i.e. internuclear constraints, and ghost vowels which appear due to Interonset

Government, i.e. intersegmental constraints. The nucleus N_2 which emerges in (11) as the head of a PG relation does count for stress, while N_3 does not. The analysis of other examples also confirms this generalisation.

All and only those ghost vowels which must be heads in PG relations are visible to stress. None of the empty nuclei which are dependents in PG feet count for stress, whether they are inaudible, like N_3 and N_5 in (8), or they are audible for reasons of Interonset, like N_2 in (11).

This situation is hardly surprising. Under the present analysis stress-visible nuclei are heads of internuclear relations, i.e. prosodic heads. It is to be expected that such heads are relevant for some other level of prosodic structure, while non-heads are not.

2.5 TONIC LENGTHENING

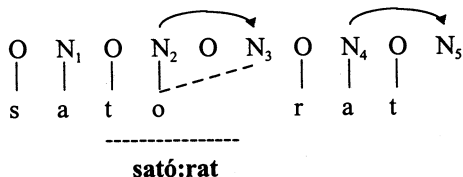
Visible and invisible ghost vowels have different effect on tonic lengthening in the preceding syllable. Stressed vowels in open syllables are generally subject to lengthening (cf. (12a)). Tonic lengthening fails to take place when a consonantal cluster follows (cf. (12b)) and before an invisible ghost vowel [e] in the next syllable (cf. (12c)). Visible ghost vowels do not block the lengthening (cf. (12d)).

- (12) a. TONIC LENGTHENING:
- | | | |
|---------------|------------|---------------------|
| /s-atorat/ | sató:rat | 'hunt'-IMPERATIVE |
| /wak-ashet-u/ | wakashé:tu | 'I have counted it' |
- b. NO TONIC LENGTHENING BEFORE A CLUSTER:
- | | | |
|----------------|------------|----------------|
| /k-atirut-ha?/ | katirútha? | 'I pull it' |
| /s-ho-ahkt-u/ | shóhktu | 'he went back' |
- c. NO TONIC LENGTHENING BEFORE AN INVISIBLE GHOST:
- | | | |
|-------------|-----------|-------------|
| /w-akra-s/ | wákEras | 'it smells' |
| /o-nraht-?/ | ónErahtE? | 'leaf' |
- d. TONIC LENGTHENING BEFORE A VISIBLE GHOST:
- | | | |
|-------------------|-----------------|---------------------|
| /te-wak-ahsutr-?/ | tewakahsú:tErE? | 'I have spliced it' |
| /yo-t-r-?/ | yó:tErE? | 'it's in the dish' |

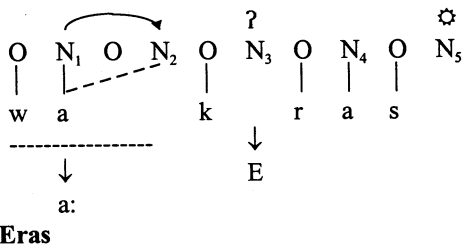
Under the present analysis, both contexts where tonic lengthening is blocked involve stress-invisible empty nuclei, either audible or inaudible. The blocking of tonic lengthening also follows from the present analysis. (It is accounted for analogically to the long vowel shortening in Turkish and Yawelmani in Rowicka, in press.)

Within the Strict CV approach, tonic lengthening is viewed as addition of an empty syllable, i.e. an empty Onset + Nucleus sequence, to the vowel under main stress (cf. Larsen 1994 on Italian). The tonic vowel (underlined below) spreads its melody to the following empty nucleus so that a long vowel results. As argued in Rowicka (in press), there must be a PG relation between the vowel spreading its melody (N_2 below) and the empty target (N_3 below).

'hunt'-IMPERATIVE



(14)



3 CONCLUSION

In this presentation I have discussed the behaviour of the ghost vowel *e* in Mohawk. Such vowels, even when audible, i.e. phonetically realised, are not always visible with respect to stress. I have analysed ghost vowels as underlying empty nuclei. I have argued that they acquire phonetic interpretation in the output as a result of two independent sets of phonotactic well-formedness conditions: internuclear ones, involving Proper Government, and intersegmental ones, involving Interonset Government. Only the Proper Government relations result in the emergence of nuclei which are both phonetically audible and prosodically visible. The requirements of Interonset Government can only trigger the phonetic presence of some nuclei. The present analysis also straightforwardly accounts for the blocking effect of invisible ghost vowels on tonic lengthening.

ACKNOWLEDGEMENTS

I am grateful to Monik Charette, Eugeniusz Cyran, Colin Ewen, Harry van der Hulst, Jonathan Kaye, Krisztina Polgárdi, Jeroen van de Weijer, and the audience of BLS 24 for helpful comments and stimulating criticism. The usual disclaimers apply.

REFERENCES

- ALDERETE, J. 1995. Faithfulness to prosodic heads. Ms, University of Massachusetts, Amherst.
- BEATTY, J. 1974. Mohawk morphology. PhD diss., City University of New York (1972). Published as Occasional Publications in Anthropology Linguistics Series 2. Museum of Anthropology, University of Northern Colorado, Greeley, CO.
- BONVILLAIN, N. 1973. A grammar of Akwesasne Mohawk. Mercury Series, Ethnology Division Paper No. 8. National Museum of Man, Ottawa.
- CHAFE, W. 1977. Accent and related phenomena in the Five Nations Iroquois languages. In L. Hyman (ed.) *Studies in stress and accent*. Southern California Occasional Papers in Linguistics 4. Department of Linguistics, University of Southern California, Los Angeles. 169-81.
- CYRAN, E. 1996. Licensing properties of nuclei and principle ranking in Irish. *The Linguistic Review* 13. 1-13.
- CYRAN, E. & E. GUSSMANN. in press. Consonantal clusters and governing relations: Polish initial consonant sequences. In H.G. van der Hulst & N. Ritter (eds.). *The syllable: views and facts*. Oxford University Press, Oxford.
- GUERSSEL, M. & J. LOWENSTAMM. 1988. Classical Arabic 'metathesis'. Paper presented at the Generative Phonology Workshop 'How phonology is different', University of Leiden.
- GUSSMANN, E. & J. KAYE. 1993. Polish notes from a Dubrovnik café. I: The yers. *SOAS Working Papers in Linguistics and Phonetics* 3. 427-62.
- HAGSTROM, P. 1997. Contextual metrical invisibility. *MIT Working Papers in Linguistics*, vol. 30.
- HULST, H. VAN DER & G.J. ROWICKA. 1997. On some parallels between (un)stressed vowels and (un)realised empty nuclei. In G. Booij & J.M. van de Weijer (eds.) *Phonology in progress - progress in phonology*. Holland Academic Graphics, the Hague. 125-49.
- KAYE, J. 1990. Coda Licensing. *Phonology* 7, 301-30.
- KAYE, J. 1996. Do you believe in magic? The story of *s*+*C* sequences. In H. Kardela & B. Szymanek (eds.) *A festschrift for Edmund Gussmann*. 155-76. The University Press of the Catholic University of Lublin.
- KAYE, J.D., J. LOWENSTAMM & J.-R. VERGNAUD. 1990. Constituent structure and government in phonology. *Phonology* 7, 193-231.
- LARSEN, U.B. 1995. Vowel length, raddoppiamento sintattico and the selection of the definite article in Modern Italian. Ms, Université Paris 7.

- LOWENSTAMM, J. 1996. CV as the only syllable type. In Jacques Durand & Bernard Laks (eds.) *Current Trends in Phonology*. Oxford University Press.
- MICHELSON, K. 1981. Stress, epenthesis and syllable structure in Mohawk. In G.N. Clements (ed.) *Harvard studies in phonology*. Vol. II. Indiana University Linguistics Club. 311-53.
- MICHELSON, K. 1983. A comparative study of accent in the Five Nations Iroquoian languages. PhD diss., Harvard University.
- MICHELSON, K. 1988. A comparative study of Lake-Iroquoian accent. Kluwer, Dordrecht.
- MICHELSON, K. 1989. Invisibility: Vowels without a timing slot in Mohawk. In D. Gerdts & K. Michelson (eds.) *Theoretical perspectives on Native American languages*. Suny Press, Albany, New York, 38-69.
- MITHUN, M. 1979. Iroquoian. In L. Campbell and M. Mithun (eds.) *The languages of native America*. University of Texas Press, Austin. 133-212.
- PIGGOTT, G. 1995. Epenthesis and syllable weight. *Natural Language and Linguistic Theory* 13.2. 283-326.
- POSTAL, P. 1968. *Aspects of phonological theory*. Harper & Row, New York.
- POTTER, B. 1994. Serial optimality in Mohawk prosody. In K. Beats, J. Denton et al. (eds.). *Papers from the 30th Regional Meeting of the Chicago Linguistic Society*, vol. 1. 347-61.
- ROWICKA, G.J. in press. On trochaic Proper Government. In J.J. Rennison (ed.) *Syllables!?*. Holland Academic Graphics, The Hague.
- ROWICKA, G.J. in prep. On ghost vowels. PhD dissertation, HIL/University of Leiden.
- SCHEER, Tobias. 1998. A unified model of Proper Government. *The Linguistic Review*.

A New Look at Japanese Conversational Styles: (In-)Direct Speech and Turn Management

Shie Sato
University of California, Los Angeles
Shigeko Okamoto
California State University, Fresno

1. Introduction*

Japanese communication style is typically described as highly harmonious, cooperative, and empathetic. These characteristics are said to be particularly evident in frequent uses of indirect expressions and supportive turn-management devices, such as backchannels (or *aizuchi*) and other reactive tokens (Mizutani 1988; Matsuda 1988; Iwasaki 1997; Maynard 1997). This characterization concurs with the cultural ideology which emphasizes relational interdependence rather than individualism (Doi 1971; Lebra 1976; Markus and Kitayama 1991; Wierzbicka 1991). However, empirical support for this view seems to be inadequate due to the narrow sampling of conversational situations.

This study reexamines the use of "cooperative" interactional styles in Japanese conversations. The data consist of six audio-taped dyadic conversations carried out in two different social contexts--three between family members, all involving mother-daughter relations, and three between female close friends. The speakers recorded a conversation, sitting face to face; they chose the topics of conversations as they wished. The participants in each conversation are shown in Table 1 below:

Table 1: Participants in the six audio-taped conversations

Conversation A: mother (47 years old) and her daughter (17 years old)

Conversation B: mother (52 years old) and her daughter (22 years old)

Conversation C: mother (56 years old) and her daughter (23 years old)

Conversation D: two female close friends (Both are 23 years old; one of them is the same person as the daughter in conversation C.)

Conversation E: two female close friends (Both are 22 years old, one of them is the same person as the daughter in conversation B.)

Conversation F: two female close friends (Both are 17 years old; one of them is the same person as the daughter in conversation A.)

The participants in conversations B1 and B2 speak standard Japanese. The other participants speak regional dialects. The lengths of the conversations vary from 20 minutes to one hour. All conversations were transcribed, using the transcription conventions shown in Table 2 below.

Table 2: Transcription conventions

{ }	overlapping talk	?	rising intonation
.	falling/utterance-final intonation	,	continuing intonation
(0.0)	length of pause/silence in tenths of a second	(.)	micro-pause
()	non-linguistic activity or change in pitch/volume		

We examined the "cooperativeness" of the participants' interactional styles in these six conversations by analyzing the following features: (1) expressions of conflicts, (2) backchannels, (3) types of response remarks, in particular, minimal responses, and (4) silences and pauses.

2. Expressions of Conflicts

It is commonly said that due to the cultural concern for harmony and mutual dependence, Japanese tend not to assert themselves in conflict situations; that is, they avoid confrontation and the use of direct expressions of conflict (Kindaichi 1975; Lebra 1976; Watanabe 1993). Jones (1990) points out that expressions of conflict are not uncommon in Japanese conversations, but at the same time she reports that the expressions of conflict in the conversations she examined were all accompanied by some mitigating devices. However, in our data, in particular, in the family conversations, conflicts were often directly expressed with no mitigating devices.

We have analyzed all the expressions in our data that involved some kind of conflict, such as disagreement, defiance, and criticism. For each conflict-related expression, we examined whether it was accompanied by any of the conflict-mitigating devices listed in Table 3 below (For a variety of linguistic devices for mitigating direct conflicts in Japanese, English, and other languages, see Pomerantz 1984; Schegloff et al. 1977; Brown et al. 1978; Levinson 1983; Lerner 1987; Mizutani et al. 1987).

Table 3: Conflict-mitigating devices

- 1. prefaces before opposition, such as acknowledgment and self-disparagement
- 2. follow-ups after opposition, such as explanation and apology
- 3. delayed start with fillers, hesitational pause, and vowel lengthening
- 4. topic shift
- 5. style switch
- 6. clause-final qualifiers, such as *kedo* 'but' and *kamo* 'maybe'
- 7. laughter

An expression of conflict was considered direct, if it did not include any of the devices listed in Table 3; it was considered indirect, if it included one or more of these devices. Table 4 shows the number of direct and indirect expressions in the first 15-minute segment of each conversation:

Table 4: Number of direct and indirect expressions of conflicts in a 15-minute segment of each conversation

Conversation	direct	indirect
A	10	6
B	28	10
C	40	1

D	1	6
E	0	1
F	0	2

As can be seen, compared to the participants in the conversations between friends, those in the family conversations not only expressed conflicts more frequently, but they expressed them directly. Direct expressions of conflicts are illustrated in (1) and (2):

- Example 1. <From conversation B; Daughter is asking her mother to buy her an expensive bag.>
- 1 Mother: *Shaneru no baggu ja nakereba donna baggu ga ii no?*
'If it's not a Chanel's bag, what kind of bag would you like?'
 - 2 Daughter: *Ya da. Shaneru no ja nai to.*
'No. It has to be Chanel.'
 - 3 Mother: *Tatoeba no hanashi.*
'I'm talking hypothetically.'
 - 4 Daughter: *Nai.*
'(There's) none (that I like).'
 - 5 Mother: *Demo, Shaneru takai n da yo.*
'But Chanel's (bags) are expensive.'
 - 6 Daughter: [Raising her voice] *Takai kara kaitai n jan.*
'I want to buy it because it's expensive (don't you understand?)'
- Example 2. <from conversation C>
- 1 Daughter: *Kaasan Amerika (.) no hondo ni kitara (.) koto nai yo ne. Nai yo ne. Hawa{ i shika.*
'You haven't been to the mainland of America, have you? Only Hawaii, right?'
 - 2 Mother: *{Aru wake nai desho sonna baka na koto yuutara tsumaran yo. Sonnaa,*
'Of course, not. Don't say such a stupid thing. It's nonsense. Such ...'

In (1), the daughter expresses her disagreement with her mother bluntly in lines 2, 4, and 6 without any hesitations; she did not use any mitigating devices listed in Table (3). In Example (2), the mother makes a meta-communicative comment, directly criticizing her daughter's question as stupid.

These results show that speakers can be quite direct and assertive. At the same time, they indicate that speakers do not choose direct expressions randomly. Rather, they are concerned about when such expressions may or may not be used. Frequent uses of direct expressions in family conversations may index the nature of relationship as being very intimate. But if they are used in other contexts, they may convey different meanings, such as rudeness and a change in the friendship. This indicates that the social import of direct expressions is not always the same; rather they are relative to specific social situations.

3. Backchannels

We now turn to the second feature, backchannels. Compared to speakers of languages like English, Japanese are said to use backchannels much more frequently to express interactional support and involvement in the conversation (Mizutani 1988; Maynard 1986, 1997; Matsuda 1988). This, then, is considered a manifestation of the Japanese cultural concern for harmony and co-operation

(Matsuda 1988).¹

In our analysis, we examined the total number of backchannels in the first 15 minutes of each conversation. In reference to Maynard (1989) and Clancy et al. (1996), we identified reactive tokens with the following features as backchannels: (1) continuer, (2) display of understanding of content, (3) support toward the speaker's judgment, and (4) agreement. The results of our analysis are shown in Figure 1:

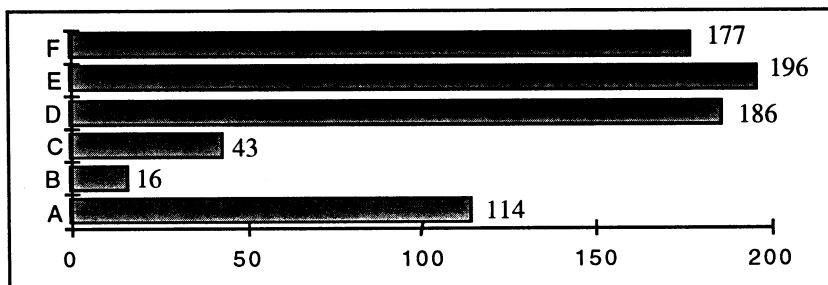


Figure 1. Total number of backchannels in a 15-minute segment of conversations A-F.

As can be seen, there is wide variability in the use of backchannels among the six conversations, ranging from 16 tokens to 196 tokens. This highest number (i.e. 13 per minute) is close to, but still below, the number of back-channels that are said to be normally used in Japanese conversations.² Such frequent uses of backchannels were seen in the conversations between friends, but not in the family conversations. The latter, particularly conversations B and C, show that the use of backchannels can be minimized. The absence of backchannels is illustrated in (3) and (4):

Example 3. <from conversation C>

- 1 Mother: *Isshoni sumu so jattara ima no hiree toko demo ee wa ne, hanbun dasasete.*
'If you guys are going to live together, the spacious apartment you have now will be fine, (you can) have her pay the half.'
- 2 Daughter: *Iya, yappari ne, kaasan.*
'No, after all, mother'
- 3 (0.2)
- 4 Daughter: *chuu-beddo-ruumu?*
'Two bedroom?'
- 5 (0.2)
- 6 Daughter: *ga nai to yappa dame.*
'without it, it wouldn't work.'
- 7 (0.6)
- 8 Daughter: *Kaasan.*
'Mother.'
- 9 (1.0)

Example 4. <from conversation B>

- 1 Mother: *Dame da yon (na).*
'You shouldn't do that.'
- 2 Daughter: *De son toki ni, soo yuu huu ni wa, hanashiteta noni, Akira nan ka, saa tanomu n dattara tanon demo ii kedo isogashii kara na toka itte. De kekkyoku sa konaida sa denwa kakatte kita toki ni, atashi ga Itoo san ni tanomoo to omotte feragamo toka, sugoi sa, hoshikatt, mukoo de datte yasui tte kiita kara toka ittara, nande tanomeba yokatta noni ku-gatsu da yo itta no toka itte. Ku-gatsu ja zenzen mae jan toka itte.*
'At that time, we were talking like that, but Akira said, "if you want to ask him (to buy something), it's O.K. to ask, but he is busy". Then, after all, when he called me the other day, I told him "I wanted to ask you to get Feragamo for me, because I heard that it's cheap there". Then he said, "why didn't you ask me? I went there in September." So I said, "September was a long time ago".
- 3 Mother: *Jooshi ni wa tanomenai.*
'(You) can't ask your supervisor.'

In Example (3) the mother did not use backchannels where they could have been used (i.e. lines 3, 5, 7, and 10); instead, she remained silent. In Example (4) during the daughter's long utterance in line 2, the mother did not use any backchannels. (See also (9) in section 5.)

These infrequent uses of backchannels seen in our data cannot be treated as simple exceptions to Japanese communication styles, because their distribution is not random, and because they are functionally important. That is, they can index the speaker's feelings toward the addressee or the situation, such as intimacy, uninhibited atmosphere, or the lack of interest in the current topic. They may also be used to imply anger toward the addressee.

4. Minimal Responses

The third feature we analyzed was the types of responses to the other speakers' utterances, in particular, minimal responses. We distinguish minimal responses from backchannels. A minimal response is a brief response to the other's speech act that expects some kind of specific response--e.g. the answer to an information-seeking question, the acceptance or decline of an offer or request, etc.; it functions as the second unit in an "Adjacency Pair" (Sacks et al. 1974). From the viewpoint of the Gricean Maxims of conversation, minimal responses provide sufficient information and hence may be regarded as cooperative communicative behaviors. However, if they are used consecutively, they may convey that the speaker is uninterested or being uncooperative. In particular, if minimal responses are used persistently by only one of the participants, asymmetry in the cooperativeness between the participants becomes evident. Such asymmetry is illustrated in Example (5):

Example 5. <from conversation C>

- 1 Daughter: *Aryaa, doko de hataraki-yocchatta n ka ne Kayo-chan.*
'I wonder where Kayo worked.'
- 2 (2.8)
- 3 Mother: *Shiran.*
'I don't know.'
- 4 (4.0)
- 5 Daughter: *Aa, ashita tanoshimi ya ne, kaasan.*
'Well, it'll be fun tomorrow. Don't you think so, mother?'
- 6 (0.5)
- 7 Daughter: *Ne.*
'Right?'
- 8 Daughter: *Anmari?*
'Not that much?'
- 9 Mother: *Han han.*
'Fifty-fifty. (Maybe, maybe not.)'
- 10 Daughter: *Nan de? Chotto erai? (.) A, ki ga haren te ii yotta ne.*
'Why? Tiring? Oh, you said you were not in good spirits.'
- 11 Mother: *Un.*
'Yeah.'
- 12 (0.6)
- 13 Daughter: *Ee wa ne, tama niyaa. Kaasan mo kii o, ki o (.) ano-,*
'It's OK. Once in a while, you also should . . .'
- 14 Mother: (Yawning)
- 15 Daughter: *Nan chuu n ya ro ka.*
'What can I say?'
- 16 (0.8)
- 17 Daughter: *Tama ni wa are sen nya.*
'You should, sometimes . . .'
- 18 (0.8)
- 19 Daughter: *A, atashi ka-yoobi Mori-san chi ni tomaru ke.*
'Oh, I'll sleep over at Miss. Mori's place on Tuesday.'
- 20 (1.2)
- 21 Mother: *Doozo.*
'Go ahead.'

Here the mother gives minimal responses in lines 3, 9, 11, and 21. Note also that her minimal response in line 9 comes only after the daughter's repeated attempts to solicit her response. Moreover, the mother yawns in the middle of the daughter's talk (line 14). A similar example is given in (6):

Example 6. <from conversation B>

- 1 Mother: *{Aa soi ja jikan-teki ni wa*
kankei nai n da.
'Oh, the time doesn't matter, then.'
- 2 Daughter: *N.*
'Yeah.'
- 3 Mother: *Jibun no ugoki de. Mariko ga denwa sun no?*
'You have a flexible schedule. Are you supposed to call her,

- 4 Daughter: Mariko?
Un.
'Yeah.'
- 5 (0.3)
- 6 Mother: *Kaette kita yo te?*
'(You'll tell her) "I'm back now", right?'
- 7 Daughter: *Soo.*
'Right.'

As we can see, the daughter's responses in (6) are minimal. Because she does not offer elaborated and helpful responses, the mother keeps asking further questions.

Persistent uses of minimal responses in (5) and (6) seems to indicate the speakers' lack of interest in actively participating in the conversation. These examples illustrate the interactional style in which one is not so concerned about showing one's involvement in the conversation or empathy for the other.

5. Pauses and Silences

Pauses and silences contribute to discourse structuring by indicating topic shifts, '(complex) transition-relevance places' (Sacks et al. 1974; Clancy et al 1996), etc. They can also be used to convey social meanings such as the speaker's attitudes and feelings towards the relationships, topics, and conversational goals (Tannen et al. 1985). We examined the use of silences and pauses in the two kinds of conversations in our data. Figure 2 shows the total amount of pauses and silences in a 15-minute segment of each of the six conversations:

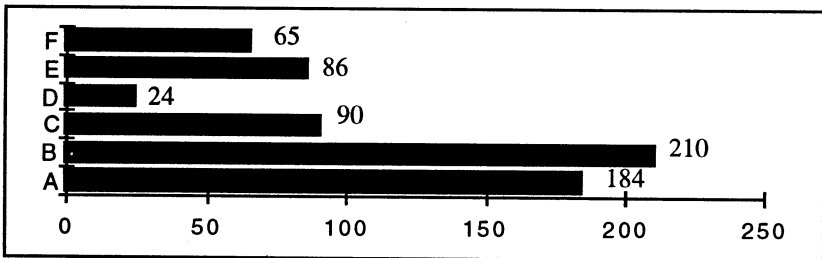


Figure 2: Total amount of pauses/silences in a 15-minute segment of conversations A-F (measured in seconds).

As we can see, the family conversations in general had a much greater amount of silences and pauses than the conversations between friends. In particular, conversations A and B contrasted sharply with conversation D, in which both speakers talked continuously at a fast pace.

There are also qualitative differences between pauses and silences in the family conversations and those in the conversations between friends. For example, in the family conversations, silence often occurred immediately after the first unit of an Ajacency Pair, such as an information-seeking question, an offer, and a request. This is illustrated in (7):

Example 7. <from conversation A>

- 1 Mother: *Huun. Hoka ni ano maikurohon ja nakute, ano, denchi kaini iku wake?*
'I see. Besides that, are you going to go buy a microphone, no, uh ... batteries?'
- 2 (1.2)
- 3 Daughter: *Denchi wa ee yan.*
'I don't need batteries.'
- 4 (1.0)
- 5 Mother: *Ja nani kai ni iku no?*
'What are you going to buy, then?'
- 6 (2.0)
- 7 Daughter: *Maiku.*
'A microphone.'
- 8 (2.0)
- 9 Mother: *Watashi no ja nakute, Keiko no maiku?*
'Not for me, but for you?'
- 10 (0.8)
- 11 Daughter: *Iya,*
'No.'
- 12 (1.0)
- 13 Mother: *Keiko no konsaato yoo ni maiku kau?*
'Are you buying a microphone for your concert?'
- 14 Daughter: *Ett, Keiko ga kau no ka.*
'What? Do I have to buy it (myself)?'
- 15 Mother: *Itt, chigau. Watashi no koto itteru no ne?*
'No, you are talking about me, aren't you?'
- 16 Daughter: *Soo. {Soo, soo.*
'Yeah, yeah, yeah.'
- 17 Mother: *{Hun. Hun. Wakatta. Wakatta. Ja ashita no ano-Hiroshima-iki no koto oshiete yo.*
'I see. I see. Well, tell me about your visit to Hiroshima tomorrow.'
- 18 (2.0)
- 19 Mother: *Ashita tomaru no wa Hukuyama?*
'Are you staying over night in Fukuyama tomorrow?'
- 20 Daughter: *Soo.*
'Yeah.'
- 21 (2.2)
- 22 Mother: *Ashita no ban ya ro?*
'That's tomorrow night, right?'

In (7) the daughter gives delayed responses (lines 3, 7, and 11) to her mother's information-seeking questions. Moreover, she gives no response to the mother's question in line 17. The mother, therefore, asks another question in line 19.

Further, silence may also follow the second unit of an Adjacency Pair. For example, the daughter's responses in (7) are minimal and followed by silence (lines 4, 8, 12, and 21) rather than by elaboration of the answers. Example (8) also illustrates the same kind of silence after a minimal response:

Example 8. <from conversation B>

- 1 Mother: *Ashita tomodachi doo sun nan ka.*
'How are you going to see your friends tomorrow?'
- 2 Daughter: *Mukae ni kite kureru.*
'They are coming to pick me up.'
- 3 (6.0)
- 4 Mother: *Okashi sukoshi motteku?*
'Do you want to take some snacks with you?'
- 5 Daughter: *Iranai.*
'(No) I don't.'
- 6 (2.1)
- 7 Mother: *Ano-, kuruma nara sa,*
'Uh, if you're going by car.'
- 8 (0.2)
- 9 Mother: *mottette yo. Ano, mikan toka.*
'Take them. Uh, oranges and so on.'

Silences may also occur as absences of backchannels. This is illustrated in (9) and (3):

Example 9. <from conversation C>

- 1 Daughter: *Moo, kaasan kono mae ne-, (.) Kaasan.*
Listen, mother. The other day, mother.'
- 2 Mother: *Un.*
'Yeah.'
- 3 Daughter: *Reeboo ga nai hoccha Karen no kuruma? Moo ne, tochuu kara ne, atsuu natte kara ne, moo ne, otagai mukuchi naru n yo ne moo. Atsuute, Atsuute ne--.*
'Karen doesn't have an air conditioner in her car, as you know. So, after a while, it became so hot that we didn't (want to) talk. It was too hot, you know?'
- 4 (0.9)
- 5 Daughter: *Moo shinisoo yatta.*
'We felt like we were going to die.'
- 6 (1.0)
- 7 Daughter: *Kaasan.*
'Mother.'
- 8 (2.0)
- 9 Daughter: *(De)mo chikai mon chau, nan ka Kokura Higashi nante sugu jaa ne, ano, koosoku de?*
'But, it won't take much time for us to go to Kokura Higashi if we take the freeway.'
- 10 (0.2)
- 11 Daughter: *Nee kaasan.*
'Right, mother?'
- 12 Mother: *Aa-. (Yawn)*
'Well ...'

In lines 4 and 6 of (9) the mother could have produced backchannels, responding to the daughter's narrative; but she is silent. So the daughter calls for her attention explicitly in line 7, which still receives no response (line 8). The daughter then introduces another topic in line 9. The mother gives no response, and finally yawns in line 12, when the daughter calls for her attention again.

Thus silences in these examples occurred as delayed responses, as lack of elaborations, and as absences of backchannels. They show asymmetry in the participants' contribution to the conversation. Through the use of such silences the speakers seem to convey that they are not so interested in actively participating in developing the current topic or the conversation itself. Such uses of silences were frequent in the family conversations but not in the conversations between friends. By this, we do not mean to say that silences and pauses always occur more frequently in family conversations than in other kinds of conversations. They may occur in the latter as well. But the point is that the meanings of silences and pauses can differ depending on the nature of social contexts.³

6. Conclusion

Our analysis demonstrates wide variability in speech styles, ranging from highly "cooperative" styles to less "cooperative" styles. While the two kinds of relations in the data (i.e. friends and family) both concern intimacy, their conversational styles were often quite different from each other. Generally speaking, compared to the conversations between friends, those between family members were much less "cooperative," characterized by direct expressions of conflict, fewer backchannels, frequent occurrences of silences and pauses, sequences of minimal responses, etc. The contrast between the two kinds of conversations that we have seen here suggests that speakers use different strategies or norms for different kinds of situations. Thus, as our examples from the family conversations show, less "cooperative" styles do not necessarily mean rudeness or a change in the relationship, since the norms are different from those used in other social contexts. Rather, they can convey that the relationship is very intimate so that the speakers can be spontaneous and express themselves more freely and directly. Note, however, if their styles become unusually uncooperative, they will convey different meanings, such as anger and a breakdown in the relationship.

Although more conversations in diverse social contexts need to be examined, the results of this study suggest that it is not accurate to characterize Japanese communication styles as highly "cooperative" and "empathetic" across all contexts. Rather, they indicate that the use and interpretation of speech styles are relative to specific social situations, and that less "cooperative" styles are not mere exceptions to Japanese conversational styles, but rather have important social functions. That is, they can index and also define the communicative context, including the nature of relationships (power, degrees of intimacy, etc.) and the speakers' attitudes toward topics and conversational goals. We argue that variations in the use of highly "cooperative" styles are best regarded as speakers' strategies for effectuating what they consider to be situationally most appropriate. The broad cultural ideology of harmony may influence the speakers' strategies, but it is not uniformly applied to all contexts. An adequate account of Japanese communication styles must address the question of under what circumstances the "dominant" norms are or are not observed. As pointed out recently (Miller 1989; Okamoto 1995, 1997), many previous studies of sociolinguistic phenomena in

Japanese, such as honorifics and "gendered" speech styles, have tended to focus on stereotypical usage based on the broad cultural and linguistic ideology, while paying little attention to "deviant" uses. In the present paper, we have examined yet another sociolinguistic phenomenon from this critical perspective.

Finally, this study also demonstrates how different ways of manipulating and coordinating a variety of linguistic features contribute to creating different interactional styles. Many conversational analyses tend to examine interactional devices such as backchannels, turn-taking, and silence/pauses individually, focusing on their local form-function units. But in order to fully understand diverse speech styles as interactional strategies, one must also investigate the meanings that emerge from the ways a set of linguistic devices are managed globally throughout conversations.

Notes

*This study was partly supported by a California State University research grant. We are grateful to those who helped us with the data collection, in particular, Takae Izumi, Misako Kure, Hatsue Sato and those who participated in recording the conversations. We would also like to thank Asif Agha, Susan Ervin-Tripp, Ritva Laury, Yoshiko Matsumoto, Elinor Ochs, Meryl Siegal, Raymond Weitzman, and Anna Wierzbicka for their valuable comments and discussions.

¹Comparing the use of Reactive Tokens (including backchannels) in English, Japanese, and Mandarin conversations, Clancy et al. (1996) found an interesting difference between English and Japanese conversations. That is, in contrast to Americans, Japanese speakers were much more likely to "give their Reactive Tokens while the primary speaker is 'in progress' rather than waiting for a completion point" (380). They suggest that "Reactive Tokens which occur frequently and are distributed throughout another speaker's turns and clauses rather than at possible completion points may constitute an especially appropriate means of providing and receiving interactional support" in Japanese conversation (ibid.:381).

²Mizutani (1988), for example, reports that in her study the average number of backchannels (*aizuchi*) was 15-20 per minute.

³Paralinguistic features such as laughter, pace of talk, tone and quality of voice are also important resources for creating different interactional styles. Although a detailed analysis needs to be done, our data indicate that compared to highly "cooperative" styles, less "cooperative" styles are characterized by fewer instances of laughter, slower pace of talk, and lower tone of voice. They also included fewer syntactic co-constructions and overlaps.

References

- Brown, Penelope and Stephen C. Levinson. 1978. *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.
- Clancy, Patricia M., Sandra A. Thompson, Ryoko Suzuki, and Hongyin Tao. 1996. The conversational use of reactive tokens in English, Japanese, and Mandarin. *Journal of Pragmatics* 26.355-387.
- Doi, Takeo. 1971. *Amae no koozoo (The anatomy of dependence)*. Tokyo: Kobundo.

- Iwasaki, Shoichi. 1997. The Northridge earthquake conversation: The floor structure and the "loop" sequence in Japanese conversation. *Journal of Pragmatics* 28.661-693.
- Jones, Kimberly. 1990. Conflict in Japanese Conversation. Doctoral dissertation. University of Michigan.
- Kindaichi, Haruhiko. 1975. *Nihon-jin no gengo-hyoogen* (Japanese verbal expressions). Tokyo: Koodansha.
- Lebra, Takie S. 1976. Japanese patterns of behavior. Honolulu: University of Hawaii Press.
- Lerner, Gene H. 1987. Collaborative turn sequences: Sentence construction and social action. Unpublished Ph.D. dissertation. University of California, Irvine.
- Levinson, Stephen. 1983. *Pragmatics*. Cambridge: Cambridge University Press.
- Markus, Hazel R. and Shinobu Kitayama. 1991. Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review* 98(2).224-253.
- Matsuda, Yoko. 1988. *Taiwa no Nihon-go kyooiku-gaku: Aizuchi ni kanren shite* (Teaching of Japanese dialogues: With reference to backchannels). *Nihongo-gaku* (Japanese Linguistics) 7.59-66.
- Maynard, Senko. 1986. On backchannel behavior in Japanese and English casual conversation. *Linguistics* 24.1079-1108.
- _____. 1989. Japanese conversation: Self-contextualization through structure and interactional management. Norwood, NJ: Ablex.
- _____. 1997. Japanese communication: Language and thought in context. Honolulu: University of Hawaii Press.
- Mizutani, Nobuko. 1988. Aizuchi-ron (On backchannels). *Nihongo-gaku* (Japanese Linguistics) 7(12).4-11.
- Mizutani, Osamu, and Nobuko Mizutani. 1987. How to be polite in Japanese. Tokyo: The Japan Times.
- Miller, Laura. 1989. The Japanese language and honorific speech: Is there a Nihongo without keigo? In *Penn Linguistics Review* 13.38-46.
- Mori, Junko. 1996. Negotiating agreement and disagreement: The use of connective expressions in Japanese conversation. Doctoral dissertation. University of Wisconsin, Madison.
- Okamoto, Shigeko. 1995. "Tasteless" Japanese: Less "feminine" speech among young Japanese women. In K. Hall and M. Bucholtz (eds.) *Gender articulated: Language and the socially constructed self*. New York: Routledge. 297-325.
- _____. 1997. Social context, linguistic ideology, and indexical expressions in Japanese. *Journal of Pragmatics* 28.795-817.
- Pomerantz, Anita. 1984. Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In M. J. Atkin and J. Heritage (eds.), *Structures of social action*. Cambridge: Cambridge University Press.
- Sacks, Harvey, Emanuel Schegloff, and Gail Jefferson. 1974. A simple systematics for the organization of turn-taking in conversation. *Language* 50.696-735.
- Schegloff, Emanuel, Gail Jefferson, and Harvey Sacks. 1977. The preference for self-correction in the organization of repair in conversation. *Language* 53.361-382.
- Watanabe, Suwako. 1993. Cultural differences in framing: American and Japanese group discussions. In Deborah Tannen (ed.), *Framing in discourse*. 197-209. NY/Oxford: Oxford University Press.
- Wierzbicka, Anna. 1991. Japanese key words and core cultural values. *Language in society* 20.333-385.

**Discourse-based solutions to quantitative problems in sociolinguistics:
The case of men's and women's speech in an Indo-Guyanese village ***

Jack Sidnell
University of California, Los Angeles

1. INTRODUCTION. In a recent paper, Penelope Eckert (1996) outlined the two main approaches to language variation pioneered by William Labov. On the one hand, a number of large-scale surveys have demonstrated highly regular and systematic correlations between language variation and the social parameters of age, sex, class and formality of style. Other studies have been based on sustained ethnographic research in smaller communities and have illustrated the way in which language variation and change is embedded in local social, economic and demographic processes. Eckert then goes on to remark that:

Broad demographic patterns of variation show us the general patterns of the distribution of variables and in some cases the patterns of the spread of change. They also suggest possible social interpretations of variation. But information about how variation actually functions as a communicative resource lies at the local level. If variation does serve as a symbolic resource at the local level, what are the nature and limits of its symbolic potential? What is the relation between the symbolic value of one variable and another in the same community? What is the nature of the relation between local symbolic value and the global patterns that show up in survey studies? And how are the summative correlations related to the actual use of variation in social practice (Eckert 1996:48).

In the following, I take up Eckert's suggestion that "information about how variation actually functions as a communicative resource lies at the local level." I argue that attention to variant forms as communicative resources can open up new kinds of research questions and allow for a higher level of integration between quantitative and qualitative approaches to language use. It is suggested that some variable phenomena require discourse-based explanations. This is to say that, without adequate attention to the way in which individual variants function as communicative resources in interaction, we will not be able to fully account for statistical and quantitative patterns of variability and observed correlations to social categories.

2. THE SETTING. All the speakers in this sample live in a village of just under 700 people about 45 minutes drive from Georgetown in East Coast Demerara, Guyana. Families vary in size and composition. Most common is the nuclear unit comprised of a coresident husband, wife and their children. Houses are situated on plots of land belonging to a patrilocal corporate group. The most important economic activity in the area is rice-growing and traditionally the patrilocal group was important for the organization of labor in this endeavor. Since the mid 20th

century small peasant land-holdings with adjoining rice fields have gradually given way to the large-scale, highly mechanized rice growing industry (Hanley 1979, Sidnell 1998). Thus wage-labor has generally overtaken in importance the traditional modes of domestic labor organization.

Generally, married women live with their husband's family in one fenced *yard* 'unit of land belonging to patrilocal group.' This has effects for the operations of domestic power. Living in the same yard with a father and mother-in-law as well as a number of other affines, a wife has to exercise a certain care in how she comports herself. Very little escapes the watchful eyes of her in-laws. Married women thus face serious restrictions in terms of movement and are generally expected to spend most of their time in and around the house doing domestic labor. Men are much more free to use the road not just as a thoroughfare but also as a place to socialize with other men. Men have strong commitments to their peer groups which are often instrumental for the organization of wage-labor. Often, married men will spend the majority of their time out on the road, at work or in the rumshop and, for several days, may come home only to eat and sleep. Although the village is stratified in terms of access to socio-economic resources the situation is rather less polarized than that described for former sugar plantation villages by Rickford (1979) and Jayawardena (1963). In such communities there is a well recognized boundary between those who work in the fields (Estate Class) and those who do not (Non-Estate Class). Historically, advancement into the Non-Estate Class was associated with the adoption of the values and expressive behavior (including language) of the colonialists. The village where the present research was conducted was never associated with a plantation economy. The property, rather, was bought by 13 formerly indentured East Indians in 1883 and at that time rice was probably already being farmed there. Because of the labor requirements of non-mechanized rice farming, the patrilocal corporate group took on a great deal of importance for the organization of everyday life. Authority was structured along kinship rather than class lines and those people that did succeed economically were never required to adopt the values and expressive behavior of the British colonials. According to current village inhabitants this situation persisted until the 1960's. Prestige and social power was more associated with economic and agricultural success than with formal schooling. Although this situation is now changing its effects are still quite evident. Most important for the present study is the fact that, just about everybody in the village uses a fairly basilectal to mesolectal variety in everyday conversation. Even school teachers, both men and women, use mesolectal varieties outside the classroom.

3. PRONOMINAL VARIATION IN A CREOLE CONTINUUM. Pronouns in Guyanese Creole, as Bickerton (1973) and Rickford (1979) have illustrated, show a kind of robust variation along a number of dimensions and this makes them particularly well-suited to variation analysis. Basilectal singular pronouns contrast

morphologically with corresponding mesolectal and acrolectal ones in a number of subcategories.¹ This is illustrated in Table 1:

TABLE 1. *Contrasting Basilectal, Mesolectal and Acrolectal Pronouns.*

	SUBJECT	OBJECT	GENITIVE	
1st	mi		mi	Basilect
	ai		mai	Acrolect
2nd			yu	Basilect
			yor	Acrolect
3rd/	masc	am	ii	Basilect
		(h) ii		Mesolect
		him	hiz	Acrolect
	fem	am	ii/shi	Basilect
		shii		Mesolect
		hor	hor	Acrolect
neut		am	ii	Basilect
		it	its	Acrolect

Speakers in the sample did not approach the acrolectal end of the continuum in trinomial subcategories. We are thus left with binomial variables in each category. Genitives show variation when the Guyanese population is considered as totality.² However the corpus used here is confined to relatively basilectal speakers who generally do not use the acrolectal forms in this subcategory. For this reason, I have included genitives in the analysis only at select points since variation is not robust enough to give a detailed analysis for this category. We are left with the variables isolated in table 2.

TABLE 2. *Contrasting Basilectal and Relatively Acrolectal Pronouns.*

	SUBJECT	OBJECT	
1st	mi		Basilect
	ai		Acrolect
3rd/	masc	am	Basilect
		(h) ii	Mesolectal
	fem	am	Basilect
		shii	Mesolectal
	neut	am	Basilect
		it	Mesolectal

The analysis that follows thus focuses on variability in two subcategories; 1st person subjects and 3rd person objects.

4. DIFFERENCES BETWEEN WOMEN AND MEN. In the community studied, men and women do not use these variants to the same extent. Thus if we take genitives, 3rd person objects and 1st person subjects together we find men using slightly more basilectal variants than women.

TABLE 3. *Distribution of Basilectal Subjects (1st), Objects (3rd) and Genitives by Sex of Speaker*

	% of basilectal variant	total	varbrul weighting for rule application (basilectal variant)
men	85	428	0.568
women	75	371	0.422

However, we find that, with regard to first person subjects (alternation between *ai/mi*), women appear more basilectal:

TABLE 4. *Distribution of Basilectal 1st Person Subjects by Sex of Speaker*

	% of basilectal variant	Total	varbrul weighting for rule application (basilectal variant)
men	92	442	0.379
women	97	251	0.663

This is consistent with the results reported in Rickford (1979). In the subcategory of objects, again consistent with Rickford’s (1979) analysis, men appear slightly more basilectal than women.

TABLE 5. *Distribution of Basilectal 3rd Person Objects by Sex of Speaker*

	% of basilectal variant	Total	varbrul weighting for rule application (basilectal variant)
men	58	178	0.554
women	48	203	0.453

The relatively minor effect of speaker’s sex is magnified if we consider only the most stigmatized forms in this group: the use of basilectal marker *am* for animate objects (cf. Rickford 1979).

TABLE 6. *Distribution of Basilectal 3rd Person Objects for Animate Referents only by Sex of Speaker*

	% of basilectal variant	Total	varbrul weighting for rule application (basilectal variant)
men	56	124	0.610
women	37	158	0.413

Now supposing that basilectal variants in each subcategory had similar pragmatic-indexical values in this community (i.e. that saying *mi* as opposed to *ai* and saying *am* as opposed to *ii/shii/it* amounted to essentially the same thing) we might agree with Rickford (1979) that there is a leveling out of differences which gives the overall impression that gender is relatively insignificant. Alternatively, if we accept that different variables within the category of pronouns may be associated with quite distinct indexical meanings we might still attempt to explain gender differences according to a single coefficient for sex along traditional Labovian lines. Thus, we might suppose that women are lagging in the use of variants that are stigmatized within the larger community (stable sociolinguistic variables) while they lead in changes that carry local, overt prestige (changes in progress). But this explanation is not appropriate either. While *am* as 3rd person object might classify as a variant stigmatized in the larger community (see below), *ai* definitely carries local overt prestige as opposed to *mi*. Women's behavior in this community thus does not conform to what has traditionally, but controversially, been labeled the "gender pattern" (cf. critique of "gender pattern hypothesis" in Eckert 1989, Haeri 1987).

5. VARIATION WITHIN AND BETWEEN CATEGORIES. So why do men and women differ in their use of these variable categories? In the following I argue that the observed distributional patterns for men and women can only be adequately explained if we attend to the particular functions of each variable category in its conversational context. Rather than attempting to analyze both variables as instances of a more general category (i.e. pronouns), I point to ways in which they are functionally differentiated and suggest that these differences have implications for the way variation is exploited by men, on the one hand, and women on the other.

6. THE PRAGMATIC VALUES OF 1ST PERSON SUBJECTS. Because first person reference is necessarily self-referential (referring, that is, to the shifting role of speaker), acrolectal variants take on special meanings in this context. Specifically, the prestige meanings associated with the acrolectal variant (*ai*) are predicated of the speaker. *Ai* usage thus involves not only an assertion of acrolectal competence,

as do other elements of the idealized acrolectal code, but simultaneously a foregrounding of the assumed identity of the speaker.

The explanation for differences here then involves the way in which men and women assume positions as authoritative speakers. For men the use of *ai* is associated with the presentation of self as respectable and removed from the rural or working-class lifestyle. Often this occurs in cases where the speaker is attempting to manipulate a hearer who might be sensitive to such forms of social distinction. The manipulation in such cases depends for its effectiveness on assuming a position of respectability. Consider the following (reported) example of a prayer to god.

NS: *if mi gu sit dong in mi alta nou. mi gu see mi noo fu taak tuu gad. mi mos kom -*
If I go and sit down in my altar, I will say I know how to talk to god. I must come,

mi mos see oo gad ai wanch yuu protek mii, ai wan yuu giv mii helt,
I most say, "Oh God, I want you to protect me. I want you to give me health

an strengt ai wan yuu protek mi hous, protek mii pikniii dem, yuu noo?
and strength, I want you to protect my house, protect my children." You know?

Women also use *ai* with manipulative predicates in request-type acts. Consider the following example in which an aunt is requesting her nephew (3 years) to talk:

SS: *ai want yuu taak. yuu mos taak le mi hiir hou i a taak.*
I want you to talk. You most talk let me hear how he talks.

Men also characteristically use *ai* when calling to women on the road. Usually the woman is someone not completely familiar with the speaker. The following is a reported example:

TS: *wel yu gu star prooch shi- yu noo?- kyeer wan- kyeer inglish langwidj tuu shi-*
Well you will start to approach her. You know? Use a-use English with her.

see heloo. ai lov yuu beebii. wats op? yu noo? ai laik di wee yuu wak-
Say, "Hello, I love you baby. What's up?" You know? "I like way you walk."

The particular effect of the acrolectal variant is, in this case, linked to its pragmatic and referential value. Because it is necessarily self-referential the assumed acrolectal, cosmopolitan identity of speaker is foregrounded. Felicitous usages of *ai* thus index authority and respectability. Women generally avoid the use of *ai* because such presentations of self are, for them, more likely to backfire. Women will often be called on their uses, and an audience may contest the assumption of such an obviously non-rural identity. In one case that I recorded in my field notes, a young women who was taking basic secretarial lessons in the village called out to a friend across the fence:

girl: *hai darling. ai goin in. ai gon sii yu leeta.*
 Hi darling. I'm going in. I will see you later.

A group of young men overheard this and proceeded to mock the girl. The boys called out "hi, hi I'm going in I'm going in for my lessons, I'm coming out later." Women, as opposed to men, are much more likely to be called on such usages if people suspect that they might be assuming a role beyond their country origins. Similarly, women are more likely to be labeled *biggity* 'arrogant' in *talking-name* 'gossip' sessions both by men and other women. Although men are also subjected to this ridicule it is less likely to have damning effects for their reputation. Furthermore, young men often use *ai* in situations where the key is fairly playful. When calling to a girl on the road with *ai lov yu* "I love you," a young man cannot expect to be taken seriously. There are too many co-occurring signs (such as dress, the location of the interaction) which indicate that he is not what he is pretending to be and that his intentions are less than completely sincere. Women who use *ai* are more likely than men to be interpreted as self-elevating - hearers will often assume that these women actually believe themselves to be a member of the class who habitually use *ai*.

7. THIRD PERSON OBJECT VARIATION. As already noted, women, as compared to men, appear to be more basilectal in their use of first person subjects, whereas, in the subcategory of third person objects, they appear more mesolectal than the men. Looking carefully at the variants involved, the third person objects present us with a rather complicated problem. As the basilectal pronominal system gives way to the mesolectal system, case distinctions are lost and animacy and gender distinctions are added (Bickerton 1973, Rickford 1979). In using the mesolectal terms, then, speakers must assess the referent's relative animacy and its gender (for animates). Not surprisingly, one finds that both the gender and the relative animacy of the referent have an effect on the frequency of variant pronouns usages. *Am* is most basilectal and is most infrequent when it is used as a referring term for feminine referents (see below). There are more and less favored environments for the use of *am* and these depend crucially on characteristics of the referent. As it turns out, although we find men and women using *am* at only marginally different frequencies (Table 5), women overwhelmingly avoid using *am* in its most creole and stigmatized contexts (i.e. for animate referents, Table 6).

8. ANIMACY AND GENDER EFFECTS. Speaking of the variable use of *am*, Rickford (1979:359-360) remarks that there is

some indication that it is acquiring (or has acquired) a certain degree of specialization as an appropriate form for neuter (or non-human) objects...*am* is still a more stigmatized form than any of the corresponding subject or possessive forms in

the basilect, but speakers seem to operate with a rule that says: if you are going to use it, use it more often for neuter objects than any other.

The data here confirms such a tendency to reserve *am* for inanimate referents. This is clear from the results of a binary coding:

TABLE 7. *Effects of animacy on the variability of 3rd person Objects (% of am usage)*

	% basilectal variant	Total	varbrul weighting for rule application (basilectal variant)
animate	46	273	0.419
inanimate	75	93	0.722

However the effect of animacy of the referent is rather more complicated than this binary distinction suggests since it interacts with gender-of-referent effects. In fact, if one codes for the gender of the referent according to a tripartite scheme which distinguishes male, female and neuter entities one finds that neuter (or inanimate) and male referents overwhelmingly favor basilectal marking but female referents do not.

TABLE 8. *Effects of referent gender on the variability of 3rd person Objects*

referent	% basilectal variant	Total	varbrul weighting for rule application (basilectal variant)
female	17	147	0.155
male	80	110	0.781
neuter	75	109	0.731

It is apparent that female referents favor the use of a mesolectal variant (*shii*) while male and inanimate favor marking with basilectal *am*. Both Bickerton and Rickford have proposed somewhat tentative explanations for this phenomena, but I cannot review their arguments here (see Sidnell forthcoming). Rather, I want to suggest a solution which attends to the way in which these variants figure in terms of conversational practice. To this point we have seen that variation in the category of 3rd person objects, is affected by the animacy of the referent and that women generally avoid the use of *am* when referring to animate referents. At the same time, all speakers favor mesolectal marking in the feminine subcategory (Table 8). It appears then that women have set a community norm here. Women, more than men, tend to reserve *am* for inanimates and prefer gender-marking pronouns for animates. If women lead in this general process it is not too hard to imagine that they set a pattern in which mesolectal marking is preferred for feminine subcategories in particular.

Some support for this suggestion comes from the way in which men and women differ in their use of *am* and the mesolectal variants. Although basilectal variants are strongly dispreferred by both men and women in the feminine subcategories, men are more likely than women to use *ii* for male referents while women are more likely than men to use *shii* for female referents. Such patterns indicate an area of contestation between men and women. *When men talk about other men they use the term marked for gender and animacy more often than women do. When women talk about other women they use a term marked for gender and animacy more often than men do.*

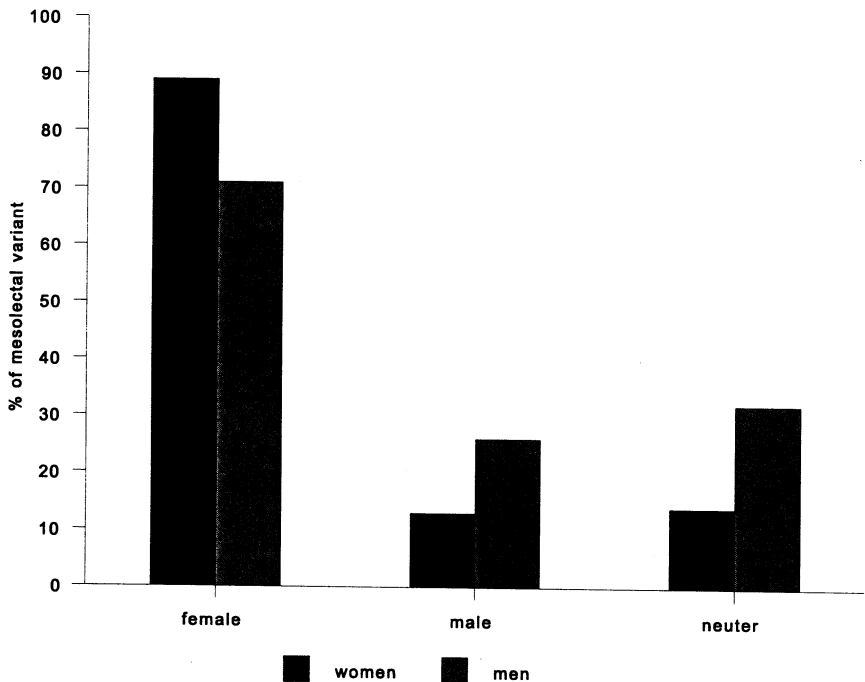


FIGURE 1. *Gender and animacy effects for men and women.*

Figure 1 shows that the influence of gender and animacy-of-referent effects is partially collapsed for men as opposed to women with differences between categories being relatively contracted for the male group. For men, the variation seems to be between more and less creole (or basilectal) ways of saying the same thing. This is to say that, *for men, the alternation between am and its mesolectal variants is determined more by the social, than the referential, significance attached to each variant. Am* is in fact strongly implicated in the expression of

working class solidarity and local values. Social and stylistic factors no doubt affect women's usage of the variants too. However, for women another, referential, dimension becomes particularly salient. In avoiding the most stigmatized uses of *am* in reference to animates women are establishing a firm association between this form and inanimates. Furthermore, women show a strong preference for mesolectal marking for members of their own sex-based category. This acts to highlight and foreground the animacy of the referent in discourse. It appears then that women and men are using the variability of the pronominal system quite differently. For women, more than men, characteristics of the referent, in addition to social and stylistic factors, play a major role in determining patterns of variability.³

The following is an example showing the general preference for mesolectal marking for feminine referents and basilectal marking for neuter/inanimate referents. S is complaining to her sister-in-law (Sh) that a young girl who lives next door was climbing on her fowl pen to pick plums and subsequently broke it but would not tell her mother (Miss):

S.: *shi a pik plom. wen mi aks shi yestodee see shi na tel mii se shii*
 She was picking plums. When I asked her yesterday why she didn't tell me that she

brook am dong shi na ansa -- shi klaim pon di foul pen an brook am dong (...)
 broke it she didn't answer. She climbed on the fowl pen and brook it down. (...)

den mi se shi na big inof fu tel Mis da shi klaim pon-
 Then I asked if she wasn't big enough to tell Miss that she had climbed on

Sh: *-oo yea*
-Oh yeah

S.: *- am brook am. mi na gu tel shii koz shi gu kos.*
- uhm broke it. I'm not going to tell her because she will curse.

The next example illustrates the alternation between the mesolectal term for feminine referents (*shii*) and basilectal term for masculine ones (*am*). Kavita is telling me how she, her mother, and her aunt had been playing with a little girl Mando (3 years) who was also present at the time this report was given. During the course of the telling, Kavita's brother Kumar asks Mando to repeat her performance from the day before. Kavita tells Mando to curse Kumar, her brother, switching to the basilectal pronoun (*am*) in doing so:

Kavita: *yestodee Mamii tel antii Beebii fu tel shii toch mii an shi toch shi so*
 Yesterday Mama told Aunty Baby to tell her "touch me" and she touched her like this

an mi tel shi toch mi so an shi toch mi. abii tel tel shi don noo
 and I told her "touch me" like this, and she touched me. We said "tell her 'right-on.'"

hiir shii mii don noo. <laughs>
This is what she said "right-on."

Kumar: *toch mi de Mandoo.*
Touch me there Mando.

Kavita: *sok yu teet pon am. kos am. tel i le i gu beed i doti skin.*
Suck your teeth on him. Curse him. Tell him to go and wash his dirty skin!

9. CONCLUSIONS. To summarize the argument, I have demonstrated that the two variables discussed show significant but seemingly inconsistent effects for the sex of the speaker, and have argued that in order to account for these inconsistencies it is necessary to look at the internal structure of each variable, particularly their pragmatic and referential qualities. In the case of *ai/mi* alternation, I have argued that differences between men and women are a result of different interactional strategies. Specifically, because these are deictic forms, indexing speech act variables, they have rather specific indexical entailments: *ai* usage foregrounds an assumed identity for the speaker - one that is relatively more difficult for women, as opposed to men, to sustain. In the case of alternation between *am* and its mesolectal variants, I have suggested that women show a rather different kind of sensitivity to the referential or semantic distinctions involved, in addition to social and stylistic factors. Patterns of use with feminine, masculine and neuter referents (figure 1.) indicate that men and women employ rather different strategies in marking animacy through the use of pronominal variants.

To conclude, let me reiterate that the argument here is meant as a first step toward a discourse-based approach to language variation. Ultimately the goal is to integrate quantitative approaches to language variation with qualitative approaches to language-use, discourse and conversation. In cases where statistical analysis reveals seemingly inconsistent distributional patterns, researchers may find explanations in the fabric of actual stretches of talk-in-interaction.

* **ACKNOWLEDGMENTS.** Funding for research in Guyana between 1994 and 1996 was provided by the Wenner-Gren Foundation for Anthropological Research and the Social Sciences and Humanities Research Council. I'd like to thank Jack Chambers, Ignasi Clemente, Sandro Duranti, Candy Goodwin, Hy Van Luong, Bonnie McElhinny and Jennifer Reynolds for their comments on previous versions of this paper. They are, of course, in no way responsible for the contents of the present one.

NOTES

¹ I have purposefully excluded variants which differ at a phonological level since such cases require levels of explanation which are beyond the scope of this paper.

Rickford (1979) gives comprehensive accounts of vowel laxing, h deletion, and t deletion in pronoun forms. Vowel laxing produces the alternation between forms such as *mi* and *mii*, *shi* and *shii*.

² There were a few exceptions to this rule. Genitives are included only in the calculations which assess overall variability for all pronoun subcategories (Table 3) in part to make the results comparable to those reported in Rickford 1979.

³ In accounting for such a situation we might suppose that either men are not attending to the relative animacy of the referent or that they are purposefully characterizing all referents unremarkable with regard to animacy.

BIBLIOGRAPHY

- BICKERTON, DEREK. 1973. The Nature of a Creole Continuum. *Language* 49: 640-69.
- ECKERT, PENELOPE. 1989. The whole woman: Sex and gender differences in variation. *Language Variation and Change* 1: 245-267.
- ECKERT, PENELOPE.. 1996. (ay) Goes to the city: Exploring the expressive use of variation. *Towards a Social Science of Language. Papers in honor of William Labov. Volume 1: Variation and Change in Language and Society*, ed. by G. Guy, C. Feagin, D. Schiffrin and J. Baugh. Amsterdam: John Benjamins.
- HAERI, NILOOFAR. 1987. Male/female differences in speech: An alternative interpretation. *Variation in Language: NWAV-XV at Stanford*, ed. by K.M. Denning et. al.. Stanford: Stanford University, Department of Linguistics.
- HANLEY, ERIC. 1979. Mechanized rice cultivation: the experience of an East Indian community in Guyana. *Peasants, Plantations and Rural Communities in the Caribbean*, ed. by A. Mauss and M. Cross. Guildford: University of Surrey.
- JAYAWARDENA, CHANDRA. 1963. *Conflict and Solidarity on a Guyanese Plantation*, London: Althone.
- RICKFORD, JOHN. 1979. *Variation in a Creole Continuum: Quantitative and Implicational Approaches*. University of Pennsylvania dissertation.
- SIDNELL, JACK. 1998. *Gender, Space and Linguistic Practice in an Indo-Guyanese Village*. University of Toronto dissertation.
- SIDNELL, JACK. forthcoming. Gender and pronominal variation in an Indo-Guyanese village. *Language in Society*.

**How 'give' and 'receive' provide structure for more abstract notions:
The case of benefactives, adversatives, causatives, and passives***

Tomoko Yamashita Smith
University of California, Berkeley

1. Introduction

This paper examines cross-linguistic categories of grammatical constructions and proposes that such categories can be polysemously related through metaphor. Lakoff and others have argued that the human conceptual system fundamentally involves metaphorical structure (e.g. Lakoff & Johnson 1980). Moreover, other work has suggested metaphorical patterns in grammar. For example, much has been written on the metaphorical extension from the space domain to the time domain (e.g. Traugott 1975, Lakoff & Johnson 1980, Svorou 1993). It has also been suggested that metaphor is one of the forces that can drive the process of grammaticalization (Sweetser 1988, 1990, Heine et al. 1991). Finally, Goldberg (1995) argues that constructions can be related by metaphorical mappings.

In this paper, I discuss how the semantics of four classes of grammatical constructions — 'benefactives', 'adversatives', 'causatives', and 'passives' — are related. I also argue that there is a cross-linguistic tendency for the notions expressed by the four constructions to be explicable in terms of 'giving and receiving' and that, therefore, 'giving and receiving' must be able to serve as concrete basis for languages to express the more abstract notions of the constructions in question. To support this claim, I show that (1) many unrelated languages have grammaticalized 'give' and 'receive' to fill a range of these related notions; (2) within one language two of the constructions can share a single syntactic form; and (3) a member of one of these classes of constructions may develop other meanings from this semantic "family" over time.

Some examples of the four construction types are given below. One kind of benefactive construction is exemplified in the English sentence (1), where the beneficiary is coded as the direct object.

- (1) *Mark baked Linda a cake.*

Adversative constructions are yet to be studied in-depth. However, it is known that they may take a variety of forms. To give just one example, Even, a Tungusic language, has adversative constructions that use an adversative marker, *-v*. These constructions express an action or event unfavorable for the subject, as in the following example (Malchukov 1993: 2).

- (2) *Hulič̣an-Ø bodel-Ø-i ene-le-v-re-n.*
fox-NOM feet-NOM-REF POS SG hurt-INCH-ADV-NONFUT-3SG
'The fox's paws began to hurt; he was negatively affected.'

Although they may not be as common as causatives or passives, languages can have distinct adversative constructions. Thus, it is important to view the adversative construction as a separate grammatical category.¹

The semantics of causatives are very complex, and there are many semantic subclasses including, for example, the 'coercive' causative as in *Mark made Linda eat the cake* and the 'permissive' causative as in *Mark let Linda eat the cake*. This paper will treat both types.²

The passive category also has various subclasses, among which are the direct and indirect passives. Japanese allows both transitive and intransitive verbs to be passivized, as in (3) and (4), resulting in a direct and indirect passive, respectively. In (3), Ken is a regular subject of the transitive passive, but in (4) Ken is an added argument, referred to as an "affectee".

- (3) *Ken* *ga* *Mari* *ni* *koros-are-ta.*
 Ken SUB Mari AGT kill-PASSIVE-PAST
 'Ken was killed by Mari.'
- (4) *Ken* *ga* *akatyan* *ni* *nak-are-ta.*
 Ken SUB baby AGT cry-PASSIVE-PAST
 'Ken was adversely affected by the baby crying.'

If syntax alone is used to compare the direct passive and indirect passive, it is difficult to treat them alike; however, using the cognitive framework I propose in section 2, their similarity can be captured.³

1.1. Previous research

There are countless studies of the semantics of both causative and passive constructions (e.g. Shibatani 1973, Talmy 1988, Kemmer & Verhagen 1994, Song 1996 for causatives and Davison 1980, Siewierska 1984, Keenan 1985 for passives). There are numerous studies on the benefactive construction, including a cross-linguistic study by Shibatani (1996). There are also some studies on the adversative construction (Oehrle & Nishio 1981, Wierzbicka 1988, Shibatani 1994). Moreover, there has been research on the relationship between some of the above.

The correlation between the causative and the passive has been much discussed (e.g. Shibatani 1985, Haspelmath 1990, Washio 1993). For example, Ikegami (1981) compares Japanese causatives and passives and gives a framework for the two constructions based on localist theory. He uses the notion of 'cause' for causatives and 'get' for passives.

Tuggy (1988) offers an account of the relationship between causatives and benefactive applicatives in an Aztec language, Tetelcingo Nahuatl, where an identical verbal suffix is used in both constructions. He argues that causatives and benefactive applicatives are similar since benefactive applicatives also involve a kind of causation, namely, the causing of 'possession'.

While the previous treatments make good points, they explain only parts of the big picture. None posits a relationship between the four constructions. Moreover, in discussions of the causative, the term 'cause' alone is often used to explain the notion of 'causation'. However, as Lakoff (1977) and later Lakoff & Johnson (1980) point out, despite the fact that causation is a basic human concept, it is not an undecomposable primitive. I argue that the notions involved in all four constructions, including the notion of causation, can be construed in terms of the prototypical 'giving and receiving' scene.

2. Object exchange model: metaphorical mappings of the constructions

I argue that causative, passive, adversative, and benefactive constructions express semantic categories that are construed metaphorically and that there is a cross-linguistic tendency for the four constructions to reflect construal as metaphorical transfers.⁴ In other words, people can construe prototypical benefactive, adversative, causative and passive scenes as structured by a 'giver'- 'object'- 'recipient' relationship, and this construal is reflected in languages. I will call this construal of the four construction types the "object exchange model". Below, I discuss the metaphorical mappings likely to be involved in it.

With both benefactives and adversatives, the agent is mapped onto the giver, and the affectee is mapped onto the recipient. What is "transferred" from the giver to the recipient is the event and its effect — a positive effect, or benefit, in benefactive constructions and a negative effect in adversative constructions.

In passives, the agent is mapped onto the giver and the patient (of the direct passive) or affectee (of the indirect passive) is mapped onto the recipient. What is transferred is the event and its direct or indirect effect on the patient/affectee.

In causatives there are two mappings. Essentially, the causer is understood as the giver and the causee as the recipient. What is transferred in this case is the effect of the primary agent's causal agency onto a secondary agent, the causee.⁵ However, the causer is also understood as the recipient of the event and its effect, and the causee is mapped onto the giver. In prototypical causative situations, the causer is a volitional entity. S/he wants the causee to carry out the event and wants an effect that the event will produce. The following is a summary of the metaphorical mappings of the four constructions.

Table 1. Mappings

Constructions	Giver	Recipient	Transferred object
benefactive	agent	affectee	event & its positively evaluated effect
adversative	agent	affectee	event & its negatively evaluated effect
passive	agent	patient/affectee	event and its effect
causative	causer	causee	effect of causal agency
	causee	causer	performed action and its effect

3. Supporting evidence

3.1. Uses of 'give' and 'receive/get' as benefactives, adversatives, causatives and passives

There are synchronic and diachronic data to support the object exchange model. First, 'give' and 'receive/get' are used either in full lexical or grammaticalized form to express the four relevant notions in a variety of unrelated languages (see Tables 2 and 3). It should be noted that with all but the causative construction the use of one verb or the other ('give' or 'receive') reflects only a difference in point of view, not in the mapping as shown in Table 1. For example, if 'give' is used in a benefactive construction, then the subject is an agent; if 'receive/get' is used, then the subject is a beneficiary. Yet, in both cases, the mapping remains the same. In causative constructions, however, the use of 'give' involves one mapping and 'receive/get' another: a 'give' causative indicates that the causee receives an effect, while a 'receive' causative indicates that the causer receives the event.

Table 2. 'give'

constructions	languages (language family)
benefactive	Akan, Ewe, Fon, Nupe, Yoruba (Niger-Congo), Saramaccan (West African Creole), Thai (Tai), Vietnamese (Austro-Asiatic), Lahu, Mandarin (Sino-Tibetan), Marathi, Sinhala (Indo-European)
benefactive and adversative	Newari (Sino-Tibetan), Japanese, Korean
passives	Mandarin, Wu, Min, Hakka (Sino-Tibetan), Buyi (Sino-Tibetan or Austro-Tai), Yao (Sino-Tibetan (?)), Li/Hlai (KadaiTai (?))
causative	English, Bulgarian (Indo-European), Nandi, Luo (Nilo-Saharan), Alawa (Australian), Somali (Afro-Asiatic), Thai (Tai), Lahu (Sino-Tibetan), Yao Samsao (Sino-Tibetan or Austro-Tai), Vietnamese, Khmer (Austro-Asiatic)

Table 3. 'receive/get'

benefactive	English (Indo-European), Sgaw Karen (Sino-Tibetan), Vietnamese (Austro-Asiatic), Japanese
passive	English, Welsh (Indo-European), Thai (Tai), Vietnamese (Austro-Asiatic), Tzeltal (Mayan)
causative	English (Indo-European)

As shown in Table 2, many languages use 'give' in benefactives, but 'receive/get' is also used in benefactives in unrelated languages, such as Vietnamese and English, shown in sentence (5).

- (5) *My father is taking me to LA, and I get to go to Hollywood.*

Moreover, some languages, such as Japanese, use both 'give' and 'receive' in benefactives. In Japanese, when 'give' is used, the agent or the giver is coded as the subject, as in (6).

- (6) *Kei ga Naoko ni hon o katte-yat-ta.*
 Kei SUB Naoko DAT book ACC buy-give-PAST
 'Kei did the favor of buying Naoko a book.'

When 'receive' is used, the beneficiary is coded as the subject, as in (7).

- (7) *Naoko ga Kei ni hon o katte-morat-ta.*
 Naoko SUB Kei DAT book ACC buy-receive-PAST
 'Naoko received the favor of Kei's buying a book.'

In Newari, 'give' is used not only in benefactives but also in adversatives, as in (8) and (9).

- (8) *Rām-āñ Gita-yāta lukha cāekā: bila.*
 Ram-ERG Gita-DAT door.ABS open.PP give.PD
 'Ram opened the door for Gita.'
- (9) *Jimi macā mhiga: sinā: bila.*
 my child.ABS yesterday die.PP give.PD
 'My child died on me yesterday.'

'Receive/get' is commonly grammaticalized into a passive marker. Languages that use 'receive/get' in passives include English, Welsh, Tzeltal, and Thai. (Siewierska 1984, Haspelmath 1990).⁶ Sentence (10) is an example from English in which Mary is seen as the recipient of the event.

- (10) *Mary got hit by a car.*

On the other hand, 'give' is also used in passives in dialects of Chinese (Hashimoto 1988). Example (11) is from Mandarin and, in this case, the agent is metaphorically seen as the giver of the action with *gěi* 'give' introducing the agent of a passive sentence.

- (11) *Nèipán cài gěi tā chīguāng-le.*
 That vegetable give him eat-finish-ASP
 'That plate of food was finished by him.' (Tiee 1986: 298)

In the case of the causative, the use of 'give' as a lexical source for causative marking is so common that it can only be explained by postulating a cognitive connection between causation and giving. In other words, it may simply be easy for humans to construe causation as a form of giving.⁷ Example (12) is from Thai and (13) is from Alawa.

- (12) *Saākahaã hây dèk wîŋ.*
 Saka give child run
 'Saka had the child run.' (Vichit-Vadakan 1976:460)
- (13) *Lilmi-r•i mar• a-muta-ya-ngur•u da an-kir•iya.*
 man-ERG carry he-give-PST-her PRT CL-woman
 'The man made the woman carry it.' (Song 1996: 32)

There are also languages in which 'give' is used in the permissive causative, as in (14) from Yao Samsao.

- (14) *Maa pun fuʔ-cuəy cap budoʔ-gway*
 mother give child cut fingernails
 'The mother let the child cut his nails.' (Matisoff 1991:428)

In all the examples (12), (13), and (14), the causer is seen as the giver and causee as the recipient. However, it is also possible to see the causee as the giver and the causer as the recipient. That is, the causer receives the action performed by the causee as well as the action's effect. English causatives with 'get', as in (15), should be understood to have the causer as the recipient. I wanted him to stop smoking and I got him to stop smoking.

- (15) *I got him to stop smoking.*

Note that English also has a causative with 'have', which implies that the causer "obtains" the event done by the causee. Thus, the use of 'have' also shows that we can see the causee as a giver and the causer as a recipient.

In summary, the fact that numerous languages have grammaticalized the verbs 'give' and 'receive/get' in benefactives, adversatives, causatives, and passives supports the object exchange model.

3.2. Benefactives and adversatives as extensions of give constructions

Syntactic evidence also supports the object exchange model. That is, in some languages, benefactives and adversatives exhibit the same syntactic pattern as 'give' constructions.

For example, English ditransitives and benefactives have the same structure, as in (16) and (17).

- (16) *Bob gave Mary a cake.*
 (17) *Bob baked Mary a cake.*

As Shibatani (1996) argues, cross-linguistically, benefactives are best construed as extensions of a schema based on the 'give' construction (see also Goldberg (1995) for a similar argument about English).

Moreover, in some Indo-European languages, the dative construction is used to convey either a benefactive or an adversative notion (Wierzbicka 1988, Janda

1993, Shibatani 1994). The affected entity is marked as dative, which suggests that the affectee is seen as a recipient. The Czech sentence (18) demonstrates a benefactive reading.

- (18) *Ludmila mu ukazala cestu domu.*
 Ludmila-NOM him-DAT showed way-ACC home
 'Ludmila showed him the way home.' (Janda 1993: 48)

So far, two types of evidence have been presented to support the object exchange model: evidence from grammaticalization and evidence from syntactic patterning.

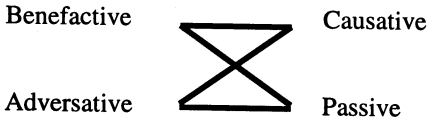
4. The relationship between the benefactive and adversative and between the causative and passive

In this section, I discuss the semantic relations between the constructions. Figure 1 is a schematization of their relations.

Figure 1.

Evaluation of an event

Internal structure of an event



Intuitively, these constructions can be grouped into two pairs, 'causative-passive' and 'benefactive-adversative'. As illustrated in the right column of Figure 1, the 'causative-passive' pair concerns the internal structure of causal events. Although the internal structures of prototypical causative and passive events differ (e.g. the prototypical number of arguments differs), in both cases the causal event includes an initiator and an endpoint.

Moreover, grouping the causative and passive together is supported by the fact that some languages (e.g. English, French, and Korean) use identical constructions to mark these two notions. For example, causative constructions in English and French can denote a passive sense, as in (19) and (20) (see Washio 1993 for more examples and analysis of ambiguity between causative and passive senses).

- (19) *John had his watch stolen.*
 (20) *Jean s'est fait voler sa montre.*

On the other hand, as shown in the left column in Figure 1, 'benefactive-adversative' denotes subjective evaluation of an event: the speaker judges the affectee as positively or negatively affected by the event performed by the agent.

Again, this grouping is supported by the fact that some languages use identical constructions to mark these two notions. For instance, in Lai, a Tibeto-Burman language, the benefactive particle *piak* may also convey the notion of 'adversity', depending on the context (see also the Newari examples (8) and (9)).

- (21) *Tsewmaŋ* *niʔ* *law* *ʔa-ka-thloʔ-piak.*
Tsewmang ERG field 3s-1s-weed-BEN
'Tsewmang weeded the field for me.'

- (22) *Tsewmaŋ* *ʔa-vok* *ʔan-thaʔ-piak.*
Tsewmang 3s-pig 3p-kill-ADV
'They killed Tsewmang's pig on him.'

The notion expressed by the causative construction contrasts with the notion expressed by the passive. Similarly the notion expressed by the benefactive contrasts with the notion expressed by the adversative. The opposed senses within each pair are mutually exclusive, so that we do not expect to find constructions that can be interpreted as simultaneously passive and causative, or benefactive and adversative. This opposition of senses within a pair also helps speakers to disambiguate, through context, sentences such as the above, since any evidence for one interpretation is evidence against the other.⁸

5. Combinations of the constructions' notions

Interestingly, since the causative-passive and benefactive-adversative pairs express conceptual and semantic parameters that are potentially independent of each other, it is possible to combine the notion expressed by a member of the causative-passive pair with a notion expressed by a member of the benefactive-adversative pair. This is indicated in Figure 1 by solid line connections. The possible combinations of the four constructions are: benefactive passive, adversative passive, benefactive causative and adversative causative.

Examples of such combinations are found among the world's languages. First, 'benefactive passives' exist in Thai, Vietnamese, and Korean (Davison 1980, Siewierska 1984), as shown in the Thai example (23).

- (23) *Phǒm* (*dâj*)-*rǎb* *chǎn.*
I (deferential) receive invite
'I was invited.' (Noss as cited in Davison 1980: 58)

The second combination, 'adversative passive', also exists in Thai, Vietnamese, Korean, Mandarin, and Japanese (Davison 1980, Siewierska 1984), as shown in the Thai example (24).⁹

- (24) *Khâw* *thùug* *chǎn.*
he suffer invite
'He was invited (against his will).'

Sentence (25) is an example from Russian of the third combination, the 'benefactive causative'. Here, the causer is also the beneficiary (Babby 1993).

- (25) *Ona sšila sebe novoe plat'e.*
 she-NOM sewed herself-DAT new dress-ACC
 'She had a new dress made (She had someone make her a new dress).'
 'She made a new dress (she did the sewing herself).'

This sentence has two possible interpretations. The first is the benefactive causative reading, which means that the agent is seen as both the giver and recipient in the causal event. The causer gives the effect of his/her causal agency to the causee and receives the event done by the causee. In addition, this received event is positively evaluated, and therefore the whole scene is understood as a benefactive causative.

An example of the fourth combination, adversative causative, is taken from Even, a Tungusic language (Malchukov 1993).¹⁰

- (26) *Bujuhemɣe-Ø buju-m ila-v-ra-n.*
 hunter-NOM wild deer-ACC stand up-ADV-NONFUT-3s
 'The wild deer stood up: the hunter was negatively affected.'
 (Malchukov 1993: 8)

An alternative translation of (26) is 'because of something the hunter did unintentionally, s/he caused the deer to stand up (indicating that it might flee), which negatively affected the hunter'. Sentences of this type have been described as nonvolitional permissive causatives, because the causal agent does not wish the event to occur but 'permits' it due to his/her own inattentiveness.

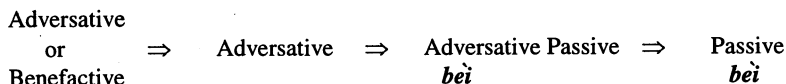
Unlike passives with benefactive or adversative readings, which are simply evaluated positively or negatively, causatives with benefactive and adversative readings seem to express different kinds of causation. Perhaps, the factor that makes the causatives either benefactive or adversative is 'volitionality'. That is, if the causer volitionally acts upon the causee, the causer *wanted* the event to happen, and getting what one wants implies a benefactive notion. This is the case with the Russian benefactive causative, as discussed earlier. However, if the causer involuntarily acts on the causee, what happens is an *accident*, which implies an adversative notion, as in Even's nonvolitional permissive causative.

6. Historical development

Finally, it is possible for one of the constructions in question to develop into another. For example, originally, Mandarin had markers to mark the direction of transitivity and to express 'affectedness', i.e. the subject's being affected by something 'beyond his/her control'. The use of these markers can be found the oracle-bone inscriptions from the 14th-11th C. BCE. Although initially this 'affectedness' could be interpreted as either 'benefactive' or 'adversative', it came to be interpreted as only 'adversative'. Then, by the 1st century BCE, these

adversative markers had been reanalyzed as passive markers, yet they retained their adversative sense as well. Now, in modern Mandarin, one of the markers, *bèi*, has started to function as a pure passive marker, without the original adversative sense, especially under the influence of and association with English passives (see Chen 1994 for a detailed explanation of the development of passive markers). The development of passive marking in Mandarin is roughly represented as follows.

Figure 2.



7. Conclusion

I have shown that, in many unrelated languages, verbs meaning 'give' and 'receive/get' are used in the conventional expression of the benefactive, adversative, causative and passive constructions; that some of the constructions' syntactic patterns can be identical to those of 'give' constructions; that syntactic patterns between pairs of constructions can also be identical; and that the constructions' notions may combine semantically and even evolve from one into another over time. All of these points provide indirect evidence for the psychological reality of the object exchange model as a basis for the development of the four constructions. If the model is a psychological reality, we can infer that one way in which the human mind structures language is by expanding a basic conceptual framework, such as the giving and receiving scene, in order to encode a broader range of meanings.

Although space does not allow a complete discussion, it should be noted here, briefly, that languages use a variety of words from within the giving and receiving scene as a lexical source for the four constructions. For example, English uses 'have' to indicate causation, Japanese can use 'from' to mark the agent of a passive; and Lahu can use 'come' in benefactive constructions ('come' is often part of receiving, because the received object can be perceived as coming to the receiver). Such words illustrate that metaphor can operate systematically within the framework, and suggest that the framework is robust.

This paper is an overview of the general relationship of the four constructions. For a better understanding of these constructions, more detailed study of particular languages is needed (for one such study of Japanese, see Radetzky & Smith to appear). In addition, to shed more light on the cognitive processes involved in the acquisition of the four constructions, studies in first and/or second language acquisition could be beneficial. Finally, it should prove instructive to analyze how the four constructions and the object exchange model fit within the larger picture of event structure.

NOTES

* An earlier version of this paper was presented at the 22nd annual meeting of the Kansai Linguistic Society, November 8, 1997 (Smith to appear). I am very grateful to Eve Sweetser, Tony Smith and Paula Radetzky for their suggestions and help throughout this paper's many revisions. I would like to thank Charles Fillmore, Yoko Hasegawa, Ben Bergen, Cindy Daugherty, and the members of the Japanese Linguistics Seminar at UC Berkeley for their helpful comments. Also, I would like to thank Kazuyuki Kiryu (Newari), Pongsak Rattanawong (Thai), and Ken Van Bik (Lai) for helping me with language data. Finally, my thanks to Carol Justus, Roula Svorou and Hana Filip for encouraging me to pursue this project during its conception and through earlier stages of its development.

1. The notion of adversity is sometimes expressed by different constructions such as passives and causatives. For instance, one type of Japanese causative expresses adversity (Oehrle & Nishio 1981). See section 4 for more about two notions co-existing in a single construction.

2. Lexical causatives can be explained with my model; however due to limitations of space, I do not include them as a part of my discussion. Nonetheless, lexical causatives can be understood as metaphorical transfers and have been discussed as such. For example, Lakoff (1993) points out that causation is seen as 'giving or taking'. Also Goldberg (1995: 144) discusses 'causal events as transfers'.

3. Passives have more than one function. One of them is to mark affectedness of the subject (Keenan 1985), and it is this function that is best explained in terms of my giving-receiving framework. Other functions of passives, such as defocusing the agent (Shibatani 1985), which appears to be behind the impersonal passive, are less clearly related to my model.

4. In his detailed cognitive linguistic study of 'give', Newman (1996) observes that 'give' is extended via metaphor cross-linguistically to express various abstract notions, including the 'causative' and 'benefactive' notions. However Newman's account of these constructions differs from mine in some important respects; moreover, Newman does not share my aim to relate these constructions.

5. In a more detailed view, the recipient and transferred object have dual status. The primary agent acts on the secondary, which also makes the secondary 'agent' an 'affectee' (as implied by 'recipient'), and the transferred object is not only the 'effect' of the primary agent's causal agency, but also the 'cause' for the secondary agent to perform the action.

6. Other lexical sources for passives include 'be', 'become', and 'stay'. Languages that use stative or inactive verbs to express passivity focus on the result of the event. This, however, does not refute the fact that passivity can be seen as a transfer.

7. For other lexical sources of causatives, see Li 1991, Pederson 1991, and Moreno 1993.

8. Because the four notions exist within the same framework, identical marking along horizontal and diagonal lines in Fig.1 should also be possible. Indeed, many languages use one morpheme to mark both causatives and benefactives (Frayzyngier 1985, Tuggy 1988, Matisoff 1991). English also uses 'get' in passives and benefactives. Identical marking in these cases, however, exists only in non-identical constructions, unlike identical marking for causative and passive pairs and benefactive and adversative pairs which can be disambiguated even when the form of the constructions is exactly the same.

9. The English get-passive also tends to express an adversative sense (Chappell 1980, Givón & Yang 1994)

10. In fact, this construction also shows characteristics of passive constructions (for a relevant discussion, see Malchukov 1993).

REFERENCES

- Babby, Leonard H. 1993. Hybrid causative constructions: benefactive causative and adversity passive. *Causatives and transitivity*, ed. by Bernard Comrie and Maria Polinsky. Amsterdam: John Benjamins. 343-68.
- Chappell, Hilary. 1980. Is the get-passive adversative? *Papers in linguistics* 13:411-52.
- Chen, Lilly Lee. 1994. The origin of adversative passives. *BLS* 20:113-24.
- Davison, Alice. 1980. Peculiar passives. *Language* 56:42-66.
- Frajzyngier, Zygmunt. 1985. 'Causative' and 'benefactive' in Chadic. *Afrika und Übersee*, 68:23-41.
- Givón, T., and Lynne Yang. 1993. The rise of the English get-passive. *Voice: form and function*, ed. by Barbara Fox and Paul J. Hopper. Amsterdam: John Benjamins. 119-49.
- Goldberg, Adele E. 1995. *Constructions: a construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Hashimoto, Mantaro J. 1988. The structure and typology of the Chinese passive construction. *Passive and voice*, ed. by Masayoshi Shibatani. Amsterdam: John Benjamins. 329-54.
- Haspelmath, Martin. 1990. The grammaticalization of passive morphology. *Studies in language* 14:25-72.
- Heine, Bernd, Ulriki Claudi and Friederike Hunnemeyer. 1991. *Grammaticalization: a conceptual framework*. Chicago: University of Chicago Press.
- Ikegami, Yoshihiko. 1981. *Suru to naru no gengogaku*. Tokyo: Taisyukan.
- Janda, Laura A. 1993. A geography of case semantics: the Czech dative and the Russian instrumental. Berlin: Mouton de Gruyter.
- Keenan, Edward L. 1985. Passive in the world's languages. *Language typology and syntactic description*, ed. by Timothy Shopen. Cambridge: Cambridge University Press. 243-81.
- Kemmer, Suzanne and Arie Verhagen. 1994. The grammar of causatives and the conceptual structure of events. *Cognitive linguistics* 5:115-56.
- Lakoff, George. 1977. Linguistic gestalts. *CLS* 13:236-87.
- _____. 1993. *The contemporary theory of metaphor*. Metaphor and thought, 2nd edn., ed. by Andrew Ortony. Cambridge: Cambridge University Press.
- _____. and Mark Johnson. 1980. *Metaphors we live by*. Chicago: University of Chicago Press.
- Li, Fengxiang. 1991. An examination of causative morphology from a cross-linguistic and diachronic perspective. *CLS* 27:344-59.
- Malchukov, Andrej L. 1993. The syntax and semantics of adversative constructions in Even. *Gengo kenkyuu* 103:1-36.
- Matisoff, James. 1991. Areal and universal dimensions of grammaticalization in Lahu. *Approaches to grammaticalization*, ed. by Elizabeth Traugott and Bernd Heine. Amsterdam: John Benjamins. 383-453.
- Moreno, Juan Carlos. 1993. 'Make' and the semantic origins of causativity: a typological study. *Causatives and transitivity*, ed. by Bernard Comrie and Maria Polinsky. Amsterdam: John Benjamins. 155-64.
- Newman, John. 1996. *Give: A cognitive linguistic study*. Berlin: Mouton de Gruyter.

- Oehrle, Richard T., and Hiroko Nishio. 1981. Adversity. *Coyote papers: Working papers in linguistics from A → Z* 2:163-85.
- Pederson, Eric. 1991. Subtle semantics: Universals in the polysemy of reflexive and causative constructions. UC Berkeley dissertation.
- Radetzky, Paula Kadose, and Tomoko Yamashita Smith. to appear. A unified account of the Japanese causative, moraw-benefactive, and passive constructions. *BLS* 23.
- Shibatani, Masayoshi. 1973. A linguistic study of causative constructions. UC Berkeley dissertation.
- _____. 1985. Passives and related constructions: a prototype analysis. *Language* 61:821-48.
- _____. 1994. An integrational approach to possessor raising, ethical datives, and adversative passives. *BLS* 20:461-85.
- _____. 1996. Applicatives and Benefactives: A cognitive account. *Grammatical constructions: Their form and meaning*, ed. by Masayoshi Shibatani and Sandra Thompson. Oxford: Clarendon Press. 157-94.
- Siewierska, Anna. 1984. *The Passive: A comparative linguistic analysis*. London: Croom Helm.
- Smith, Tomoko Yamashita. to appear. Benefactives, adversatives, causatives, and passives. *KLS* 18.
- Song, Jae Jung. 1996. Causatives and causation: a universal-typological perspective. New York: Longman.
- Svorou, Soteria. 1993. *The grammar of Space*. Amsterdam: John Benjamins.
- Sweetser, Eve. 1988. Grammaticalization and semantic bleaching. *BLS* 14:389-405.
- _____. 1990. From etymology to pragmatics: the mind-as-body metaphor in semantic structure and semantic change. Cambridge: Cambridge University Press.
- Talmy, Leonard. 1988. Force dynamics in language and cognition. *Cognitive science* 12:49-100.
- Tiee, Henry Hung-Yeh. 1986. *A reference grammar of Chinese sentences with exercises*. Tucson: The University Arizona Press.
- Traugott, Elizabeth Closs. 1975. Spacial expressions of tense and temporal sequencing; a contribution to the study of semantic fields. *Semiotica* 15.3:207-30.
- Tuggy, David. 1988. Náhuatl causative/applicatives in cognitive grammar. *Topics in cognitive linguistics*, ed. by Brygida Rudzka-Ostyn. Amsterdam: John Benjamins. 587-618.
- Washio, Ryuichi. 1993. When causatives mean passive: a cross-linguistic perspective. *Journal of East Asian linguistics* 2:45-90.
- Wierzbicka, Anna. 1988. *The semantics of grammar*. Amsterdam: John Benjamins.

Evidence for the Distinction between Resultative and Consequential Serial Verbs*

Osamuyimen Thompson Stewart
McGill University

0. Background

The general assumption within generative syntax is that serial verb constructions (SVCs) (1, 2) are distinct from covert coordinations (parataxis) (3) (cf. Déchaine 1993, Baker 1989, Carstens 1988, Collins 1997, Stewart 1996, etc.).

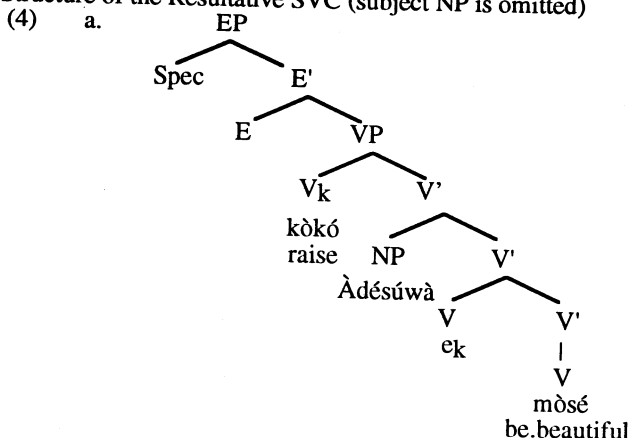
- (1) a. Òzọ kòkó àdésúwà mǝsé.
Ozo raise Adesuwa be.beautiful
'Ozo raised Adesuwa to be beautiful.'
b. Òzọ sùá úyì dé.
Ozo push Uyi fall
'Ozo pushed Uyi down.'
- (2) a. Òzọ lé èvbàré ré.
Ozo cook food eat
'Ozo cooked the food and ate it.'
b. Òzọ dé èbé tié.
Ozo buy book read
Ozo bought a book and read it.'
- (3) a. Òzọ gbǝǝ ívìn bòlò ọkà.
Ozo plant coconut peel corn
'Ozo planted coconut and [he] peeled corn.'
b. Òzọ lé ízè ríí ọrè.
Ozo cook rice eat it
'Ozo cooked rice and [he] ate it.'

One basic difference between the sentences in (1, 2) and those in (3) is the fact that there is a single surface object that is assigned the internal thematic roles of both verbs in the former, while each verb has a separate (overt) object in the latter (cf. Déchaine 1993). This phenomenon in which two or more verbs share a single surface object is usually referred to as object sharing, which has been generalized as an internal argument sharing criterion that defines SVCs (cf. Déchaine 1993, Baker 1989, Collins 1997, etc.).

However, there are different views on how best to account for this internal argument sharing condition. For example, Baker (1989, 1991) posits a double-headed VP structure whereby object sharing is derived from a phrase structure condition: the projection principle. Under this analysis, SVCs are assumed to have true internal argument sharing in which the verbs directly theta-mark a single object NP position without an empty category. In contrast, Déchaine (1993) Collins (1997), Stewart (1996), etc. all assume that each verb heads a distinct phrasal projection. Specifically, Collins (1997) posits that object sharing is mediated by an empty category, *pro*. Under this approach, there is actually no true internal object sharing as in Baker (1989).

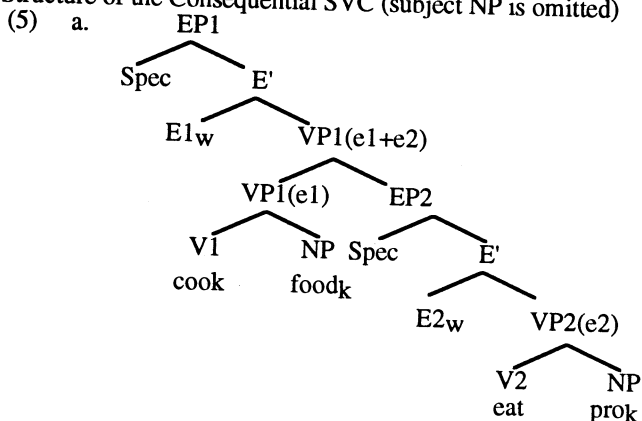
Against this background, this paper explores the basic problem with all of the previous analyses of SVCs by challenging the spurious unification of 1 and 2 as a single class of object sharing SVCs to which a unified analysis is often provided. Based on empirical evidence from Edó language (a Kwa language spoken in Nigeria), I will argue that there are two kinds of SVCs with distinct syntactic structures: (a) Resultative SVCs such as those in (1) have the structure and defining characteristics in (4), (b) Consequential SVC exemplified by (2) have the typical syntactic features that are summarized in (5).¹

Structure of the Resultative SVC (subject NP is omitted)



- b. The second verb is always unaccusative.²
- c. There is a single object NP with no empty category (true object sharing).
- d. The two verbs form a co-headed VP (both verbs are non-distinct).
- e. Both verbs express a single event that is existentially quantified-over by a single event operator, head of EP (Event Phrase).

Structure of the Consequential SVC (subject NP is omitted)



- b. Both verbs must be transitive.
- c. Each verb heads a distinct VP and expresses unique events (e1, e2).
- d. Each event is licensed by a separate event operator, head of EP.
- e. The two (functional) E heads are asymmetric (E1 quantifies over the two events (e1 + e2) and binds E2).
- f. Object sharing is mediated by an empty category, *pro*.

The initial evidence for the distinction between resultative and consequential SVCs comes from observations about verb restrictions and their interpretations discussed in Section 1. Section 2 examines different forms of adverbial modification that are used to introduce some of the fine-grained differences regarding the relations between the verbs with respect to object sharing and the nature of events that the verbs express in the different kinds of SVCs. Section 3 concludes this paper by using the results from the previous sections to show that differences in predicate clefts also correlate with the distinction between resultative and consequential SVCs.

1. Verb restrictions and interpretations

SVCs are generally described as constructions in which two or more verbs occur within a single clause without any form of subordination or coordination. According to Stewart (1963), there are restrictions on the verbal combinations. Therefore, this section examines the issue of restrictions on verb sequencing based on verb type in SVCs (cf. Stewart 1963, Baker 1989, 1991, Campbell 1996, Collins 1997, etc.). For present purposes, I assume that SVCs are minimally composed of two verbs and so the relevant questions to be addressed are: what sort of restrictions regulate verb serialization? Under what interpretations do such restrictions hold? Consider the following:

- (6) a. *Òzó gbé àdésúwà khòṣ.
Ozo beat Adesuwa fierce
'Ozo beat Adesuwa and She (Adesuwa) became fierce.'
- b. *Òzó sùá úyì só.
Ozo push Uyi shout
'Ozo pushed Uyi and he (Uyi) shouted.'
- (7) a. Òzó gbé àdésúwà wú. b. Òzó sùá úyì dé.
Ozo beat Adesuwa die Ozo push Uyi fall
'Ozo beat Adesuwa to death.' 'Ozo pushed Uyi down.'
- (8) a. *Òzó bọ́ọ̀ isókèṅ vbié.
Ozo comfort Isoken sleep
'Ozo comforted Isoken and she slept'
- b. *Òzó dé ẹ́mí!lá khián.
Ozo buy cow walk
'Ozo bought a cow and walked it.'
- (9) a. Òzó bọ́ọ̀ isókèṅ khú --- làdián.
Ozo comfort Isoken pursue out
'Ozo comforted Isoken and drove her out.'

- b. Òzọ dẹ ẹmí!lá kien ---
 Ozo buy cow sell
 'Ozo bought a cow and sold it.'

(6) shows that an unergative verb cannot occur as the second verb in SVCs that express a resultative meaning. A resultative SVC is one in which the second verb expresses the result of the action of the first verb based on an action-result relation. For example, in (6a) the intended meaning is that Ozo's (constant) beating of Adesuwa made her (Adesuwa) become fierce, and this sentence is ungrammatical as a SVC. Furthermore, the sentences in (6) also show that an unergative verb cannot occur in the second position in an object sharing SVC such as a resultative. For example, (6b) cannot have the interpretation in which the subject of the unergative verb 'so' (shout) is the object of the transitive first verb 'sùá' (push). However, if we replace the unergative verbs in (6) with unaccusatives as in (7), then the sentences become grammatical as both object sharing and resultative readings are now permitted. For example, (7a) has the reading in which Ozo beat Adesuwa to death, where the unaccusative second verb expresses the result of the action of the first verb. The generalization that emerges from this contrast is that in SVCs expressing result where the first verb is transitive, the second verb is typically unaccusative, but never unergative. This accounts for the ungrammaticality of (8a) but not (8b) and (9) where the verbs express sequences of events based on a precedence-consequence (consequential) relation rather than action-result (resultative).

I propose that (8b) and (9) illustrate a distinct class of object sharing SVC: the consequential SVC. Just as with SVCs expressing resultative meaning, those that express sequential precedence-consequence relation also exhibit verb sequencing constraints. Thus, it is ungrammatical for an unergative verb to occur in the second verb position. For example, (8b) cannot have the object sharing consequential SVC reading where Ozo bought a cow and walked it (the cow). We can generalize, therefore, that unergative verbs never occur in the second verb position of object sharing SVCs. On the other hand, when a transitive verb occur in similar position, the sentences are grammatical as shown in (9a,b). Consequently, while the second verb of the resultative SVC must be unaccusative, that of the consequential SVC must be transitive. I turn very briefly to the first verb position.

As it turns out, the transitive verb constraint extends beyond the second verb in the consequential SVC. It is important to note that whereas the first verb in SVCs expressing result could be of any verb type (transitive as in (7), unaccusative as in (10a), and possibly unergative in sequences such as bird fly exit), only transitives are allowed in SVCs expressing sequential actions. These facts regarding the restriction on both the first and second verbs in object sharing SVCs reflect the underlying difference between the resultative and the consequential: the verb positions in resultative SVCs are rigidly defined aspectually to express an accomplishment so they must contain a process first verb part (transitive, unaccusative, or unergative) and a result second verb part (unaccusative). On the other hand, there are no rigid aspectual constraints to be observed in consequential SVCs. Rather, there is only the syntactic constraint that the verbs match in transitivity values. The immediate consequence of these observations about verb sequencing and interpretations is the conclusion that resultative SVCs must be composed of two verbs (maximally and minimally) as shown in 10, while consequential SVCs may allow stacking of transitive verbs as illustrated in 11.

- (10) a. àkhé ghuòghò kànmwán.
 pot break be.short
 'The pot broke into small pieces.'
- b. *Òzó gbé àkhé ghuòghò kànmwán.
 Ozo hit pot break be.short
 'Ozo broke the pot into small pieces.'
- (11) a. Òzó dé iyán lé.
 Ozo buy yam cook
 'Ozo bought the yam and cooked it.'
- b. Òzó dé iyán lé - ré -
 Ozo buy yam cook - eat -
 'Ozo bought the yam, cooked and ate it.'

Thus, we conclude based on the contrasts from verb sequencing and interpretation that resultative SVCs are different from consequential SVCs both only similar in that they are constrained by an object sharing condition which rules out unergatives from occurring in the second verb position.

2. Adverbial Modification

2.1. Object sharing and empty category

This section provides evidence that reconciles the divergent views between Baker (1989) and Collins (1997) on the unified analysis of object sharing in SVCs. This is based on the behavior of an adverbial particle 'tòbòrè' which has the properties of an anaphor, and whose distribution is consistent with my proposed distinction between resultative and consequential SVCs. Consider the following:

- (12) a. Àdésúwà_k dé t_k tòbòrè_k
 'Adesuwa fell, herself (alone).'
- b. Òzó gbé ékítà_k nè úyì tá wẹẹ ìsòkẹn hòó *pro*_k tòbòrè_k
 Ozo hit dog that Uyi say that Isoken want *pro* itself
 'Ozo beat the dog that Uyi said that Isoken wants (itself).'

I adopt a general analysis of anaphors based on the simplest assumption about Condition A of the binding theory (Chomsky 1981) that requires it to be locally bound within the governing category. In particular, I assume a strict locality condition such that 'tòbòrè' is licensed as a right-adjunct to an NP that is its antecedent. Thus, in (12a) 'tòbòrè' can occur after the unaccusative verb in a simple sentence and ambiguously takes either the subject or the trace of the object of an unaccusative verb as its antecedent. Similarly, in (12b) we observe that 'tòbòrè' can also be licensed by an empty category, *pro*, that is the object of the transitive verb in the embedded clause. Conclusively, we observe that the distribution of 'tòbòrè' can be used to reveal the position of an otherwise null NP. Against this background, let us now examine its distribution in the two kinds of SVCs:

- (13) a. *Òzó sùá ògók dé pro_k tòbòrè_k
 Ozo push bottle fall itself
 b. Òzó dé iyánk dùnmwún pro_k tòbòrè_k
 Ozo buy yam pound pro itself
 'Ozo bought the yam and pounded it (itself).'

(13a) shows that it is ungrammatical for 'tòbòrè' to occur after the unaccusative verb in the resultative SVC. This contrasts with the simple sentence in (12a) where we observe that the particle can occur after an unaccusative verb. This contrast implies that there is no empty category (trace or *pro*) in resultative SVCs. This conclusion is consistent with my proposal that the two verbs form a syntactic compound and assign their internal theta roles (object sharing) to a single object NP, i.e. true internal object sharing that does not involve any empty category. Contrastively, in consequential SVCs where the second verb must be transitive we observe that the particle can occur after the verb (13b). This implies that there is an empty category, *pro*, to which the particle right-adjoins.

In conclusion, we note that the licensing and distribution of the adverbial particle 'tòbòrè' provides clear and consistent evidence for the distinction between resultative and consequential SVCs: the former involves sharing of a single NP, while object sharing in the latter is mediated by an empty category serving as the null object of the (transitive) second verb.

2.2. The distribution of manner adverb

This section deals with the interaction between adverbial modification and relations between the verbs in the different kinds of SVCs, both in terms of syntactic structure and the representation of the event or events that they express. One type of manner adverb in Èdó is illustrated in (14) (cf. Stewart 1996).

- (14) a. Òzó giégìé kòkò ògó (*giégìé). non-past tense
 Ozo quickly gather bottle (*quickly)
 'Ozo is quickly gathering the bottles (*quickly).'
 b. Òzó gié!gié kó!kó ògó (*gié!gié). past tense
 Ozo quickly gather bottle (*quickly)
 'Ozo quickly gathered the bottles (*quickly).'

The adverb in (14) can only occur in the INFL position between the subject and the verb, i.e. to the left of a verb, but never in sentence final position. Observe further that the adverb in (14) may vary for tense as realized by the different tones in (14a, b). For descriptive clarity, I propose to classify this adverb as an INFL-type adverb (henceforth I-type adverb) because it exhibits signs of being linked with tense both by its position and by varying for tense tones like verbs. The generalization arising from this description is that there are syntactic conditions that regulate the distribution of the I-type adverb. It can only occur as an adjunct to a syntactic position that is tense-related, i.e., a syntactic position where tense features can be checked. Consequently, I propose that the I-type adverb is licensed as a left-adjunct to the head of EP which is a tense-related functional projection (cf. Travis 1994). Furthermore, I follow Parsons (1990) and assume an account in which adverbs are predicates of events that are denoted by verbs. This is based on Davidson (1967), who proposes that verbs explicitly stand for kinds of events, so

that a sentence containing such a verb states implicitly that there is an event of that sort. Therefore, the presence of an adverb indicates that there is an event of which it is predicated, i.e., an adverb modifies (is predicated of) an event that is denoted by the verb. Given these assumptions about the licensing of I-type adverb, let us now examine what they can tell us about SVCs.

2.2.1 I-type adverb: Events and functional structure

The section illustrates the structural prediction that the distribution of the I-type adverb will provide evidence for functional structure, i.e. EP. The assumption is that since the EP is taken to represent an event in the syntax, the distribution of I-type adverb will match up with the number of EPs to reflect one or two events with respect to the distinction between resultative and consequential SVCs. There are two possible positions for the adverb: before both verbs (pre-first verb) and before the second verb. I begin with the pre-first verb position:

- (15) a Òzọ giẹ!giẹ kó!kó Àdésúwà mọ!sé
 Ozo quickly raise Adesuwa be-beautiful
 'Ozo quickly raised Adesuwa to be beautiful'
- b. Òzọ giẹ!giẹ dú!nmwún èmà khién
 Ozo quickly pound yam sell
 'Ozo quickly pounded the yam and sold it'

(15) shows that the I-type adverb can occur in the position before the verbs in both kinds of SVCs. This is compatible with the interpretation of the adverbs. For example, in both examples in (15) the adverb modifies the actions denoted by both verbs necessarily. Thus, (15a) has the reading that the raising and the becoming beautiful were quick, while (15b) has the reading that both the pounding and the selling were quick. I conclude, based on Parsons (1990), that there must be a sense in which the verbs in (15) express a single event of which the I-type adverb is a predicate. This implies that there is always a projection of EP that dominates the verbs in SVCs. Now let us examine the possibility of an I-type adverb before the second verb which is more interesting because we will be able to tease apart a double-headed VP structure from that involving separate VP projections.

- (16) a *Òzọ sùá Àdésúwà giẹ!giẹ dé.
 Ozo push Adesuwa quickly fall
 'Ozo pushed Adesuwa down quickly.'
- b. Òzọ dùnmwún èmà giẹ!giẹ khién
 Ozo pound yam quickly sell
 'Ozo pounded the yam and quickly sold it'

(16) shows a clear contrast between resultative and consequential SVCs. In the resultative SVC (16a), we observe that an I-type adverb cannot occur before second verb in contrast to the consequential SVC (16b). (16b) has the interpretation that Ozo pounded the yam (perhaps rather slowly) and quickly sold it. Now, the ungrammaticality of (16a) suggests that there is no EP dominating the second verb in the resultative SVC. In contrast, (16b) implies that consequential SVCs have an EP before the second verb. The generalization from this in terms of meaning is that although both kinds of SVCs may express a single event based on (16), they differ, however, with respect to the internal composition of the single event. This is

consistent with the previous generalization in Section 1 that resultative SVCs express an accomplishment, i.e. the meaning/function of each verb is determined by the event ontology of an accomplishment. Consequential SVCs on the other hand are composed as a single 'complex' event without any rigid aspectual conditions and each verb express a unique event which are then linked together.

2.3. Iterative morpheme and event composition

This section provides further evidence in support of the distinction between the two kinds of SVCs based on the distribution and interpretation of the iterative morpheme *ghá*. This will also shed some light on the exact nature of the licensing of the complex single event in the consequential SVC. One quirky fact in the Èdó language is that the iterative morpheme requires the presence of an I-type adverb to occur in the higher EP position. Now, consider the following:

- (17) a. Òzó fẹ̀!kó ghá sùá ọ̀gọ́ dé.
 Ozo slowly Iter. push bottle fall
 'Ozo carefully pushed the bottle down repeatedly.'
- b. *Òzó sùá ọ̀gọ́ (fẹ̀!kó) ghá dé.
 Ozo push bottle slowly Iter. fall
- (18) a. Òzó fẹ̀!kó ghá dún!mwún èmà khiến. = a whole event (e1+e2)
 Ozo slowly Iter. pound yam sell
 'Ozo pounded the yams and sold it repeatedly.'
- b. Òzó dúnmwún èmà fẹ̀!kó ghá khiến. = part event, only (e2)
 Ozo pound yam slowly Iter. sell
 'Ozo pounded the yams and repeatedly sold it.'

Following Jackendoff (1990) I assume that iteration applies to events and so I propose an analysis for the iterative morpheme by base-generating it in the head of the functional projection of EP. In (17a) and (18a), we observe that the iterative morpheme *ghá* can occur before the two verbs, modifying both of them. For example, (17a) means that Ozo pushed the bottle down again and again (push-fall is iterated). Similarly, (18a) implies that Ozo pounded the yam and sold it again and again (pounding and selling are iterated). Therefore, we generalize that the operator in the head of the (higher) EP obligatorily takes scope over the event(s) which the verbs express in both SVCs. However, the iterative morpheme cannot occur before the second verb in the resultative SVC (17b), while (18b) shows that this is possible in the consequential SVC. This difference between (17b) and (18b) is consistent with the proposal that there is no EP projection between the verbs in resultative SVCs compared with consequential SVCs. Therefore, resultative SVCs have a representation as in (19).

- (19) $(\exists e) [\text{Push-Fall}(e) \ \& \ \text{Agent}(e, \text{Ozo}) \ \& \ \text{Theme}(e, \text{ọ̀gọ́})]$.

(19) illustrates the fact that resultative SVCs express a single event and that the verbs are non-distinct from each other; they form some sort of syntactic compound as represented in the structure in (4). Thus, the ungrammaticality of (17c) is consistent with the fact that we would not expect to iterate one part of a resultative (accomplishment). However, nothing prevents this sort of discrete modification in

the consequential SVC that is made up of unique events. The question, then, is what licenses the complex single event in the consequential SVC. Consider the following:

- (20) a. *Òzó fẹ̀!kó ghá dún!mwún èmà ghá khién.
 Ozo slowly Iter. pound yam Iter. sell
 'Ozo repeatedly pounded the yams and repeatedly sold it.'
- b. *Òzó fẹ̀!kó dún!mwún èmà fẹ̀!kó khién.
 Ozo slowly pound yam slowly sell
 'Ozo carefully pounded the yams and carefully sold it.'
- c. Òzó gẹ̀llé dún!mwún èmà fẹ̀!kó khién.
 Ozo truly pound yam slowly sell
 'Ozo truly pounded the yams and carefully sold it.'

(20a) shows what happens when the heads of the two EPs are simultaneously filled by tokens of the same thing, in this case the iterative morpheme. The ungrammaticality of (20a) is unexpected under an analysis in which the events in the consequential SVCs are in a conjoined structure. However, (20b) provides evidence that there is an asymmetric relation between the EPs and, by implication, the events that they represent. This is based on the simultaneous distribution of the same I-type adverb which is ungrammatical (20a, b). Crucially, however, when there are different tokens of the manner adverb as in (20c), the sentence is grammatical. Thus, (20c) shows that what is relevant is not the fact that the head of two EPs cannot be filled, nor is it a problem of adverb hierarchy as in Cinque (1997), whereby adverbs are ordered with respect to each other. If the latter approach were correct, then we would expect that the reversal of the adverbs in (20c) would be ungrammatical which is contrary to fact as illustrated in (21).

- (21) Òzó fẹ̀!kó dún!mwún èmà gẹ̀llé khién.
 Ozo slowly pound yam truly sell
 'Ozo carefully pounded the yams and truly sold it.'

Consequently, I propose that the ungrammaticality of (20a, b) provides the vital clue regarding the internal structure of the complex single event. This can be derived from a parallel problem in Formal Semantics as shown in (22).

- (22) $\forall x (Fx \text{ -----} \rightarrow \exists x G(x))$ Formal Semantics parallel

According to (22), there are two quantifications present and both target the same variable $G(x)$. However, it is only the closest (lower) quantifier that counts as the binder for $G(x)$. I adopt a similar approach to the ungrammaticality of (20a, b) and propose a parallel representation as shown in (23) for the consequential SVC.

- (23) $\exists x (Fx \text{ -----} \rightarrow \exists x G(x))$ consequential SVC

In simple terms, (23) expresses the fact that there are two quantifications over one complex event, one which quantifies over the entire event and the other which is restricted in scope to the lower (part of the) event. Therefore, a consequential SVC like (2a) will have the formal representation in (24).

- (24) $\exists_E (\exists e_1 \exists e_2 [\text{Buying}(e_1) \ \& \ \text{Agt}(e_1, \text{Ozo}) \ \& \ \text{Th}(e_1, \text{èbè})]$
 $\& [\text{reading}(e_2) \ \& \ \text{Agt}(e_2, \text{Ozo}) \ \& \ \text{Th}(e_2, \text{èbè})]$
 $\& [E \text{ "consists of" } (e_1, e_2)]$

(24) implies that there is a single complex event in the consequential SVC that is made up of two events (e_1) and (e_2) expressed by each of the verbs. Both events are thus linked together by the fact that the first EP obligatory takes scope over the two verbs ($e_1 + e_2$), while the lower EP takes scopes over only the lower verb. These are realized structurally via an adjunction structure where the lower EP adjoins to the first VP as illustrated in (5), rather than a conjoined structure.

In conclusion, we see that the distribution and licensing of I-type adverb and iterative morphemes distinguishes single event resultative SVCs from two-event consequential SVCs.

3. Predicate clefts and event(s)

This section shows that the distinction between resultative and consequential SVCs and the analysis proposed have significant consequences for the interaction between predicate clefts and the representation of event(s). The phenomenon of predicate cleft is a way of focusing a verb that involves moving a category that is associated with it (cf. Koopman 1984, Lefebvre 1994, Manfredi 1993, etc.). Consider the following:

- (25) a. $\text{Òzó sùá} \quad \text{Àdésúwà.}$
 'Ozo pushed Adesuwa.'
- b. $\text{ùsuámwèn} \quad \text{òré} \text{ Òzó sùá} \quad \text{Àdésúwà.}$
 nom-push-nom Foc. Ozo pushed Adesuwa.'
 'It is pushing that Ozo pushed Adesuwa'
- (26) a. $\text{Àdésúwà dé. b.} \quad \text{ùdémwèn} \quad \text{òré} \text{ Àdésúwà dé.}$
 Adesuwa fall nom-fall-nom Foc. Ade. fall
 'Adesuwa fell.' 'It is falling that Adesuwa fell.'

As (25b) and (26b) show, predicate cleft is the movement to sentence initial position of some item that is morphologically cognate to the verb. This applies to all types of verbs (compare the transitive verb in (28b) and the unaccusative verb in (29b)). I assume that the morphologically cognate item that is moved in predicate cleft is the nominal argument of the event denoted by that verb (cf. Bamgbose 1972, Manfredi 1993, Lefebvre 1994). I propose an analysis in which the nominal argument of an event is base-generated as a complement within the VP. However, when predicate cleft applies this event argument must move at LF through the Specifier of EP, after the relevant verb has been raised into the functional head, in order for it to be licensed under Spec-head. In this way, I equate the possibility of predicate cleft with the presence of EP. Crucially, I assume that there is no adjunction to Specifier. Thus, both (25b) and (26b) contain a single EP projection and predicate cleft is properly licensed under Spec-head by the LF raising of the event argument and the verb which denotes such an event. Now, consider predicate cleft possibilities from resultative and consequential SVCs respectively:

- (27) a. *ùsuámwèn òré Òzó sùá Àdésúwà dé
 nom-push-nom Foc. Ozo push Adesuwa fall
 b. *ùdémwèn òré Òzó sùá Àdésúwà dé
 nom-fall-nom Foc. Ozo push Adesuwa fall
- (28) a. ùlémwèn òré Òzó lé èvbàré ré
 nom-cook-nom Foc. Ozo cook food eat
 'It is cooking that Ozo cooked the food and ate, (not shredding it).'
- b. ùrémwèn òré Òzó lé èvbàré ré
 nom-eat-nom Foc. Ozo cook food eat
 'It is eating that Ozo cooked the food and did, (not selling it).'

(27a, b) shows that it is ungrammatical to cleft either of the verbs from the resultative SVC, while similar clefts from the consequential SVC are grammatical (28a, b). The generalization that emerges from this contrast is that predicate cleft is constrained in a single-event (resultative) SVC but possible in two-event consequential SVC. Based on my analysis, this difference between the two kinds of SVCs comes from the fact that there is a single EP in the resultative SVC, while there are two in the consequential SVC. The two verbs in the resultative SVC form a syntactic compound that act as a unit, therefore when the event argument of one of the verbs raise to Spec EP at LF there can be no appropriate Spec-head matching since there are two verbs that must have raised to the functional head and so one of the verbs will fail to license its event argument. Consequently, the predicate cleft of either verbs in the single-event resultative SVC will be ungrammatical as shown in (27a,b). On the contrary, there are two EPs in the two-event consequential SVC and so it is possible to cleft either of the verbs since the event argument will be properly licensed under Spec-head in the different EPs. Thus, once again we observe that resultative SVCs pattern differently from consequential SVCs.

4. Conclusion

In this paper, I have presented clear empirical evidence from Èdó which distinguish resultative SVCs from consequential SVCs. Several theoretical consequences are derived from this distinction including the unaccusative vs. transitive verb contrast, the nature of object sharing, and the representation of event(s) denoted by the verbs. Based on these facts, two distinct syntactic structures are proposed, contrary to the generally assumed unified analysis.

Notes

* I thank Mark Baker, Lisa Travis, Sam Mchombo, BLS audience, and Alma Mater student travel grant from McGill University for various contributions to this paper. The usual disclaimer applies.

¹ Structurally, I omit the functional projections TP and VoiceP simply because the subject NP is not directly relevant to the issues raised in this paper, however see Stewart (1998) for discussion.

¹ This includes stative verbs. See Baker and Stewart (1997)

References

- Baker, Mark. 1989. Object sharing and projection in serial verb construction. *Linguistic Inquiry* 20, #4: 513-553.
- Baker, Mark. 1991. On the relation of serialization to verb extensions. In *Serial verbs: Grammatical, comparative, and cognitive approaches*, ed. by Claire Lefebvre, 79-102. Amsterdam: John Benjamins.
- Baker, Mark, and Osamuyimen T. Stewart. 1997. Unaccusativity and the Adjective/Verb Distinction: Edo Evidence. In *Proceedings of NELS 27*, 33-47, GLSA, McGill University.
- Bamgbose, Ayo. 1972. (ed.) *The Yoruba verb phrase*. Ibadan: IUP.
- Campbell, Richard. 1996. Serial verbs and shared arguments. *The Linguistic Review* 13, 83-118.
- Carstens, Vicki. 1988. Serial verbs in Yorùba. Paper presented at the 2nd workshop in Niger-Congo syntax and semantics. Cambridge, Mass: MIT Press.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Cinque, Guglielmo. 1997. *Adverbs and functional heads: a cross-linguistic perspective*. New York: Oxford.
- Collins, Chris. 1997. Argument sharing in serial verb construction. *Linguistic Inquiry*, 28: 461-497.
- Davidson, David. 1967. The logical form of action sentences. In *The logic of decision and action*, ed. N. Rescher. Pittsburgh.
- Déchine, Rose-Marie. 1993. *Predicates Across Categories: Towards a Category-Neutral Syntax*. Ph.D Dissertation, University of Massachusetts, Amherst.
- Jackendoff, Ray. 1990. *Semantic structures*. Cambridge, MA: MIT Press.
- Koopman, Hilda. 1984. *The syntax of verbs: From verb movement in Kru languages to Universal grammar*. Dordrecht: Foris.
- Lefebvre, Claire. 1994. On spelling out E. In *Travaux de recherche sur le créole haitien*. No 23 : 1-29. Université du Québec à Montréal.
- Manfredi, Victor. 1993. Verb focus in the typology of Kwa/Kru and Haitien. In *Focus and grammatical relations*, ed. by Francis Byrne and D. Winford, 3-51. Amsterdam: John Benjamins.
- Parsons, Terence. 1990. *Events in the semantics of English: A study in subatomic semantics*. Cambridge, Mass.: MIT Press.
- Stewart, John. M. 1963. Some Restrictions on Objects in Twi. *Journal of African Languages* 2, #2, 145-149.
- Stewart, Osamuyimen. T. 1996. Adverb placement and the structure of the serial verb construction. *Proceedings of NELS 26*: 409-422 MIT/Harvard GLSA.
- Stewart, Osamuyimen. T. 1998. *Serial verb construction parameter*. Ph.D dissertation, McGill University.
- Travis, Lisa deMena. 1994. Event phrase and a theory of functional categories. *Proceedings of CLA*: 559-570.

Telescope Anaphora

Janine Toole
Simon Fraser University

1.0 Introduction

In this paper I am concerned with the analysis of a particular class of definite noun phrases known as 'telescope pronouns'. According to Poesio & Zucchi (1992: 347), a telescope pronoun is a singular pronoun that is anaphorically related to a universal quantifier in a previous clause. Some examples of telescope pronouns are given in (1). These examples are problematic because most theories of discourse semantics do not allow universal noun phrases to bind pronouns that are outside their scope. While this approach excludes the invalid examples in (2), it provides the wrong prediction for the telescope examples in (1).

- 1a. Each degree candidate_i walked to the stage. *He_i* took his diploma from the dean and returned to his_i seat.
- b. Each student_i in the syntax class was accused of cheating on the exam and *he_i* was reprimanded by the dean.
- c. Every story_i pleases these children. If *it_i* is about animals, they are excited, if *it_i* is about witches, they are enchanted, and if *it_i* is about humans, they never want me to stop.
- 2a. #If every cat_i purrs, *it_i* is happy
- b. #John likes every dog_i and Sam feeds *it_i*.
- c. #If John owes every man_i money then Sam pays *him_i*.

In this paper I deal with two issues concerning telescope anaphora. First, I provide a semantic analysis which avoids the problems found with previous analyses. Secondly, I consider the conditions under which a telescope interpretation is licensed. I assume an understanding of the basic constructs of Discourse Representation Theory (Kamp & Reyle 1993).

2.0 Previous Analyses of Telescope Anaphora

Previous attempts to account for telescope anaphora include the restrictor reconstruction approach of Roberts (1987, 1989) and Poesio & Zucchi (1992), the numberless pronoun approach of Neale (1990) and the scope extension approaches of Sells (1985) and Dekker (to appear).¹ In this section I introduce two problems with these previous analyses.²

Problem 1: Anaphora to an NP within the scope of the telescope antecedent

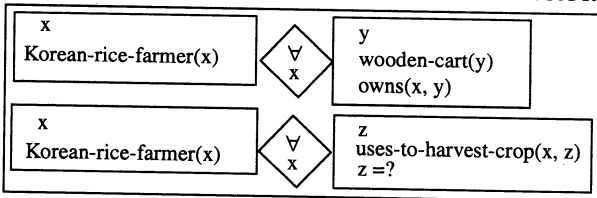
Both the numberless pronoun and restrictor reconstruction approaches provide incorrect analyses for examples that include two pronouns: one a telescope pronoun, and one a pronoun which is anaphoric on a noun phrase within the scope

of the telescope antecedent. Such an example is given in (3). In addition to the telescope pronoun *he*, there is a pronoun *it* which is anaphoric on the indefinite noun phrase *a wooden cart*.

3. Each Korean rice farmer_i owns a wooden cart_j. *He_i* uses *it_j* to harvest the crop.

Poesio & Zucchi's DRT-based restrictor reconstruction analysis of this example produces an invalid DRS, illustrated in (4). Their analysis cannot provide an accessible antecedent for the discourse referent *z* that is introduced by the pronoun *it*.

4. Every Korean rice farmer owns a wooden cart. He uses it to harvest the crop.



Neale's numberless pronoun approach is similarly problematic as it provides incorrect truth conditions for this example. On Neale's analysis (3) is true only if each farmer uses *every* wooden cart he owns to harvest the crop. However, native speakers I have consulted consider this sentence to be true as long as each farmer uses *at least one* wooden cart he owns to harvest the crop.

Problem 2: Possible antecedents for the telescope pronoun

The examples in (5) have the same structure as the example in (1) except that the subject quantifier is different. Yet, the examples in (5) are not acceptable telescope examples. This indicates that there are restrictions on the type of noun phrase that can be the antecedent of a telescope pronoun. Ideally, this restriction should fall out naturally from an analysis of telescope anaphora. However, this is not the case with previous approaches to the telescope construction, all of which need to *stipulate* that only quantificational determiners which combine with a singular noun phrase can be the antecedent of a telescope pronoun.

- 5a. Most degree candidates walked to the stage. #He took his diploma from the dean.
 b. All degree candidates walked to the stage. #He took his diploma from the dean.

3.0 An Alternative Approach

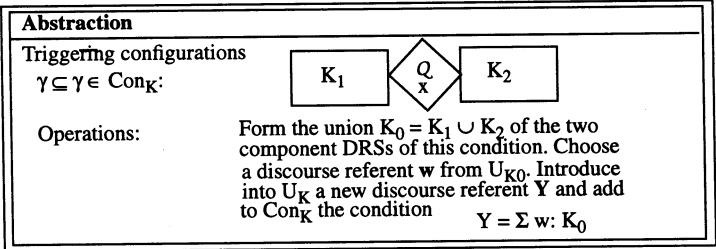
Various proposals have been made to account for the problems posed by telescope anaphora. Poesio & Zucchi accommodate an antecedent, Sells relocates the relevant conditions so that an antecedent is accessible, and Neale introduces a numberless operator that accesses predicates and variables from previous clauses. Each of these

analyses requires the addition of some special mechanism (accommodation/relocation/ambiguous pronoun) to account for the behavior of telescope noun phrases. I propose instead to account for the telescope construction using minimally modified versions of two constructs already introduced in Discourse Representation Theory. As a first step I introduce these constructs. Following this I illustrate how these constructs can be combined to account for telescope anaphora.

3.1 Technical Apparatus

The first construct I assume is the process of abstraction. On Kamp & Reyle’s interpretation, as given in the construction rule given in (6a), this process introduces a plural discourse referent that is derived in a specified way from a duplex condition already existing in the DRS. For example, given a duplex condition of the form found in (7a), abstraction introduces a plural discourse referent consisting of the set of men who sing. This is illustrated in (7b). The semantic interpretation rule for a condition of the form $Y = \Sigma w: K_0$ is given in (6b).

6a. Kamp & Reyle 1993: 344

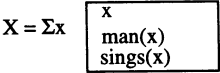


b. $M \models_f x = \Sigma zK \text{ iff } f(x) = \oplus_M \{b: b \in U_M \ \& \ M \models_f \cup \{<z,b>\} K\}.$

7a.



b.



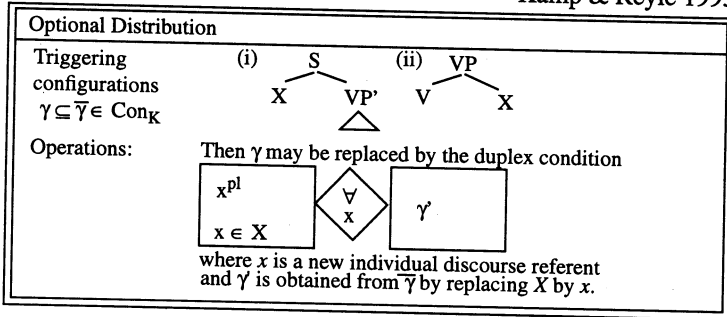
Implicit in Kamp & Reyle’s rule is the assumption that abstraction only applies when it is needed. That is, this process only applies when there is the possibility that the sentence may be uninterpretable because there is no antecedent for a noun phrase.

The second construct I assume is the process of optional distribution. This process is required in order to obtain the distributed reading of a sentence like (8). Kamp & Reyle’s default analysis of (8) produces the collective reading. On their approach, the distributive reading is obtained by applying an optional distribution rule. This rule is given in (9) and specifies that the discourse referent introduced by a plural noun phrase may be replaced by a duplex condition that quantifies over the individual members of the plural discourse referent.³

8. Several men bought a donkey.

9.

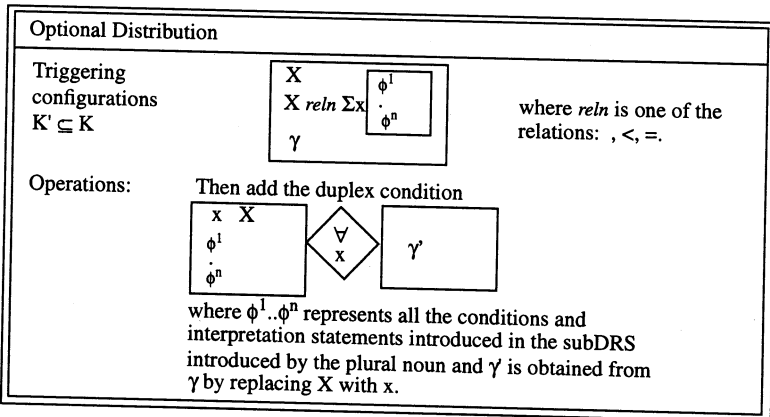
Kamp & Reyle 1993: 348



In order to be able to combine abstraction and distribution, I make two minimal modifications to Kamp & Reyle's rule. Firstly, the original rule is triggered by the existence of a plural discourse referent in a certain structural position. On the analysis I propose the distribution rule is triggered by the existence of an *abstraction condition*, like that introduced by the abstraction process or by a cardinality determiner.^{4, 5} This is illustrated in the triggering conditions in the modified distribution rule given in (10).

Secondly, I modify the format of the duplex condition that the rule introduces. On my analysis the restrictor DRS contains all the conditions found in the abstraction subDRS that licensed the rule. In addition, it contains a condition of the type $x \in X$, where x is the discourse referent that is summed over in the abstraction condition and X is the individual on the left hand side of the abstraction condition. These changes are found in the updated rule given in (10).

10.



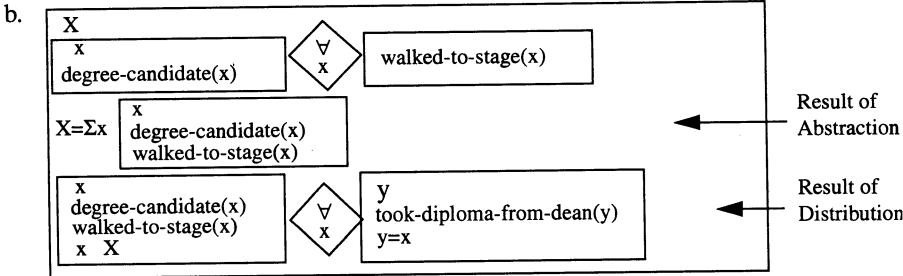
3.2 Accounting for the Telescope Construction

I now illustrate how the processes of abstraction and distribution can be used to account for telescope constructions. In section 3.1 it was noted that abstraction applies only when there is no other antecedent for a definite noun phrase. A telescope construction meets this condition since there is no apparent antecedent for the telescope pronoun.

However, applying abstraction does not solve the problems posed by the telescope example since the *plural* discourse referent introduced by the abstraction process is not a possible antecedent for the *singular* discourse referent introduced by the telescope pronoun. For example, although the process of abstraction in (11b) introduces a plural discourse referent *X* which consists of the set of degree candidates who walked to the stage, this discourse referent cannot be the antecedent for the singular discourse referent *y* (the discourse referent introduced by the telescope pronoun).

However, as a result of abstraction, the triggering conditions for the revised distribution rule are met. Applying the distribution rule produces a duplex condition that distributively quantifies over the individual members of the plural discourse referent. This provides a singular discourse referent that can be the antecedent for *y*. As can be seen from the third condition in (11b), the singular discourse referent introduced by the distribution rule is an appropriate antecedent for the telescope pronoun.

11a. Each degree candidate walked to the stage. He took his diploma from the dean.



The truth conditions for this DRS are met whenever each degree candidate both walked to the stage and took his diploma from the dean. These truth conditions match those that native speakers associate with the telescope example in (11a). Hence, by applying abstraction and distribution we have provided an antecedent for the telescope noun phrase and produced truth conditions that match native speaker intuitions.⁶

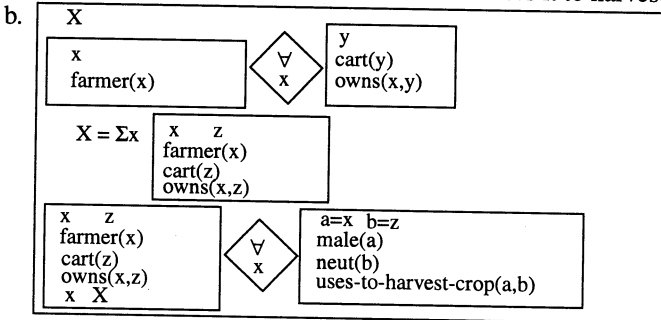
3.3 Accounting for Previous Problems

In the previous section I illustrated the means by which abstraction and distribution provide an appropriate antecedent for a telescope noun phrase. In this section I evaluate this analysis with respect to the two problems introduced in section 2.0.

Firstly, the abstraction/distribution approach avoids the problems previous analyses have with sentences that have two pronouns. The example discussed in section 2.0 is repeated in (12a). Recall, Poesio & Zucchi's analysis produces an invalid DRS for this example while Neale's analysis provides incorrect truth conditions: on Neale's approach (12a) is true only if each Korean rice farmer uses *every* cart he owns to harvest the crop.

The abstraction/distribution analysis of this sentence avoids this problem and produces a valid DRS that is true if each Korean rice farmer uses at least one of his carts to harvest the crop. The DRS for the discourse in (12a) is given in (12b) (abstraction and distribution have applied). This DRS has the same structure as (11b). The first condition is derived from the first sentence in (12a) in the standard manner. The second condition is the result of abstracting over the duplex condition. In turn, the existence of the abstraction condition licenses distribution. The final duplex condition is the result of applying distribution. This DRS is true if each Korean rice farmer uses at least one cart he owns to harvest the crop. Thus, the abstraction/distribution approach readily accounts for examples which include an additional pronoun anaphoric on a noun phrase within the subject relative clause.

12a. Each Korean rice farmer owns a cart. He uses it to harvest the crop.



The abstraction/distribution analysis also avoids much of the stipulation that was necessary in previous accounts. For example, previous analyses have to stipulate which noun phrases can be the antecedent of a telescope noun phrase and which cannot. Recall, in section 2 we saw that a noun phrase such as *each man* or *every man* can be a telescope antecedent while a noun phrase such as *most men* or *all men* cannot. The examples used to illustrate this point are found in (5).

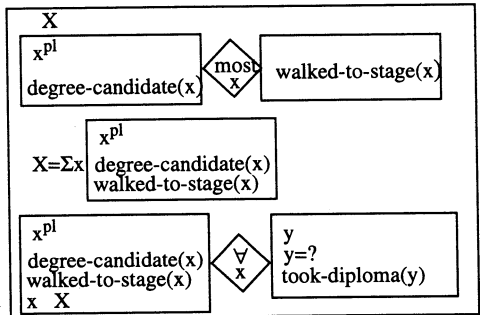
However, if Kamp & Reyle's (1993) approach to plural noun phrases is incorporated into the abstraction/distribution approach then it is not necessary to

make such stipulations about the antecedent of a telescope noun phrase. In order to account for the behavior of pronouns in sentences like (13), Kamp & Reyle mark an *individual* discourse referent that is introduced by a *plural* noun phrase with the addition of a superscript *pl* (1993: 347). A discourse referent with a *pl* superscript can only be the antecedent for a plural noun phrase. Hence, the sentence in (13a) with a plural pronoun is grammatical while the sentence in (13b) with a singular anaphoric pronoun is not.

- 13a. Few lawyers hired a secretary they liked.
 b t #Few lawyers_i hired a secretary he_i liked.

If we assume Kamp & Reyle's approach to plural noun phrases then it is not necessary to stipulate that only singular quantified noun phrases can be the antecedent of a telescope noun phrase. A proportionally quantified noun phrase such as *most men* cannot be a telescope antecedent because, even with abstraction and distribution, there is no accessible antecedent for the telescope noun phrase. For example, the DRS for the sentence in (5a) is given in (14). In this DRS abstraction and distribution have applied in an attempt to provide an accessible antecedent for the telescope noun phrase. However, the only potential antecedent is x^{pl} and this discourse referent cannot be the antecedent for the singular pronoun. Since there is no accessible antecedent, there is no valid interpretation for this example.

14.



In sum, the semantic interpretation of telescope constructions can be accounted for by combining the previously motivated constructs of abstraction and distribution. This approach provides the correct semantic interpretation for telescope examples and avoids the problems found with previous analyses.

4.0 Towards Licensing Conditions

In previous sections I have argued that a telescope interpretation is available when the processes of abstraction and distribution can apply to produce an appropriate antecedent for the telescope anaphor. While this accounts for the data we have discussed so far, this licensing condition is much too general because it allows abstraction (and subsequently distribution) in examples that do not have a telescope

interpretation. For example, the current analysis predicts that the example in (15) should have a telescope interpretation. However, the pronoun in this sentence cannot be interpreted as a telescope noun phrase.

15. #John likes every dog and it likes cheese.

Unfortunately, the development of licensing conditions is no easy matter. One issue that becomes apparent when dealing with telescope anaphora is that there is considerable variation among native speaker intuitions as to whether a particular example can be interpreted telescopically or not. Although the literature contains no mention of telescope intuition variation, among my colleagues are people who find the examples in (16) acceptable, which I do not. Similarly, I find the examples in (17) acceptable whereas others I have consulted do not.

- 16a. If every cat purrs, it is happy. [self-constructed]
 b t If each candidate for the space mission meets all our requirements then he has a Ph.D. in physics and extensive flight experience. [self-constructed]
- 17a. Each candidate for the space mission meets all our requirements. He has a Ph.D. in physics and extensive flight experience. [Poesio & Zucchi 1992]
 b t Every dog came in. It lay down under the table. [Poesio & Zucchi 1992]

A further confounding issue is the fact that intuitions are almost always obtained from examples in isolation. However, Poesio & Zucchi point out that the example in (17b) is often deemed more acceptable when it is put in a context, like (18), which emphasizes the sequential nature of the dogs' actions. Hence, the intuitions that informants are providing may depend to some extent on the ability of the informant to come up with a suitable context in which to distributively interpret the telescope noun phrase. The problem of intuition variation is exacerbated by the fact that all of the examples discussed in the literature are constructed examples.⁷

18. I went to the circus last night. They had a number involving dogs that went like this: The circus performers put a table on some supports. Then, every dog came in. It lay down under the table, stood on its back paws, and lifted the table with its front paws.

This variation poses a problem for the development of licensing conditions since no licensing conditions will account for all native speakers' intuitions (or even for the intuitions of the small set of people who publish on this topic). These problems led me to take an alternate approach. As an initial step towards developing licensing conditions I chose to review the characteristics of a collection of naturally occurring telescope examples.⁸ My premise being that features of these examples may provide clues to their licensing conditions.

Indeed, this review proved enlightening. Whereas there was considerable variation among informants in their judgements of the constructed examples, the judgements on the naturally occurring examples were unanimous: in every case, the informants judged the naturally occurring examples to be perfectly acceptable. In addition, the

characteristics of these sentences do provide some clues towards the factors which license a telescope reading. I make three observations below.

The Telescope NP

Firstly, a review of the telescope examples discussed in the literature gives the impression that it is only pronouns that can occur in a telescope construction. Indeed, Neale's analysis explicitly restricts this type of construction to pronouns. However, a study of naturally occurring examples indicates that other definite noun phrases can also occur in telescope constructions. This is illustrated in (19) where the telescope NP *the group* is anaphoric on the universally quantified subject noun phrase *each group*.

19. Each group has a leader, a project and a budget. The group is free to spend its budget with virtually no upper-management control, provided it stays within its allotment.

Verb Feature Parallelism

Secondly, the telescope examples that I reviewed exhibit parallel tense, aspect, and voice. I define an example as having parallel tense/aspect/voice if the main clause containing the antecedent and the main clause containing the telescope noun phrase have the same values. That is, they have the same tense, are both progressive or both non-progressive, both perfective or both non-perfective, and both active or both passive. This parallelism can be seen in the naturally occurring example in (20a). In contrast, the unacceptable example in (20b) does not have parallel tense.⁹

- 20a. Each group *has* a leader, a project and a budget. The group *is* free to spend its budget with virtually no upper-management control, provided it stays within its allotment.
b t Every rice farmer in Korea *owns* a wooden cart. #Usually, he *got* it from his father.

Most interestingly, it is the verb in the main clause that exhibits the parallelism, not the verb of the clause that contains the telescope anaphor/antecedent. This is illustrated by the examples in (21). The verbs in the main clauses are italicized and the verbs in the antecedent/anaphor clauses are underlined. The underlined verbs do not have the same features. The underlined verb in the antecedent clause in (21a) is in the subjunctive, while the underlined verb in the anaphor clause is not. In (21b) the underlined verbs differ in their tense feature. The antecedent verb is present tense while the anaphor verb is in the past tense. In contrast, the italicized verbs share the same features.

- 21a. Hungarian law, it seems, *requires* that every child have a recognized father. If he doesn't, the state simply *invents* one.
b t When a magnetic field is applied, each mobile electron *is forced* to spiral about the field's direction, although its energy is not changed. This additional angular motion imposed by the field *is quantised*, exactly as if the electron was constrained in an orbit around an atom.

Grammatical Function Parallelism

Finally, the naturally occurring examples also exhibit a certain amount of parallelism in the grammatical function of the anaphor and antecedent. If the antecedent is a subject, the telescope anaphor will be a subject. If the antecedent is the object of a preposition, the anaphor will be the object of a preposition, etc. These two specific examples are illustrated in (22).

- 22a. Under Mr. Harel's system, *each business* must operate as a free-standing unit. *It* must assign a value to each job, recognizing that a manager costs more than a laborer. [WSJ]
- b t In our tests, the hobs were dirtied by letting milk boil over on two burners of *each cooker*. All the milk was then cleaned from *the cooker*, and whatever parts were necessary to do this were moved. [LOB]

The only exception to this claim is the example in (23). In this example the antecedent functions as the subject of the subordinate clause while the anaphor functions as a possessive. However, it is still possible to claim a certain amount of grammatical parallelism since the anaphor, while possessive, is within the subject noun phrase, *the taxi's memory*.

23. At the centre, dispatchers keep track of taxis on monitor screens, which can display up to five cabs per pole. Colour codes show how far *each cab* is from each pole. The system updates the information every 30 seconds, and *the taxi's memory* can retain a pole's code for up to three minutes. [BC]

It is this feature of the naturally occurring examples that allows us to explain why (15) is not an acceptable telescope example, i.e. this example is not an acceptable telescope construction because it does not have grammatical function parallelism.

5.0 Conclusion

To account for the analysis of telescope noun phrases I proposed that the potential lack of interpretation for the telescope noun phrase licenses the antecedent-forming process of abstraction. In turn, abstraction licenses the optional rule of distribution. When combined these two previously motivated processes provide an antecedent for the telescope noun phrase.

This approach avoids the problems that were identified with previous analyses. It provides the correct analysis for a class of more complex examples which contain an additional noun phrase anaphoric on a noun phrase within the scope of the telescope NP's antecedent. Furthermore, with the abstraction/distribution approach it is not necessary to stipulate the type of noun phrase that can be a telescope antecedent.

As a first step towards developing licensing conditions, in the second part of the paper I reviewed some of the characteristics of naturally occurring telescope examples. My analysis revealed (i) that the telescope NP is not restricted to pronouns, other definite noun phrases can also be interpreted telescopically, (ii) telescope constructions exhibit verb feature parallelism and (iii) telescope

constructions exhibit grammatical function parallelism. We saw that even these initial observations go partway towards providing an account of when a telescope construction is licensed.

Endnotes:

1. The analyses of Roberts, Poesio & Zucchi, and Dekker are the only ones which deal specifically with telescope noun phrases. The other papers have in common the fact that examples of the telescope construction are included in their discussion (although they are not recognized as such).
2. For further discussion of problems with previous analyses see Toole (1997).
3. The use of the *pl* superscript in this rule can be ignored for the present. Its function will be discussed later in this paper.
4. I call a statement of the form $x \text{ reln } \Sigma x K$ an abstraction condition because this is the type of condition introduced by the process of abstraction.
5. The term 'cardinality quantifier' was introduced by Partee (1988) and identifies those quantifiers that assert that a certain set meets specific cardinality conditions (e.g. that it contains one member, no members etc.). Determiners which introduce cardinality quantifiers are referred to as cardinality determiners. On Kamp & Reyle's analysis cardinality determiners introduce conditions of the form $x \text{ reln } \Sigma K$.
6. One disadvantage of the analysis described above is that, like Kamp & Reyle's approach in general, it is not compositional. However, this same analysis can be represented compositionally by incorporating Reyle's (1993, 1995, 1996) Underspecified Discourse Representation Theory. See Toole (1997) for further details.
7. I make this assumption since no sources are provided for the examples in the literature (the nature of the examples seems to confirm this assumption).
8. As a first step it was necessary to compile a set of naturally occurring examples. I examined the London-Oslo Burgen corpus, the Brown Corpus, the Associated Press corpus, and a section of the Wall Street Journal corpus. The London-Oslo Burgen, Brown, and Associated Press corpora were kindly made available by the Oxford text Archive. The examples were found by searching the corpora for the words *every* and *each*. I then manually examined the surrounding context to see if there were any singular pronouns antecedent on the universally quantified NP.
9. I classify this example as unacceptable since, given a scale of *fine-mediocre-no good*, three native speaker informants judged this sentence to be mediocre and one native speaker informant judged the sentence to be *no good*.

References:

- Dekker, Paul (to appear). Scope in discourse. *Journal of Language and Computation*. Available on-line at <http://turing.wins.uva.nl/~pdekker/abstracts.html>.
- Kamp, Hans & Reyle, Uwe (1993). *From Discourse to Logic*. Dordrecht: Kluwer Academic Publishers.
- Neale, Stephen (1990). *Descriptions*. Cambridge Massachusetts: MIT Press. A Bradford Book.
- Partee, Barbara (1988). Many quantifiers. In Proceedings of ESCOL.
- Poesio, Massimo & Sandro Zucchi (1992). On telescoping. In Chris Barker and David Dowty (eds.) *Proceedings of SALT II*. Columbus Ohio: Ohio state University. 347-366.
- Reyle, Uwe (1993). Dealing with ambiguities by underspecification: construction, representation, and deduction. *Journal of Semantics* 10. 123-179.
- Reyle, Uwe (1995). On Reasoning with Ambiguities. In *Proceedings of the European Association for Computational Linguistics*. Also in cmp-lg e-print archive at <http://xxx.lanl.gov>, document cmp-lg/9502026.

- Reyle, Uwe (1996). Co-indexing labelled DRSs to represent and reason with ambiguities. In Stanley Peters and Kees van Deemter (eds.), *Semantic Ambiguity and Underspecification*. Stanford: Centre for the Study of Language and Information. 239-268.
- Roberts, Craige (1987). Modal Subordination, Anaphora, and Distributivity. Ph.D. Dissertation, University of Massachusetts at Amherst.
- Roberts, Craige (1989). Modal subordination and pronominal anaphora in discourse. *Linguistics and Philosophy* 12: 683-721.
- Sells, Peter (1985). *Restrictive and Non-restrictive Modification*. Report no. CSLI-85-28. Stanford: Centre for the Study of Language and Information.
- Toole, Janine (1997). *The Semantic Interpretation of Definite Noun Phrases: Uniqueness Intuitions, Telescope Anaphora, and the Strong and Weak Readings in Donkey Sentences*. Ph.D. Dissertation, Simon Fraser University.

The Natural Classes of Two-Handed Signs

Janine Toole
Simon Fraser University

Linda Uyechi
Stanford University

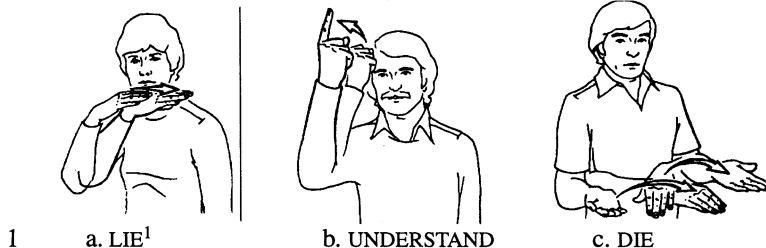
1.0 Introduction

In this paper we focus on identifying the natural classes of two-handed signs in American Sign Language (ASL). To this point, the overall goal of our work has been to develop a formal representation of signs and formal definitions of sign symmetry. This has led us in previous work (Uyechi & Toole 1997) to argue for a set of formal binary symmetry features for each of the parameters of a sign: namely, handshape, orientation, location, and movement. In this paper we take this approach one step further. We argue that the symmetry features are not independent - rather, they are related by a feature hierarchy. Recognizing this interdependency leads us to a proposal for the natural classes of two-handed signs that not only accounts for the distribution of two-handed signs but also accounts for some properties of diachronic change.

We begin by giving a brief introduction to the basic properties of two-handed signs and introduce the two puzzles that motivate our work. We then review the symmetry features introduced in Uyechi & Toole (1997) and introduce our approach to natural classes in ASL.

2.0 Basic Properties of Signs

Signs are described in terms of four parameters: handshape, location, orientation, and movement. Following Stack (1988) and Hayes (1993), we take the view that only handshape, location and orientation are phonological primitives. On this analysis, movement is not a primitive. Rather, it is derived from changes in the other three parameters, as illustrated by the signs in (1).



To articulate (1a), the hand starts at one side of the chin and ends at the other side. This is an example of a change in location. In (1b) the hand starts in a position in which all the fingers and thumb are folded into a fist with the hand held at the side of the signer's head and the palm of the hand facing towards the signer. To articulate the sign the index finger is extended until it points straight up. This is an example of a change in handshape. To articulate (1c) the hands start in a position in which one palm faces up and the other down. The hands are then rotated 180 degrees so that each palm faces the other way. This is an example of a change in orientation. All movements articulated in monomorphemic lexicalized signs can be represented by these three components of movement: change in location, change in handshape, and change in orientation (Uyechi 1995).

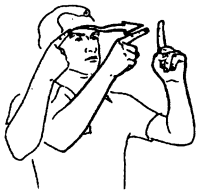
The earliest typology of two-handed signs is Battison's (1978) three-way typology. Although we note elsewhere that Battison's definitions are too informal and imprecise to form the basis of an adequate typology (see Toole & Uyechi 1995, Uyechi & Toole 1997), it is a well-known system that provides a convenient way to describe two-handed signs. For this reason, we informally introduce Battison's typology and use it as a reference system in the remainder of this paper.

Briefly, Battison classifies two-handed signs into one of three classes: Type 1, Type 2, and Type 3 signs. In a Type 1 sign, (2a), both hands have the same handshape, orientation, and location, and both hands move. In a Type 2 sign, (2b), both hands have the same handshape, but only one hand moves. In a Type 3 sign, (2c), the hands have different handshapes and only one hand moves.

2.



a. ALIKE



b. GOAL



c. DISCUSS

3.0 The puzzles

3.1 Puzzle 1: Diachronic Variation

Previous research has noted that some two-handed signs exhibit a diachronic change towards more symmetrical forms (Frishberg 1975, Sandler 1993, Brentari 1995). For example, the original sign for WHISKEY is a Type 3 sign, as shown in (3a). The handshapes are different, and only one hand moves. However, the more common articulation of the sign is now the Type 2 pronunciation in (3b): in this sign

the handshapes are the same, although only one hand moves.

3.



a. WHISKEY[Type 3]
older version



b. WHISKEY[Type 2]
newer version

Of interest to us is the fact that many Type 3 signs do NOT undergo this migration. For example, the sign in (2c), DISCUSS, is a Type 3 sign - the hands have different handshapes and only one hand moves - and it gives no indication of migrating to a Type 2 sign. Rather, it is clearly a stable Type 3 sign.

The puzzling behavior we need to address is how to predict which Type 3 signs are “stable” Type 3 signs and which are likely to migrate towards the more symmetrical form of a Type 2 sign. The general classes of Battison’s typology and similar proposals such as Sandler (1993) and van der Hulst (1993) cannot account for this important difference in Type 3 signs.

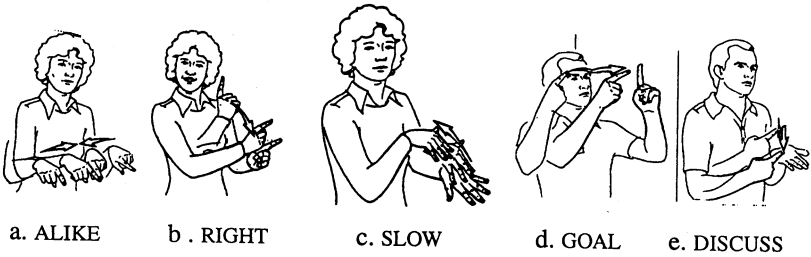
3.2 Puzzle 2: The Distribution of Two-handed Signs.

In previous work we provided a formal definition of symmetry for each of the four parameters of a sign (Uyechi & Toole 1997). We proposed that each sign can be classified according to whether or not it has symmetric handshape, symmetric location, symmetric orientation, and symmetric movement. A consequence of this approach is that there are sixteen logically possible combinations of symmetry features, and therefore, sixteen possible classes of two-handed signs.

However, in a random survey of over 100 two-handed signs, only five of these logical possibilities are in evidence. These classes are shown in (4), where a ‘+’ indicates that this parameter of the sign is symmetric. A representative example from each class is given in (5). The second puzzle, then, is to account for this empirical distribution of signs.

4.		HS	LOC	OR	mov		
	A	+	+	+	+	ALIKE	Maximally symmetric
	B	+	+	+	-	RIGHT	
	C	+	+	-	-	SLOW	
	D	+	-	-	-	GOAL	
	E	-	-	-	-	DISCUSS	Maximally non-symmetric

5.



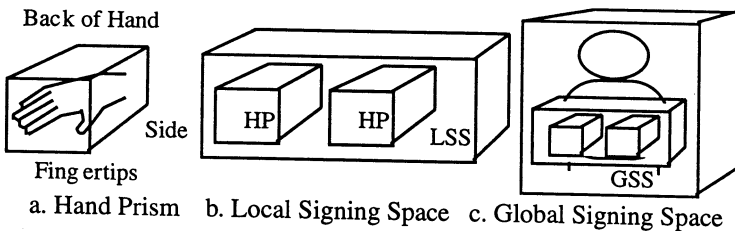
4.0 Symmetry Features

In order to account for the two puzzles introduced in the previous section, we begin by introducing Uyechi's (1996) geometry-based model of signs and Uyechi & Toole's (1997) formalization of symmetry. Following this we introduce the Feature Symmetry Hierarchy which we use to account for the two puzzles.

4.1 The Geometry-based Model

The geometry-based model of sign phonology is based on a three-dimensional rectangular representation of signing space. As illustrated in (6), the hands are represented as hand prisms which are embedded in rectangular prisms of signing space. Monomorphemic signs are articulated in local signing space, (6b), which is in turn embedded in global signing space, (6c). Morphologically complex signs are specified for positions in both local and global signing space. At an even higher level of representation, discourse is represented as instances of global signing space embedded in a discourse signing space. Of interest here is the representation of monomorphemic signs, hence the discussion will focus on the relation between the hand prisms and local signing space

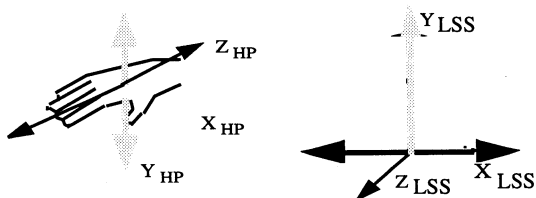
6. Geometry Based Representation of Articulators and Signing Space



In particular, we focus on the relation between the hand prisms within local signing space, and global signing space. The position of local signing space (LSS) does not change during the articulation of a monomorphemic sign, rather, it carves out the relevant part of signing space. The places of local signing space then provide

reference planes for determining the relative posture of the hands. Movement of the hands is represented as changes in position of the hand prisms within local signing space.

Finally, we note that the dimensions of the hand prisms, local signing space, and global signing space can also be represented in terms of the axes which define the prisms. For example, the dimensions of the hand prism in (6a) can be represented by the three axes as in (7a). The dimensions of the local signing space in (6b) can be represented as in (7b).



7. a. Hand Prism b. Local Signing Space

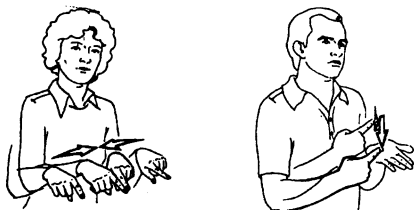
Using this system we are able to formalize symmetry for each of the components of a sign, namely handshape, location, orientation, and movement

4.2 Formalizing Symmetry

In this section we use the constructs of the geometry-based model to formalize the definition of symmetry. We define symmetry for each of the components of the sign described above, namely handshape, location, orientation, and movement. The brief definitions we give here are based on the more detailed definitions given in Uyechi & Toole (1997).

HANDSHAPE:

Handshape is symmetrical if both hands have the same handshape. For example, (8a) is articulated with both hands in the same configuration. In contrast in (8b) the moving hand has a different configuration from the non-moving hand. Hence (8a) has symmetrical handshape, but (8b) does not.



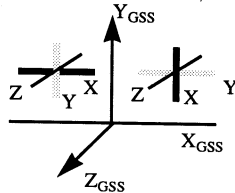
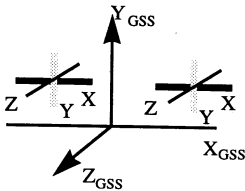
8. a. ALIKE (+symHS) b. DISCUSS (-symHS)

ORIENTATION:

A sign has symmetric orientation if it meets the constraint given in (9) (where “corresponding axes” refers to the X-axes of both hand prisms, the Y-axes of both hand prisms, and the Z- axes of both hand prisms).

9. Constraint on Orientation in Two-Handed Signs
In a two-handed sign, corresponding axes must be parallel.

For example, in (10a) all three axes are parallel to each other throughout the articulation of the sign. In contrast, in (10b) only two sets of corresponding axes are perpendicular to each other. Hence, (10a) is articulated with symmetric orientation, but (10b) is not.²



10. a. ALIKE (+symOR)

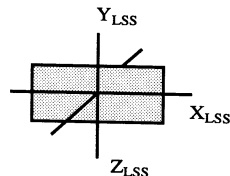
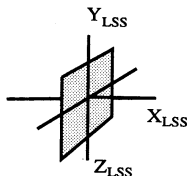
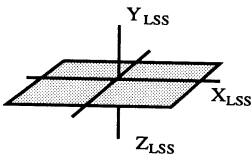
- b. GOAL (-symOR)

LOCATION:

Using the relative position of the hand prisms in local signing space, we define symmetric location in (11) where we use the formal definition of the “origin” of the hand prism as the point at which the X-, Y-, and Z- axes of the hand prism intersect each other.

11. A sign has symmetrical location if, at some point in the sign, the location of the origin of one hand prism is a reflection of the location of the origin of the other hand about one plane of symmetry in LSS.

The planes of symmetry for the location parameter are the planes defined by the axes of local signing space, namely, the X-Z plane, the Y-Z plane, and the X-Y plane, illustrated in (12).

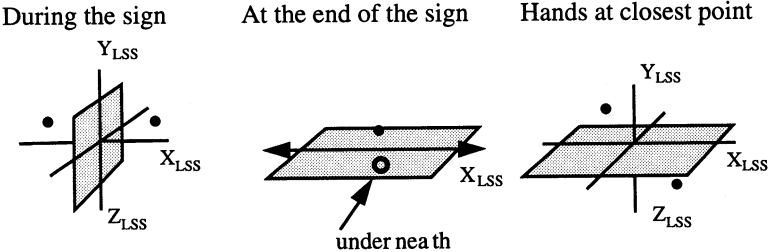


12. a. X-Z place

- b. Y-Z plane

- c. X-Y plane

For example, in the sign ALIKE, (2c), the location of the hands is symmetric throughout the articulation of the sign. In (13a) we illustrate this by showing the origins of the hands relative to each other at the beginning of the sign. As is evident from this diagram the origins of the hand prisms are reflected about the centre (Y-Z) plane. In (13b), the locations are symmetric only at the end of the gesture because one hand is static throughout the articulation of the sign. At the end of the gesture the origins of both hand prisms are reflected about the base (X-Z) plane. In (13c), even at the points where the hands are closest, at the end of the sign, the origins of the hand prisms are skewed with respect to the base plane, and are not reflected about either of the other planes. Hence (13a) and (13b) are articulated with symmetric location but (13c) is not.



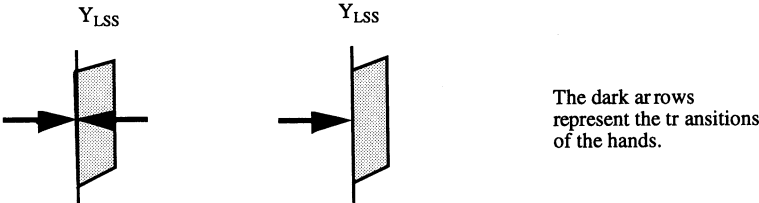
13. a. ALIKE (+symLOC) b. RIGHT (+symLOC) c. DISCUSS (-symLOC)

MOVEMENT:

In the geometry-based system “movement” is represented as a transition in one of the primary components. Specifically, movement is represented as either a change in handshape, change in location, or change in orientation. Hence, symmetric movement is defined as in (14).

14. A sign has symmetric “movement” if the transition(s) of the hands are a reflection about a plane of symmetry in LSS.

For example, both signs in (15) are articulated with a change in location. The movement vectors in (15a) are reflections about the centre plane. In contrast, only one hand moves in (15b), so the vectors are not reflections about any plane. Hence (20a) is articulated with symmetric “movement” but (20b) is not.



15. a. ALIKE(+symMOV) b. GOAL(-symMOV)

In sum, for *each* of the four parameters, a sign can be classified as symmetric or non-symmetric. Having established the required background, we now return to the two puzzles that motivate our work.

5.0 Accounting for the Puzzles

To account for the two puzzles introduced in section (3.0), we propose that the symmetry of each feature is not independent. Rather, the symmetry features are related by the Feature Symmetry Hierarchy in (16). In this hierarchy, the symmetry of any feature assumes the symmetry of all features to its left. Thus, if a sign has symmetric movement, then it necessarily has symmetric handshape, orientation, and location. If a sign has symmetric location, then it necessarily has symmetric orientation and handshape, and so on.

16.

Feature Symmetry Hierarchy

HS > OR > LOC > mov

The hierarchy provides both explanatory and predictive power for the analysis of two-handed signs: it explains the occurrence of only five of the sixteen possible classes of two-handed signs, and it provides the basis for predicting the migration of two-handed signs towards more symmetric forms. We start by describing the explanatory power of the hierarchy.

The five classes allowed by the hierarchy are listed in (17). Simply stated, no other classes meet the requirements of the Feature Symmetry hierarchy.

17.	class	HS	LOC	OR	mov	
	A	+	+	+	+	ALIKE Maximally symmetric
	B	+	+	+	-	RIGHT
	C	+	+	-	-	SLOW
	D	+	-	-	-	GOAL
	E	-	-	-	-	DISCUSS Maximally non-symmetric

Comparing these classes with those presented in section 3.0, we see that the predicted classes in (17) correspond exactly with the classes in (4) that emerged from an examination of over one hundred randomly chosen signs. Hence, the hierarchy provides an account of the empirical distribution of data.

The predictive power of the hierarchy is illustrated by applying it to the puzzling data from diachronic change described in section 3.1. Specifically, we address the question of why some Type 3 signs undergo diachronic change and become Type 2

signs (WHISKEY), whereas other do not (DISCUSS). Again, the hierarchy accounts for this problem. In (18) we compare the symmetry features for the two signs, WHISKEY and DISCUSS.

18.		HS	LOC	OR	mov
	DISCUSS S	-	-	-	-
	WHISKEY S	-	+	+	-

Although both of these signs are classified as Type 3 signs under Battison’s typology, when presented in terms of their symmetry features these signs are clearly distinct types of signs. The symmetry features of DISCUSS (18a), indicate that none of its parameters are symmetric. This is a classification that is consistent with the Feature Symmetry hierarchy: thus predicting that the signs is stable and unlikely to undergo diachronic change.

In contrast, the configuration of the symmetry features for WHISKEY, (18b), violates the Symmetry Feature hierarchy. It has symmetric location and orientation but not symmetric handshape. Because the hierarchy requires that handshape be symmetric if location and orientation are symmetric, it predicts that the form in (18b) is unstable; hand shape must be symmetric in order for this sign to be consistent with the hierarchy. In fact, the more recent form of WHISKEY is one in which the static hand has the same handshape as the moving handshape. Having acquired symmetric handshape, WHISKEY now meets the Feature Symmetry hierarchy.

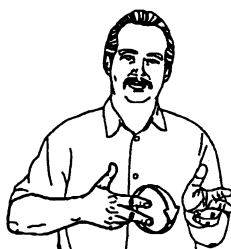
This prediction holds of other signs that have entered into the lexicon, e.g., WORLD and RESIDENTIAL-SCHOOL. In their original form these signs have the same symmetry features as WHISKEY and do not adhere to the Feature Symmetry hierarchy. Each of these signs has undergone changes that parallel the changes to WHISKEY so that each is now consistent with the hierarchy.

In addition, there is cross-linguistic evidence to support this view. In (19) we show both old and new versions for the sign for TRAIN in Italian Sign Language (LIS) (Radutzky 1990). The older version violates the Feature Symmetry hierarchy; it has symmetric location and orientation but not symmetric handshape. The newer form of the sign has acquired symmetric handshape and is now consistent with the hierarchy. We claim, then, that signs that are not consistent with the feature hierarchy are unstable and are likely to undergo changes that bring them into line with the hierarchy.

19.



a. TRAIN (old)



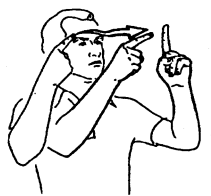
b. TRAIN (new)

Hence, the feature hierarchy explains the presence of precisely five classes of signs in the data as well as correctly predicting the migration of two-handed signs towards more symmetric forms. Thus, we present these five classes as candidates for the natural classes of two-handed signs.

As natural classes, we would expect other phonological processes to be sensitive to them. Indeed this is the case. As illustrated in (20), the feature hierarchy and the natural classes of two-handed signs that it predicts account for the synchronic alternation noted by Uyechi & Toole (1997): namely, that some Type 2 signs have a type 1 pronunciation. In other words, some signs that are articulated with the hands in the same handshape but one hand static can also be articulated with both hands moving symmetrically. For example, RIGHT, shown in (20a), can be articulated with one hand static, or with both hands moving symmetrically. In contrast, a Type 2 sign such as (20b) does not have a Type 1 pronunciation.



a. RIGHT



b. GOAL

20

Again, comparing the symmetry features of the Type 2 signs reveals a significant underlying difference. As shown in (21), all but the movement parameters are symmetric for RIGHT, (21b), whereas only the handshape parameter for GOAL, (21a) is symmetric.

21.

	HS	OR	LOC	mov
a. GOAL	+	-	-	-
b t RIGHT	+	+	+	-

Hence, the feature hierarchy again provides an explanation for the alternation for

RIGHT, and the lack of alternation for GOAL. For RIGHT, all of the parameters that are lower in the feature hierarchy are symmetric, so it can acquire symmetric movement and still be consistent with the hierarchy. In contrast, GOAL lacks symmetric location and orientation. It cannot acquire symmetric movement because this would make the sign inconsistent with the hierarchy (handshape, location, and orientation must be symmetric before movement can be symmetric).

In sum, in this section we have introduced a symmetry hierarchy that dramatically restricts the gestures that can be valid two-handed signs. This hierarchy predicts that there are five classes of signs, which we identify as class A-class E. In class A all the parameters of a sign are symmetric, in class B all parameters except movement are symmetric, in class C only handshape and orientation are symmetric, in class D only handshape is symmetric. Finally, in class E signs none of the parameters are symmetric. As well as accounting for the characteristics of two-handed signs and excluding gestures which are not natural two-handed signs, we found that the hierarchy also provides insight into cases of diachronic change and synchronic variation. Furthermore, there is cross-linguistic evidence to support this claim.

6.0 Some Consequences for Universal Phonology

Our approach raises interesting questions with respect to universal phonology. Specifically, the sign symmetry features that we are proposing are associated with a whole sign gesture. Yet, a large number of sign phonology proposals support the idea that there is a one to one mapping between the sign gesture and the syllable. This suggests that for sign language some features are associated with the *syllable*. This is a highly irregular finding from the perspective of spoken language phonology. However, we believe it is premature to focus on these apparent differences between spoken language and signed languages. Rather, a more productive approach is to continue to study signed languages using theories and constructs which seem most relevant to the visual mode.

As we have seen, the geometry-based approach we take has allowed us to identify classes which (i) reflect sign distribution, and (ii) which are relevant to phonological processes in ASL. We believe the best approach is to study sign phonology independently of spoken language phonology so that ultimately each will provide independent evidence for a *truly* universal theory of phonology.

Endnotes:

1. In this paper we follow the convention of glossing signs with small capitals.
2. One additional point needs to be made with respect to symmetric orientation. First, when articulating a sign like (RIGHT/CORRECT), the axes of the hands

are skewed articulatorily. This is arguably induced by the physiology of the hands. The environment for this adjustment is restricted, and shifting the hands so the axes are aligned does not change the meaning of the sign (as happens for some signs, e.g., NAME). Hence, this variation can arguably be captured by a phonetic adjustment rule. With respect to the phonology, the orientation for these signs is symmetric.

References:

- Battison, R. 1978. *Lexical Borrowing in American Sign Language*. Silver Spring: Linstok Press.
- Brentari, D. 1995. *A prosodic account of two-handed signs*. University of California, Davis. ms.
- Frishberg, N. 1975. Arbitrariness and iconicity: historical change in American Sign Language. *Language* 51. 696-719.
- Hayes, B. 1993. Against Movement: Comments on Liddell's Article. In Geoffrey Coulter (ed.) *Phonetics and Phonology*, Volume 3: Current Issues in ASL Phonology, Academic Press.
- Humphries, T., Padden, C., & O'Rourke, T. J. 1980. *A Basic Course in American Sign Language*. Silver Spring: T. J. Publishers.
- Radutzky, E. 1990. The changing handshape in Italian Sign Language. In W. H. Edmondson and F. Karlsson (eds.) *Papers from the 4th International Symposium on Sign Language Research*.
- Sandler, W. 1993. Hand in hand: The roles of the nondominant hand in sign language phonology. *The Linguistic Review* 10. 337-390.
- Stack, K. 1988. *Tiers and Syllable Structure in American Sign Language: Evidence from Phonotactics*. MA Thesis, University of California, Los Angeles.
- Toole, J. & Uyechi, L. 1995. *Another perspective on hand orientation in American Sign Language*. In Proceedings of WECOL.
- Uyechi, L. 1995. *The Geometry of Visual Phonology*. Ph.D. Thesis, Stanford University.
- Uyechi, L. 1996. *The Geometry of Visual Phonology*. Stanford: CSLI Publications.
- Uyechi, L. & J. Toole 1997. Formal Symmetry in American Sign Language. In *Proceedings of the Berkeley Linguistics Society*, Feb 14-17, 1997.

(Pictures from Humphries, Padden, and O'Rourke (1980), illustrated by Frank Paul.)

***Bueno*: A Spanish Interactive Discourse Marker¹**

Catherine Travis
La Trobe University, Australia

—*Vayase al carajo— balbuceó Trueba sin convicción.*

—*Bueno, por allá vamos. Usted viene conmigo.*

'Go to hell,' stammered Trueba without conviction.

'*Bueno*, that's where we're going. You're coming with me.'

(Isabel Allende. *The House of the Spirits*: 319)

1. Introduction

This paper presents a semantic analysis of the Spanish Discourse Marker (DM) *bueno* (something similar to 'well', 'OK', 'alright' in English). Using recordings of conversational Colombian Spanish, and drawing from two major works on this Marker (Bauhr 1994 and Fuentes 1993), four discourse functions of *bueno* will be identified and discussed. On the basis of these functions, it will be suggested that *bueno* is polysemous, and following the Natural Semantic Metalanguage approach, two definitions to account for its range of use will be proposed. It will be argued that these definitions reflect the way in which the meaning of the DM *bueno* is related to the adjective *bueno*, meaning 'good'. Finally, the role *bueno* may play in the communicative realization of the norms and values of the cultures in which it is used will be considered.

2. Natural Semantic Metalanguage Approach

The theoretical framework used here is the Natural Semantic Metalanguage approach, as developed by Wierzbicka (1972, 1980, 1996 and references therein), and colleagues, principally Goddard (1989, 1994). This approach is based on the notion of reductive paraphrase: the use of a limited set of maximally simple and maximally universal words, combined in accordance with a limited set of maximally simple and maximally universal syntactic patterns, to form a kind of script, outlining the meaning of the item under consideration. The lexicon of the Metalanguage numbers around 60 words and includes, for example, 'I', 'you', 'good', 'bad', 'think', 'know', 'say', 'want', 'this', 'something', 'same' etc. These words are combined to form sentences such as: 'you say something'; 'I think this is good'; 'I feel something bad'; 'I want you to know this', etc. These words and patterns are believed to be found in most, if not all, of the world's languages, and thus the definitions should be readily translatable cross-linguistically. Such definitions are of use for language-learning, for cross-linguistic studies, in that they enable apparently related items to be compared across languages, and for linguistic studies comparing related items within the one language. Comparative work will be left for further research, and this study will concentrate on just the one marker, *bueno*.

3. Data

The data on which this study is primarily based are recordings of conversational Spanish, made in Colombia in 1997. This comprises three hours and 36 minutes of conversation, and presents a total of 81 tokens of *bueno*. The majority of

examples given here are from this corpus, although examples will occasionally be drawn from other sources to support points made in the discussion. These include interviews conducted in Australia with Colombians in 1997, novels by Latin American authors, and other studies of *bueno*, which have focused on its use in Spain. The source of each example is indicated above it. Although there are differences in terms of frequency of occurrence of *bueno* in different dialects and genres, there is no apparent difference in the semantics of *bueno*, and the definitions proposed are intended to apply equally well to these other data.

The conversational data has been transcribed in accordance with the Du Bois et al. transcription method (1993). Each line of the transcription represents one intonation unit, defined as a string of words that occur under one coherent intonation contour (1993:47). The transcription conventions used are given in the appendix.

4. Functions of *bueno*

The polyfunctionality of DMs has been widely noted in the literature (cf. Schiffrrin 1987, Fraser 1990, Jucker 1993, Lenk 1998), and has been noted specifically for *bueno* in the work of Bauhr (1994) and Fuentes (1993). Applying the analyses of Bauhr and Fuentes to the Colombian conversational data, four basic functions that account for all the cases presented in these studies have been identified. These are used to mark: acceptance; a response (that is not pure acceptance); a reorientation (or topic change); and a correction. Although acceptance is clearly distinguishable from the other functions, the other functions are interrelated, and are not always mutually exclusive (cf. Bauhr 1994:79). That is, some cases of *bueno* marking a response, for example, could also be seen to be marking a reorientation, or a correction. It will be argued that these three functions share the same meaning. One definition will therefore be proposed to account for the use of *bueno* indicating acceptance, and another to account for that of marking a response, reorientation and correction.

The following table presents the frequencies of occurrence of *bueno* on the basis of function and meaning for the Colombian conversational data. It is interesting to note that each of the two different meanings to be proposed account for roughly the same number of tokens.

Functions	acceptance	response	reorien- tation	correction	TOTAL
Tokens	39 (48%)	14 (17%)	25 (31%)	3 (4%)	81 (100%)
		42 (52%)			

These functions will now each be considered individually, and meanings to account for them proposed.

4.1 Acceptance

Indicating acceptance is the most common use of *bueno* in the data, accounting for 39 of the 81 tokens, or 48%. *Bueno* can be used to mark acceptance of an

offer, or a proposal, or also information receipt, indicating that the speaker has understood and accepted what another has said. Some examples are given below.

In (1), Angela uses *bueno* to accept Sara's order for two pizzas.

- | | | |
|---------|--------------------------------|--------------------|
| (1) | | 27: insurance (24) |
| SARA: | [Yo encargo dos]. ² | [I'll order two]. |
| | [2@@@2] ³ | [2@@@2] |
| ANGELA: | [2Ah, | [2Oh, |
| | bueno2]. | bueno2]. |
| | @@ | @@ |
| | Listo. | OK. |

In (2), *bueno* is used to indicate Santi's acceptance of what Angela has told him; namely that the tape recorder should be left to continue recording.

- | | | |
|---------|----------------------------|------------------------------------|
| (2) | | 3: almuerzo (10) |
| ANGELA: | .. Tiene que seguir así. | .. It has to keep going like that. |
| SANTI: | .. Sí? | .. It does? |
| ANGELA: | Sí. | Yes. |
| SANTI: | .. Ah, | .. Oh, |
| | bueno . | bueno . |
| | chao. | ciao. |
| | ((ANGELA LEAVES THE ROOM)) | |

Bueno can also be used to indicate acceptance of an offer, as shown in (3).

- | | | |
|-----|------------------|-----------------------|
| (3) | | (Bauhr: ONE71 UT:201) |
| H1 | —¿Quieres fumar? | Do you want a smoke? |
| H2 | — Bueno . | Bueno . |

Bauhr and Fuentes claim that the use of *bueno* to mark acceptance encodes reluctance to comply, and some kind of concession on the part of the speaker (Bauhr 1994:92ff, Fuentes 1993:208ff). Both compare it with other markers such as *claro*, *por supuesto*, *seguro*, all meaning something like 'sure, of course' (Fuentes 1993:209), or even *con mucho gusto* 'with pleasure' (Bauhr 1994:93), which certainly are more enthusiastic ways of replying. This notion of reluctance can be seen in (4) below. This comes from a conversation between three participants, arranging to go somewhere that afternoon. Angela and Patricia (the third participant) both want to go at two, and Clara wants to go at 2.30. Here, Clara concedes to going at two.

- | | | |
|---------|--------------------------------------|---------------------------------|
| (4) | | 36: Colombo (31) |
| ANGELA: | Por qué no vamos a las dos? | Why don't we go at two? |
| | .. @@ | .. @@ |
| CLARA: | ... Ay, | ... Oh, |
| | <VOX no puede ser a las dos y media? | <VOX can't it be at two thirty? |
| | Yo= quiero VOX> -- | I want VOX> -- |
| | ((AS THOUGH CRYING)) | |
| | ... Bueno , a las dos pues. | ... Bueno , at two then. |

Note that Clara encodes her reluctance here through voice quality, as she is pretending to be crying, and thus *bueno* alone is not carrying this implication. Also note that in (1) above, there is clearly no notion of reluctance or concession. Following *bueno*, Angela laughs, and agrees with *listo* 'OK', and thus this is quite an enthusiastic acceptance. It therefore cannot be the case that reluctance is inherent in the meaning of *bueno* itself. Rather, what seems to be the case is that *bueno* encodes a neutral or non-committal response, and it is this that allows for 'reluctance' as a possible interpretation. In accepting an offer, for example, with *bueno*, the speaker indicates that this is not something they have been eagerly awaiting or expecting (for example, in (3), that they be offered a cigarette), but something that, now proposed, is accepted. Its neutrality also means that prosodic and other features can be used to encode reluctance (or enthusiasm), in accordance with the pragmatics of the situation.

I propose the following semantic formula to represent the meaning of *bueno* as it is used to indicate acceptance, which will be called *bueno*1.

1. you say something (X)
2. I say: this (X) is good

The first component captures the fact that this refers to the use of *bueno* as a response to something someone else has said. ('I' and 'you' are used here to refer to the speaker, who accepts with *bueno*, and the interlocutor, respectively.) The second component captures the fact that *bueno* is a positive comment on what another person has said, and it is in this sense that it is related to the adjective 'good'. Note that this formula encodes nothing about the speaker's real (versus portrayed) attitude to the preceding discourse, as would need to be included in more (and less) enthusiastic responses. For example, something like *muy bien* 'very good' may encode a notion of 'I think this is good', and *con mucho gusto* 'with pleasure' may encode 'I feel something good about this', while no such notions are encoded in *bueno*. *Bueno* is a 'neutral' way of accepting something, by saying 'what you have said is good', without committing oneself further.

4.2 Prefacing a response

The use of *bueno* to preface a response is similar to that of indicating acceptance in that both respond to something someone else has said. These two functions differ in one immediately apparent way, however, in that in one function (that considered above), *bueno* itself forms the response, and in the other (that to be considered in this section), *bueno* is used to mark the response which follows.

Included under the heading of prefacing a response is the use of *bueno* to preface a response to a comment, and an answer to a question. In both (5) and (6), *bueno* is prefacing a response to a comment.

In (5), Milena is discussing the work done by environmental organizations in the small town where she lives, and is saying that in the light of all there is to be done, there isn't enough money. Rosario accepts this with *bueno*, but adds that even small projects, when timely, can make a difference.

(5)

53: contamination (36)

MILENA: .. *No es suficiente.*
ROSARIO: *Sí.*
... **Bueno**,
pero a veces cositas
pequeñitas no?
Hay proyeticos así=,
como=,
.. *puntuales,*
.. *No?*

.. It's not enough.
Yes.
... **Bueno**,
but sometimes little things don't you
think?
There are projects like that,
kind of,
.. timely,
.. Don't you think?

The use of *bueno* here indicates that Rosario agrees with Milena, but that her agreement is only partial. This is similar to what has been seen for *bueno* encoding acceptance, in that, by agreeing, it marks acceptance of another's comment. Here, however, *bueno* indicates that some modification of that comment is needed. As was proposed for *bueno*₁, this use also implies a positive evaluation, but in this case, it is not of what Milena has said, as clearly Rosario does not want to say that it's good that there isn't enough money. Rather, the positive evaluation seems to be of the fact that the speakers are in agreement.

In (6), Santi asks Angela if she knows how the age of trees is measured, and she replies with an answer he hadn't thought of.

(6)

14: restaurant (24)

SANTI: [*Sabes como le miden la*
edad] *a un arbol?*
ANGELA: *Con .. el carbono catorce?*
SANTI: **Bueno**,
también.
No me acordaba de eso.

[Do you know how they measure the
age] of a tree?
With .. the carbon-14?
Bueno,
that too.
I didn't remember that.

Santi responds to Angela's answer with *bueno*, indicating that he agrees with her, while it is not what he had in mind. The use of *también* 'also' shows that there is another method he was thinking of, and he then admits that he had forgotten the method Angela has proposed.⁴ Thus, again, *bueno* is used to indicate that the speaker agrees with their interlocutor's comment, but only partially.

In (7), *bueno* is used to preface an answer to a question about an insurance policy the participants are discussing.

(7)

20: insurance (11)

ANGELA: *Y cuál es la diferencia,*
O sea,
en plata.
@ @
SARA: .. **Bueno**,
.. *Y entonces,*
Eso también es super
importante.

And what's the difference,
I mean,
in money.
@ @
.. **Bueno**,
.. And so,
That too is very important.

In this excerpt, Angela asks about the difference in cost for the different options she has available to her. Before going on to answer this, Sara continues with what she

had been saying before Angela asked this question, outlining one very important element of the policy, and only then turns to answer Angela's question.⁵ Thus, the response immediately following the question does not in itself constitute an appropriate answer, although Sara does go on to answer Angela's question following this. *Bueno* is used in this context to acknowledge the validity of the question (and in this sense is similar to the notion of agreement seen above), and to indicate at the same time that it is not going to be answered without some further information being given.

One final example to illustrate this is the following.

- (8) (Bauhr: ONE71 UT:158)
 H1 —*Hay cosas más importantes que* —There are more important things to
hacer. [...] (sic) ¿No crees? do. [...] (sic) Don't you think?
 H2 —*Bueno, depende* —*Bueno*, it depends.

Thus, in prefacing an answer to a question, *bueno* is used to acknowledge the validity of the question, but indicates that the answer is not a straightforward one, or cannot be given without some further information being presented first.

What is common to the use of *bueno* prefacing both responses and answers to a question is an implication of: agreement (cf. Bauhr 1994:92ff), either with the comment, or with the validity of the question; a positive evaluation of the fact that the interlocutors are in agreement; and that the agreement is only partial, and needs to be qualified in some way. Before presenting a definition to capture this use of *bueno*, its use prefacing a reorientation in the conversation will be considered.

4.3 Reorientation

The most common use of *bueno* in the data following that of acceptance, is to mark a reorientation in the conversation, which accounts for 25 of the 81 tokens, or 31%. Included under this heading are such uses as initiating a new topic, closing a topic, returning to an earlier topic following a digression, or other kinds of breaks in the flow of conversation (cf. Bauhr 1994:216, Fuentes 1993:216). Some examples are given below.

- (9) 49: contamination (7)
 OMAR: ... *No me acuerdo cómo se* ... I don't remember what it's called.
 llama.
 ROSARIO: ... *Bueno.* ... *Bueno.*
 ... *Una vez fuimos allá,* ... We went there once,
 llegando unos amigos when some German friends had
 alemanes? arrived?

Here, Rosario is telling a story about when she took some friends of hers to a beach which she can't remember the name of. None of her interlocutors can remember the name, but they all seem to know which beach she is talking about, and she goes ahead and tells the story. *Bueno* marks a reorientation from trying to remember the name of the beach (which is background information to her story), to the story itself. The meaning of *bueno* here is similar to what has been seen above: it marks the speaker's acceptance of the fact that nobody remembers the name of the

beach, and indicates that this doesn't matter, that she's going to go on and tell her story anyway.

(10)

ANGELA: *Mirá que,*
.. este señor,
... (2.0) Hm,
Parece que tiene es como --
.. un problema,
o yo no sé.
Bueno.
... Estaba ahorita= --
así como,
hablando .. solo?

15: restaurant (26)

Look,
 .. this man,
 ... (2.0) Hm,
 He seems to have kind of --
 .. a problem,
 or I don't know.
Bueno.
 ... Just now he was --
 kind of,
 talking .. to himself?

Here Angela presents her opinion about a customer in the restaurant where this conversation was recorded (that he seems to have some kind of problem). Following *o yo no sé. Bueno.* ('or I don't know. *Bueno.*'), she gives an example of why she thinks that (because he was talking to himself). *Bueno* marks her reorientation from her opinion about the man, to the reason for why she has that opinion. It encodes acceptance of what precedes (that Angela doesn't really know if the man has a problem), and implies that she believes she and her interlocutor are in agreement (that neither really know), but that she has something more to say about it (that she has some reason for thinking that he may, because he was talking to himself).

In what has been coded as marking a reorientation, *bueno* can function as a response to an interlocutor's comment, as in (9), or as a response to the speaker's own speech, as in (10). It is, however, often impossible to say precisely whose talk it is responding to, as in the following example. Here, Sara is going through the options available in an insurance policy she is describing to Angela. One of the options is a daily 'allowance' if she were to be hospitalized, of 110,000 pesos.

(11)

SARA: *[Ciento diez] .. diarios.*
 ANGELA: *.. Ah,*
diarios!
 SARA: *Diarios.*
 ANGELA: *... Bastante.*
 SARA: *.. Sí.*
.. Bueno.
Cobertura para el cancer es
importante.

24: insurance (15)

[A hundred and ten] .. daily.
 .. Oh,
 daily!
 Daily.
 ... That's a lot.
 .. Yes.
 .. **Bueno.**
 Coverage for cancer is important.

Sara uses *bueno* to return to the list of options she is going through, following the digression regarding the generous daily allowance. *Bueno* here could be a response either to Angela's comment, *Bastante* 'That's a lot', or to her own expression of agreement with this, *Sí*, 'Yes.' Thus, it is important that the definition capture this notion: that although *bueno* is a response to prior discourse, it is not inherent in the meaning of *bueno* whose discourse it responds to.

The meaning of *bueno* in (11) is as we have seen above: it indicates positive evaluation of agreement between the interlocutors, and that there is more to be said about this, in this case, that there are other options of the policy to go through.

In the light of these examples, I propose the following definition to account for *bueno* as a marker of response and of reorientation.

1. someone says something
2. I say: I think the same as you
3. this is good
4. there is something more to say

Component one reflects the fact that this use of *bueno* is a response to something someone has said. (Note that this differs from what has been proposed for *bueno*₁, which has been defined specifically as a response to a comment by the interlocutor, with the component 'you say something'.) Component two reflects the fact that *bueno* is a comment by the speaker indicating assumed agreement between the interlocutors. Component three captures the positive evaluation of that agreement (which is a reflection of the relation between the DM and the adjective *bueno*, 'good'), and component four, the fact that what precedes is not 'the full story', and that something needs to be added.

One final main function of *bueno* will now be considered; that of prefacing a correction.

4.4 Correction

In the Colombian conversational data, there are just three examples of *bueno* used to mark a correction. One is given below.

(12)

50: contamination (22)

((OMAR IS REPORTING A
DIALOGUE BETWEEN TWO
SPEAKERS,
LABELED S1 AND S2))

OMAR: <S1 ... Pero,
un= pingüino no mide
tanto S1>.

<S1 ... But,
a penguin isn't that tall S1>.

ROSARIO: [Mhm].

[Mhm].

OMAR: <S2 [*Bueno*],
qué sería,
uno cincuenta S2>.

<S2 [*Bueno*],
what would it be,
one fifty (cm) S2>.

Omar is telling a joke here, about a person (S2) who claims (prior to this excerpt) to have shot a penguin that was over 160 cm tall. (The joke is that what they have shot was not in fact a penguin, but a nun.) *Bueno* is used to mark correction of the height to 150 cm, in response to S1's comment that a penguin can't be that tall. As has been seen for its use as a preface to a response and a reorientation, *bueno* here encodes agreement with the interlocutor, but only partial agreement, as the prior discourse needs to be modified, or corrected.

The following is another example of *bueno* marking a correction, from an interview conducted in Australia with two Colombian males. Here, Alonso is reading a question regarding punctuality from a question/topic sheet he was given.

- (13) (WR:2.2.2)
 ALONSO: <READING Hay situaciones en las cuales, la puntualidad sea más importante que otras READING>? <READING Are there situations in which, punctuality would be more important than in others READING>?
 Si es una cita para encontrarse con una niña. If it's with a date to meet up with a girl.
Bueno, **Bueno,**
 quiero decir una mujer no? I mean a woman don't I?
 No es una niña. It's not a girl.
 ... **Bueno,** **Bueno,**
 nosotros decimos niña no? we say girl don't we?

The speaker corrects himself twice here; the first time from *niña* 'girl' to *mujer* 'woman', and the second time to correct his statement *No es una niña* 'It's not a girl', because to a Colombian, it is in fact *una niña* 'a girl'; that is, that's what Colombians call women. Note that in this example, as in (12), some kind of acceptance of what is to be corrected is implied, and thus, just as has been seen above, *bueno* is marking a partial modification.

In these examples, *bueno* can be seen to have the same meaning as evident in its use to mark a response and a reorientation. It indicates an assumption that the speakers are in agreement ('we think the same') (thus in (13), for example, 'we all know that I don't mean girl'), a positive evaluation of this ('this is good'), and a qualification, which is the correction ('there is something more to say').

5. Conclusion

The four functions listed above of marking acceptance, prefacing a response, indicating a reorientation in the conversation, or a correction, can be accounted for under the two formulae proposed, reproduced below.

*bueno*1

1. you say something
2. I say: this is good

*bueno*2

1. someone says something
2. I say: I think the same as you
3. this is good
4. there is something more to say

Definition (1) captures the meaning of *bueno* used to mark acceptance be that of an offer, a suggestion, or of information. It implies that what the interlocutor has said is accepted without qualification. This is reflected in component two, which states that *bueno* in this use makes a positive comment on something someone else

has said, without encoding anything more about the speaker's attitude to that proposition.

Definition (2) accounts for the use of *bueno* as a marker of response, reorientation or correction. This use of *bueno* functions as a comment on a prior utterance or stretch of discourse; it indicates that the speaker assumes agreement between their interlocutors, that they view this agreement positively, but that the agreement is only partial, and some modification is needed.

These two meanings are related, in that both are a response to prior discourse, encoding a positive evaluation. The main difference between them is that *bueno2* encodes only partial acceptance, indicating that there is something more to be said. It may be the case that the notion of 'reluctance' discussed in the literature in the use of *bueno* indicating acceptance (see section 4.1 above) is related to, or even has its basis in, this notion of 'something more to say', evident in *bueno2*.⁶ As this notion is not always encoded, however, it cannot be attributed to the meaning of *bueno*, and for this reason, I have posited polysemy.⁷

In closing, I would like to consider what role, if any, *bueno* may play in enacting the norms and values of the cultures in which it is used. The literature on Latin American culture characterizes it as a culture in which the maintenance of harmonious interpersonal relations, the avoidance of personal conflict and the recognition of similarity with others are of central importance (Diaz-Guerrero and Szalay 1991, Fitch 1989, Markus and Kitayama 1991, Triandis et al. 1984). From the analysis presented here, *bueno* can perhaps be seen to be one way in which these cultural values are realized communicatively. It is a Marker which can be used to minimize conflict in conversation, in that, by first indicating acceptance of what another has said, an upcoming qualification is mitigated, or softened. It is also a way of stressing shared ground between the interlocutors, by acknowledging the validity of others' contributions, before outlining potential differences of opinion. Although a study of just one DM is of course very limited in terms of the kinds of cultural insights it can offer, this brief account suggests that, as has been argued by Wierzbicka (1994), a detailed study of a range of DMs in a given language may prove to be a valuable resource for tapping into cultural values, and how they are realized in conversation.

Appendix

Transcription Conventions (Du Bois et al. 1993:88ff)

.	final intonation contour	...(N)	long pause
,	continuing intonation contour	...	medium pause (> 0.7 secs)
?	appeal intonation contour	..	short pause (about 0.5 secs)
!	emphasis	@	one syllable of laughter
--	truncated intonation contour	((WORDS))	researcher's comment
-	truncated word	<VOX word VOX>	word pronounced with a certain voice quality
=	lengthened syllable	<S1 word S1>	speaker attribution in reported speech
[]	speech overlap	<READING word READING>	speaker is reading aloud
[2 2]	(used to distinguish consecutive overlaps)		

Notes

¹ I would like to thank Alan Baxter, Hilary Chappell, Tim Curnow, Carmen Fought, Barbara Kelly and Ivo Sanchez for help with this paper, in its final and earlier versions, and Sandy Thompson for advice on issues concerned with the treatment of Discourse Markers. I would also like to thank the audience at the presentations of this paper at UCSB, and BLS 24, for their valuable comments. And I would like to thank my research assistants in Colombia, Maria Elena Rendon and Marianne Dieck, for the data they collected, and for help with the transcriptions.

² The brackets here indicate that this Intonation Unit overlaps with a prior one, not reproduced here as it is not relevant to the discussion.

³ Numbered brackets are used when overlaps occur over consecutive Intonation Units, to indicate what overlaps with what. Thus, this laughter overlaps with Angela's *Ah, bueno* 'Oh, *bueno*'.

⁴ Presumably Santi was thinking of something along the lines of looking at the number of rings a tree has (they have just been talking about how large trees get as they get old), although precisely at this point, his mobile phone rings, and the conversation ends.

⁵ *Eso* 'that' here refers to another element of the policy that Sara had been describing, when Angela asked her about the cost, and not to the cost itself.

⁶ A study of the grammaticalization of this Marker could help establish whether this is the case.

⁷ The polysemy seems to be reflected in the prosody, but this remains to be looked at in detail

References

- Allende, Isabel. 1985. *La casa de los espíritus* ('The house of the spirits'). Buenos Aires: Editorial Sudamericana.
- Bauhr, Gerhard. 1994. Funciones discursivas de *bueno* en Español moderno ('Discourse functions of *bueno* in modern Spanish'). *Lingüística Española Actual* 16.1:79–124.
- Díaz-Guerrero and Lorand B. Szalay. 1991. *Understanding Mexicans and Americans: Cultural perspectives in conflict*. Cognition and Language Series. New York: Plenum Press.
- Du Bois, John, Stephan Schuetze-Coburn, Susanna Cumming, Danae Paolino. 1993. Outline of discourse transcription. In: Jane A. Edwards, Martin D. Lampert, eds. *Talking data: Transcription and coding in discourse*, 45–89. Hillsdale, NJ: Lawrence Erlbaum Assoc.
- Fitch, Kristine. 1989. *Communicative enactment of interpersonal ideology: Personal address in urban Colombian society*. Unpublished PhD thesis, University of Washington.
- Fraser, Bruce. 1990. An approach to discourse markers. *Journal of Pragmatics* 14:383–395.
- Fuentes, Catalina Rodríguez. 1993. Comportamiento discursivo de *bueno*, *bien*, *pues bien* ('Discourse behavior of *bueno*, *bien*, *pues bien*'). *Estudios de Lingüística*, Universidad de Alicante 9:205–221.
- Goddard, Cliff. 1989. Issues in Natural Semantic Metalanguage. *Quaderni di Semantica* 10.1:51–64.

- Goddard, Cliff. 1994. Semantic theory and lexical universals. In Cliff Goddard and Anna Wierzbicka, eds. *Semantic and lexical universals*, 7–29. Amsterdam: John Benjamins.
- Jucker, Andreas. 1993. The discourse marker *well*: A relevance theory account. *Journal of Pragmatics* 19:435–452.
- Lenk, Uta. 1998. *Marking discourse coherence: functions of discourse markers in spoken English*. Tübingen: Gunter Narr Verlag.
- Markus Hazel and Shinobu Kitayama. 1991. Culture and the self: Implications for cognition, emotion and motivation. *Psychological Review* 98.2:224–253.
- Schiffrin, Deborah. 1987. *Discourse markers*. Cambridge: Cambridge University Press.
- Triandis, Harry, Gerardo Marín, Judith Lisansky and Hector Betancourt. 1984. *Simpatía* as a cultural script of Hispanics. *Journal of Personality and Social Psychology* 47.6:1363–1375.
- Wierzbicka, Anna. 1972. *Semantic primitives*. Frankfurt: Athenäum.
- Wierzbicka, Anna. 1980. *Lingua mentalis: The semantics of natural language*. Sydney: Academic Press.
- Wierzbicka, Anna. 1994. 'Cultural scripts': A new approach to the study of cross-cultural communication. In: Martin Pütz, ed. *Language contact and language conflict*, 69–87. Amsterdam: John Benjamins.
- Wierzbicka, Anna. 1996. *Semantics: Primes and universals*. Oxford University Press: Oxford.

A Theory of Agreement and Disagreement

Stephen Wechsler and Larisa Zlatić
University of Texas at Austin

1. Introduction. A striking fact emerging from studies of agreement is that the expected agreement relations often break down, resulting in 'mixed' or 'hybrid' agreement, or, as we call it, *disagreement*. For example, certain common nouns trigger feminine agreement on modifiers but masculine on predicates and pronouns; or a noun may trigger masculine agreement, but is itself inflected in the pattern normally reserved for feminines. In this paper we propose a theory of agreement which explains this phenomenon of disagreement and predicts which classes of disagreement are possible, and which are impossible. We apply this theory to Serbo-Croatian (SC), and also show briefly that this theory predicts the main typological generalizations due to Greville Corbett (1983; 1991) and confirmed by further studies such as Barlow (1988). Our descriptive framework is Head-Driven Phrase Structure Grammar (HPSG, Pollard and Sag, 1994).

2. Two types of agreement: CONCORD and INDEX. Regular agreement in SC is illustrated in (1):

- (1) Ova stara knjiga stalno pada.
this.F.SG old.F.SG book(fem.) always falls.3.SG
- Molim vas podignite je.
please you pick.2pl it.F.SG

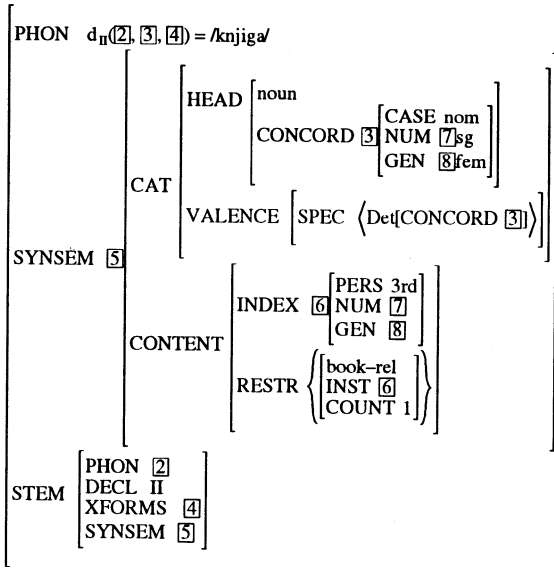
'This old book keeps falling. Please pick it up.'

The noun *knjiga* 'book' triggers feminine singular (f.sg) agreement on determiners (*ova*) and modifiers (*stara*), third person singular (3.sg) on finite predicates (*pada*), and third person feminine singular (3.f.sg) on coreferential pronouns (*je*). Participles also show gender and number agreement (see Section 7 below).

Following the basic lead of Kathol (to appear), Pollard and Sag (1994), and, in different variants, numerous other researchers, we distinguish two agreement features on the noun, called CONCORD (Kathol's AGR) and INDEX. CONCORD is a head feature, while INDEX is part of the noun's semantic content, as illustrated by the HPSG lexical sign for *knjiga* 'book.nom.sg' in (2). The CONCORD feature effects determiner and modifier agreement as a side-effect of functor-argument combination: the CONCORD value of the noun unifies with that of its specifier, as shown in (2); similarly, an attributive adjective specifies the features of the N' it modifies. INDEX agreement, on the other hand, arises through the sharing of referential indices between pronoun and antecedent. As shown in (2), the index contains person, number, and gender (PNG) features. Thus the matching condition on pronouns is a side-effect of the coreference ('coindexation') mechanism (see Pollard and Sag, 1994). The PNG features are grammaticalizations of referential anchoring conditions— PERSON: speaker → 1st, addressee → 2nd, other → 3rd; NUMBER: aggregate → pl, other → sg; GENDER: male → masc, female → fem, other → neut — but cannot be reduced

to semantics (cp. grammatical gender, pluralia tantum, etc.). In addition, a finite predicate specifies the INDEX features of its nominative subject, enforcing agreement as a side-effect of functor-argument combination.

- (2) Lexical sign for the feminine noun *knjiga* 'book.nom.sg':



We now motivate this split between concord and index, first theoretically and then empirically.

First, the feature sets differ: *person*, number, and gender for index but *case*, number, and gender for concord. Second, the syntactic domains differ. As explained above NP's and pronouns have referential indices, hence index agreement. Verb agreement, which is derived historically from incorporated pronouns, also involves index agreement. Concord agreement is typically (but not necessarily) NP-internal. Unlike index agreement, it includes non-referential elements such as determiners and adjectives.

Third, and most relevant here, the relation to other lexical properties differs for index and concord respectively. As we will show, INDEX features are closely related to semantics, while CONCORD features are closely related to phonology (declension class). In what follows, we will flesh out these index-semantics and concord-phonology relations. Then we turn to those exceptional noun classes where these relations break down.

3. SC gender and number constraints.

3.1. Phonology-concord constraints. The concord-phonology relation is explicated via the notion of declension class. We define a *declension* as a function d from stem phonology Φ_{stem} , CONCORD value C , and exceptional forms X (and animacy A , for d_I), to the inflected noun's phonology Φ_{noun} .

Following traditional descriptions, SC has three declensions, which we designate d_I , d_{II} , and d_{III} . Here is a partial statement of the functions:

$d_I(\psi, c, x, a) = \psi$ if $c = m.sg.nom$ or $[c=m.sg.acc \ \& \ a=-anim] \ \& \ x = \{ \}$
 $d_I(\psi, c, x, a) = \psi + /a/$ if $c = m.sg.gen$ or $[c=m.sg.acc \ \& \ a=+anim] \ \& \ x = \{ \}$
 $d_I(\psi, c, x) = \psi + /o-e/$ if $c = nt.sg.nom \vee acc \ \& \ x = \{ \}$
 ...etc.

Sample output forms are shown in this Table I at the end of the paper. There are other ways to set up the declension classes. The crucial distinction between declension and gender is that syntactic processes refer to gender but not declension; hence in (2) we make gender but not declension a head feature.

Concord and phonology are systematically related since concord *number* is an argument of all Declension functions, while concord gender is an argument of Declension class I. Moreover concord gender strongly influences assignment to declension class. As illustrated in Table I, declension class I nouns are either masculine or neuter, while Class II and III nouns are feminine. We capture this rule by positing a subsort of *noun-sign* called *dec=con* and formulating the following constraint:

$$(3) \ dec=con: \begin{bmatrix} PHON & d_I \\ CONCD & m \vee n \end{bmatrix} \vee \begin{bmatrix} PHON & d_{II} \vee d_{III} \\ CONCD & f \end{bmatrix}$$

Most nouns belong to the sort *dec=con*; but some exceptional ones discussed below do not.

Evidence that (3) is part of SC competency grammar comes from *inquorate* nouns (Corbett, 1991:173), nouns with different gender in singular and plural. Interestingly, declension class and gender follow *dec=con* in both singular and plural. Some inquorates have a different stem in singular and plural, such as *ok-o* ('eye' nt., d_I) ~ *oč-i* ('eyes' f., d_{III}). Note that the singular, being neuter, is declined with the Class I pattern, while the plural, being feminine, takes the Class III declension. Another example is *mač-e* ('kitten' nt., d_I) ~ *mačič-i* ('kittens', m., d_I). Certain collective nouns like *deca* 'children' (see Section 6 below) are also inquorates:

	<u><i>det-e</i> ('child')</u>	<u><i>dec-a</i> ('children')</u>
DECLENSION:	d_I	d_{II}
CONCORD:	nt.sg	f.sg
INDEX:	nt.sg	nt.pl

Once again, declension and concord covary in lockstep across the singular/plural distinction, following the constraint in (3): *deca* 'children' has feminine singular concord and is declined in class II. Moreover, the index features do not vary, as we demonstrate below (Section 6). This pattern, where concord but not index covaries with declension, is consistent across the SC language.

Note also the phonologically defined subclass of inquorates comprising /o/-final loans such as *auto*, *radio*, *kino*, and *torzo* (masculine in the singular, neuter in the plural). Summarizing, our basic observations are (i) the

Concord(gender)-Declension rule (3) is observed even in the face of irregularity across singular/ plural; and (ii) Concord (gender) irregularity is related to phonology.

3.2. Concord-index constraints. In regular nouns such as *knjiga* 'book' in (2) above, the number and gender values of the concord match those of the index. We posit that such nouns belong to a subsort of *noun-sign* called *con=ind* and formulate the following constraint:

$$(4) \quad con=ind: \left[\begin{array}{c} \text{CONCORD} \left[\begin{array}{c} \text{NUM } \boxed{1} \\ \text{GEN } \boxed{2} \\ \text{CASE} \end{array} \right] \\ \text{INDEX} \left[\begin{array}{c} \text{NUM } \boxed{1} \\ \text{GEN } \boxed{2} \\ \text{PERS} \end{array} \right] \end{array} \right]$$

Note that the boxed numerals ('tags') indicate structure sharing, each numeral representing a single node in a directed acyclic graph. Hence a noun belonging to this class has only one number value and one gender value (not two of each). Each of these values can be 'accessed' in two ways, via concord and index.

3.3. Index-semantics constraints. There are two constraints relating index to semantics, one for number and one for gender. Taking number first, in the normal case count nouns with aggregate reference are plural, while other nouns (mass and non-aggregate count) are singular. We formulate this with a feature COUNT declared only by count nouns, whose value encodes the cardinality of the reference. Here we need only consider two COUNT values, '1' (cardinality one) and '>1' (cardinality greater than one). The constraint is given in (5), where *ind=sem_N* is yet another subsort of *noun-sign*:

$$(5) \quad ind=sem_N: \left[\begin{array}{c} \text{INDEX } \boxed{2}pl \\ \text{RESTR } \left\{ \left[\begin{array}{c} \text{INST } \boxed{2} \\ \text{COUNT } >1 \end{array} \right] \dots \right\} \end{array} \right] \vee \left[\begin{array}{c} \text{INDEX } \boxed{1}sg \\ \text{RESTR } \neg \left\{ \left[\begin{array}{c} \text{INST } \boxed{1} \\ \text{COUNT } >1 \end{array} \right] \dots \right\} \end{array} \right]$$

A BLS audience member raised the problem of applying the notion of (non-) aggregate reference to nouns like *flock*, *committee*, etc. It is important to distinguish *reference* (a relation between language and an item in the world) from the *referent* itself (an item in the world). For us a noun with 'aggregate reference' is one such that its referent consists of more than one item of the sort described by that noun's stem: *apples* denotes more than one apple so it has aggregate reference. But *flock* does not denote more than one flock, so it has non-aggregate reference. Pluralia tantum nouns like *scissors* clearly also lack aggregate reference. It is this notion of aggregate reference that the COUNT feature encodes.

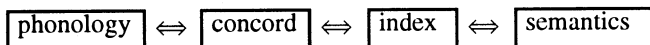
Turning to gender, SC nouns restricted to female (/male) referents have feminine (/masculine) index, a constraint apparently valid for Slavic generally. Hence nouns denoting 'man', 'husband', and 'bull' are masculine, while 'woman', 'wife', and 'hen' are feminine. This is formulated here:

(6) $ind = sem_G$:

$$\left(\left[\text{RESTR } \{ [female(\mathbb{I})] \dots \} \right] \Rightarrow [\text{INDEX } \mathbb{I} fem] \right) \wedge \left(\left[\text{RESTR } \{ [male(\mathbb{I})] \dots \} \right] \Rightarrow [\text{INDEX } \mathbb{I} masc] \right)$$

It is important to note that this constraint is ‘vacuously satisfied’ by any noun which is not sex-specific, either because it has no sex at all (‘book’, ‘table’, etc.) or because it is unspecific (‘child’, ‘human’, ‘president’, etc.)

Collecting the constraints in (3), (4), (5), and (6), we arrive at a ‘chain’ of dependencies between features of a common noun, where ‘ \Leftrightarrow ’ represents one or more constraints.



This chain can be broken in any of the three places depending on which constraint is violated. Technically the constraints are inviolable, but we capture such apparent violations through the system of sorts: for example, a noun violating the concord-index constraint simply does not belong to the subsort $con=ind$, and so on for the others. In the remainder of the paper we illustrate these three types of violations.

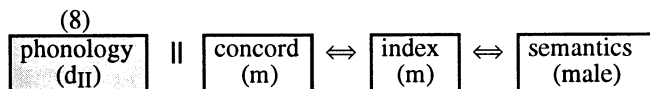
4. Phonology-concord mismatches. As our first example of a mismatch between phonology (i.e. declension) and concord, certain male names in /-a/ such as *Steva* and *Mika* trigger masculine agreement but nonetheless belong to declension II.

(7) Vratio mi je ovaj ludi
 returned-1SG me AUX.3SG this-NOM.M.SG crazy-NOM.M.SG

Steva violinu koju sam mu pozajmio.
 Steve(m.sg).NOM violin-ACC which AUX him-DAT.M.SG loaned

‘This(M) crazy(M) Steve returned to me the violin which I loaned him(M).’

Hence these names violate constraint (3). We show this situation schematically as follows:



Also certain nouns such as *sudija* ‘judge’, *sluga* ‘slave’, *gazda* ‘master’, *mušterija* ‘customer’, *kolega* ‘colleague’, etc. trigger feminine or masculine agreement depending on whether they refer to females or males:

(9) a. Taj stari sudija_i je dobro sudio. On_i...
 that.M old.M judge AUX well judged.M he.M
 'That old (male) judge judged well.'

b. Ta stara sudija_i je dobro sudila. Ona_i...
 that.F old.F judge AUX well judged.F she.F
 'That old (female) judge judged well.'

Since *sudija* belongs to declension class II, it violates (3) when referring to a male judge, as in (9)a.

How are we to capture the covariation between gender and sex shown in (9)a,b? The problem with assuming 'semantic agreement', where the sex (semantics) directly determines agreement, is that the syntactic mechanism of agreement would not be uniform: clearly sex does not determine agreement in (1) above, for example. Our constraints explain this covariation if we assume that *sudija* is sex-specific but disjunctive, as shown by the RESTRICTIONS value in (12)a below: it is either restricted to being male or restricted to being female. The agreement facts follow from a chain of deductions. Taking the male disjunct, by the constraint *ind=sem_G* (see (6) above), *male(x)* entails [INDEX masc]; hence, by *con=ind* (see (4)), we also get [CONCORD masc]. (Constraint *dec=con* is violated, as noted). Mutatis mutandis for the female disjunct. In contrast, a noun like *pas* 'dog' is grammatically masculine and has no sex specification at all:

(10) taj verni pas
 that.M loyal.M dog
 'that loyal (male or female) dog'

The noun *mušterija* 'customer' has an optional disjunctive specification, hence it is midway between 'judge' and 'dog'. It exhibits dog-like behavior in (11)a (grammatically feminine but unspecified for sex) but is judge-like in (11)b (male and masculine). The female/feminine variant of (11)b overlaps with (11)a.

(11) a. Ta mušterija je došla.
 that.F customer AUX came.F
 'That (male or female) customer came.'

b. Taj mušterija je došao.
 that.M customer AUX came.M
 'That (male) customer came.'

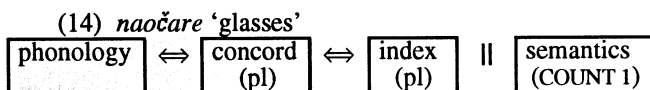
We summarize the three types as follows:

- (12) a. *sudij*-. {judge(x)} ∪ {male(x) ∨ female(x)}
 b. *pas*-. {dog(x)}
 c. *mušterij*-. {customer(x)} ∪ {male(x) ∨ female(x)}

5. Index-semantics mismatches. Pluralia tantum nouns, when used in the non-aggregate sense (hence specified [COUNT 1]), violate constraint (5) *ind=sem_N*. Examples include *kola* 'car', *vrata* 'door', *ledja* 'back', *makaze* 'scissors', *naočare* 'glasses', and *novine* 'newspaper'.

- (13) Ove naočare_i su nove. One_i ...
 this.PL glasses are.3PL new-PL 3PL...
 'These glasses_i are new. They_i...'

This example is ambiguous, as in English, referring to one pair of glasses or more than one pair. On the former reading it violates (5), as shown here schematically:



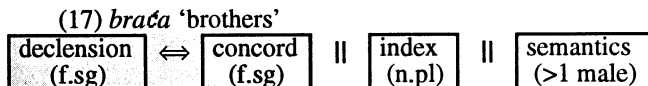
Violations of (6) *ind=sem_G* are rare, but *braća* 'brothers' is one case. While male-only, it has a nt. pl index:

- (15) ... braća_i ... Ona_i su došla
 ... brothers ... they.N.PL AUX.3.PL came.N.PL

Also, adjectives and determiners show f.sg:

- (16) ovu dobru braću
 this-ACC.F.SG good-ACC.F.SG brothers-ACC

Hence *braća* violates two constraints:



6. Index-concord mismatches. Collective nouns like *deca* 'children', *braća* 'brothers', and *gospoda* 'gentlemen' violate (4) *con=ind* (see Corbett, 1983). Such nouns trigger feminine singular on modifiers, plural on finite verbs and auxiliaries, and neuter plural on pronouns:

- (18) Posmatrali smo ovu dobru decu_i.
 watched.1.PL AUX this.F.SG good.F.SG children.ACC(f.sg[nt.pl])

Ona_i su spavala.
 they-N.PL AUX.3.PL slept-NT.PL

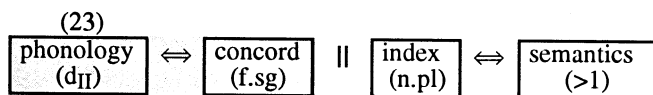
'We watched those good children_i. They_i slept.'

- (19) ...deca_i... Mi smo ih_i videli.
 ...children... we Aux-PL them-ACC.PL saw
 '...children_i... We saw them_i.'
- (20) Ta dobra deca su došla.
 that.F.SG good.F.SG children(F.SG) Aux-PL came-PPRT.F.SG/N.PL
 'Those good children came.'
- (21) Ta dobra deca dolaze.
 that.F.SG good.F.SG children(F.SG) come.3.PL
 'Those good children came.'

We posit the following CONCORD and INDEX gender/number features for *deca*:

- (22) *deca*: $\begin{bmatrix} \text{CONCORD} & \text{fem.sg} \\ \text{INDEX} & \text{nt.pl} \end{bmatrix}$

This situation is shown schematically as follows:



Note that the concord and phonology are still mutually consistent, as are index and semantics. Regarding the former, these nouns are inflected in the traditional second declension class, the norm for feminine (concord) nouns. Regarding the latter, *deca* 'children' is an aggregate and has a plural index, so it satisfies *ind=sem_N*, and it is not sex-specific so it vacuously satisfies *ind=sem_G*.

Relative pronouns show an interesting mixture of sensitivity to both concord and index features (Corbett, 1983). The nominative relative pronoun is *koja*, the nt.pl form.

- (24) deca koja su / *je tada bila...
 children who-N.PL Aux-PL/*SG then were...
 'the children who were...'

But an accusative relative pronoun takes a f.sg form, and oblique cases can take either f.sg (*koje*) or plural (*kojih*) form.

- (25) deca koju sam video
 children who-ACC.F.SG Aux-1.SG saw
 'the children whom I saw'
- (26) deca koje/kojih se svi plače
 children who-GEN.F.SG./GEN.PL REFL all fear
 'children whom everyone fears'

We analyze these facts as follows. Relative pronouns, being heads of N'-modifiers, show CONCORD agreement (like other modifiers); being bound pronouns, they also show INDEX agreement. But the relative pronouns are not

specified for all features: most of the plurals do not mark (CONCORD) gender, and nominative and accusative relative pronouns are unmarked for (INDEX) PNG features. Details are omitted for lack of space (see Wechsler and Zlatic, 1997).

Relative clauses introduced by the complementizer *što* employ ordinary pronoun forms functioning as 'resumptive pronouns'. Since they are ordinary pronouns they show index agreement, irrespective of case.

- (27) *moja deca, što ih/*je svi hvale.*
 my children, that 3.ACC.PL/*SG all praise
 'my children, whom everyone praises'

- (28) *moja deca, što ih/*je se svi plače.*
 my children, that 3.GEN.PL REFL all fear
 'my children, whom everyone fears'

7. Participles and predicate adjectives. Participles and predicate adjectives agree in gender and number with a nominative subject. But do they show CONCORD or INDEX agreement? That is, do they pattern with modifiers or with pronouns, when those differ? There is at least some evidence for INDEX agreement. The plural form *vladike* 'bishops' belongs to a group triggering feminine concord (although bishops are male). In this dialectal example the participle shows masculine rather than feminine agreement:

- (29) Jekavski dialect (cited in Marković, 1954)
ove dobre vladike su došli.
 these-F.PL good-F.PL bishops-F Aux came-M.PL
 'These good (male) bishops came.'

Coordination resolution also provides some evidence for index agreement. When items of different genders are coordinated, the coordinate structure triggers masculine plural agreement; when like gender is coordinated, the gender of the conjuncts is triggered. Coordinating *deca* with the female name *Jelena* in (30) yields masculine plural agreement, indicating mixed gender (the subscript in the gloss indicates index value; the non-subscripted item is concord value). Concord values on the conjuncts are both feminine singular, while index values are mixed (f. and nt.), so this means that the index value must be the relevant one. (This argument is due to Corbett 1983). Secondly, when *deca* is conjoined with the neuter plural NP *ta čudovišta* 'those monsters', neuter plural is favored, again suggesting index rather than concord agreement (31).

- (30) *Jelena i deca su došli/*došle.*
*Yelena(F.SG[F.SG]) & children(F.SG[N.PL]) Aux.PL came.M.PL/*F.PL*
 'Yelena and the children came.'

- (31) *Ta deca i ta čudovišta su se lepo igrala/?igrali.*
 that children & those monsters Aux.PL REFL well played-N.PL/?M.PL
 (F.SG[N.PL]) (N.PL[N.PL])
 'Those children and those monsters played well.'

8. **Conclusion.** We summarize our findings in Table II below (not all of the examples in the table are discussed above). The first four rows represent regular agreement; the remaining rows show disagreement, where the break in the pattern is indicated by the double horizontal line and the shift from shaded to non-shaded cells. Since contiguous cells in a row are related by constraints, we predict exactly the three types of mismatch pattern shown. Many hypothetically possible mismatch types are unattested: e.g., where declension and index correlate but concord is the odd man out. That situation is unexpected since according to our theory the correlation between declension and index is not direct but rather mediated by concord.

Our results resemble Greville Corbett's cross-linguistic Agreement Hierarchy (Corbett, 1983:10; 1991:226):

attributive < predicate < relative pronoun < personal pronoun

←syntactic agreement—

—semantic agreement→

As we move rightwards along the hierarchy, the likelihood of semantic agreement will increase monotonically (that is, with no intervening decrease).

For us the relative ranking of attributives, relative pronouns, and personal pronouns depends on whether they show concord or index agreement. As noted in Section 2 above, personal pronouns are referential while attributive modifiers are not, hence the former but not the latter show index agreement. Relative pronouns occupy a middle ground since they head modifiers but are also pronominal (Section 6). Hence this part of the hierarchy is explained.

As for predicate agreement, we make no general prediction but perhaps expect some relation to the provenience of the agreement morphology (those derived from incorporated pronouns may show index agreement); however, some predicates, such as adjectives, function alternatively as modifiers and so might be expected to show concord. In fact, SC finite predicates clearly show index agreement, and for participles and predicate adjectives the evidence is less clear but they also seem to show index agreement. If anything, this should place them to the right of relative pronouns on the hierarchy. So the extra detail in the hierarchy which we fail to predict appears not to be justified by the data.

Corbett (1998) and Barlow (1988) argue against the sort of 'split' theory of agreement advocated by us (Concord vs. Index) and by numerous others (e.g. Bresnan and Mchombo, 1987). Space does not permit a full discussion here, but the data cited do not appear to pose problems for the present account. Most of the problematic cases involve a single agreement type (e.g. pronouns or finite verbs) alternating between 'grammatical' and 'semantic' agreement. For example, pronouns taking *deca* 'children' as antecedent can take either neuter plural form reflecting the 'grammatical gender' of *deca* (see (18)-(19) above) or alternatively masculine plural, reflecting 'semantic agreement'. But our index agreement cannot be equated with semantic agreement in this sense. Both concord and index are lexical (hence 'grammatical') features of common nouns. But the use of masculine plural pronouns for *deca* 'children' reflects the fact that masculine plural is used for groups which are male or of mixed or unknown sex in SC. The latter type does not involve common noun features and thus falls outside the scope of the present paper, but is discussed in detail in Wechsler and Zlatić (to appear).

Appendix.

	Class I		Class II	Class III
	'window' (m)	'village' (nt)	'woman' (f)	'thing' (f)
Singular				
Nominative	prozor	sel-o	žen-a	stvar
Accusative	prozor	sel-o	žen-u	stvar
Genitive	prozor-a	sel-a	žen-e	stvar-i
Dative/Locative	prozor-u	sel-u	žen-i	stvar-i
Instrumental	prozor-om	sel-om	žen-om	stvar-i
Vocative	prozor-e	sel-o	žen-o	stvar-i
Plural				
Nominative	prozor-i	sel-a	žen-e	stvar-i
Accusative	prozor-e	sel-a	žen-e	stvar-i
Genitive	prozor-a	sel-a:	žen-a	stvar-i:
Dative/Inst/Loc	prozor-ima	sel-ima	žen-ama	stvar-ima

Table I. Serbo-Croatian Declension Classes

example	DECL	CONCD	INDEX	SEM
<i>knjiga</i> 'book'	II(f)	f	f	—
<i>rad</i> 'work'	I(m)	m	m	—
<i>žena</i> 'woman'	II(f)	f	f	F
<i>mužu</i> 'husband'	I(m)	m	m	M
<i>mušterija</i> 'customer' (when used for male)	II(f)	m	m	M
<i>sudija</i> 'judge' (when used for male)	II(f)	m	m	M
<i>Steva</i> 'Steve'	II(f)	m	m	M
<i>Jovanova</i> 'John's-f.sg'	II(f)	f	m	M
<i>Marijin:</i> 'Mary's-m.sg'	I(m)	m	f	F
<i>deca</i> 'children'	II(f.sg)	f.sg.	n.pl	>1
<i>gospoda</i> 'gentlemen'	II(f.sg)	f.sg	m.pl	M
<i>braća</i> 'brothers'	II(f.sg)	f.sg	n.pl	M, >1
<i>makaze</i> 'scissors' (COUNT 1)	pl	pl	pl	1

(m = masculine, f = feminine, M = male, F = female)

Table II. Declension, CONCORD, INDEX, and semantic features of sample Serbo-Croatian common nouns

References.

- Barlow, Michael. 1992. *A Situated Theory of Agreement*. (1988 Stanford PhD Dissertation). New York: Garland Publishing.
- Bresnan, Joan, and Sam A. Mchombo. 1987. Topic, pronoun, and agreement in Chichewa. *Language* 63 (4):741-782.
- Corbett, Greville. 1983. *Hierarchies, Targets and Controllers: Agreement Patterns in Slavic*. London: Croon Helm.
- Corbett, Greville. 1991. *Gender*. Cambridge: Cambridge University Press.
- Corbett, Greville G. 1998. Agreement in Slavic. *Position paper for the conference on Comparative Slavic Morphosyntax, Bloomington, Indiana, 5-7 June, 1998*.
- Kathol, Andreas. to appear. Agreement and the Syntax-Morphology Interface in HPSG. In *Readings in HPSG.*, edited by R. Levine and G. Green. Cambridge: Cambridge University Press.
- Marković, Svetozar. 1954. O kolebljivosti slaganja u rodu kod imenica čiji se prirodni i gramatički rod ne slažu (i o rodu ovih imenica). *Pitanja književnosti i jezika* 1:87-110.
- Pollard, Carl, and Ivan Sag. 1994. *Head Driven Phrase Structure Grammar*. Stanford and Chicago: CSLI Publications and University of Chicago Press.
- Wechsler, Stephen and Larisa Zlatic. 1997. (Dis)agreement in Serbo-Croatian. Manuscript. University of Texas at Austin.
- Wechsler, Stephen and Larisa Zlatic. to appear. Agreement in Discourse. In *Proceedings of the Conference on the Structure of Non-narrative Texts, University of Texas, Austin, February 13, 1998*.

Lexical and Grammatical Universals as a Key to Conceptual Structures

Anna Wierzbicka
Australian National University

1. Introduction.

According to Leibniz (1981[1703]:333), 'languages are the best mirror of the human mind, and (...) a precise analysis of the significations of words would tell us more than anything else, about the operations of the understanding'. Over the centuries, and especially in this century, many students of mind have repeated Leibniz's observations, often, however, changing his plural form ('languages') to the singular ('language'). But in fact the profundity of Leibniz's insight lay precisely in that plural.

'Language' in the singular (a particular language) is not a good mirror of 'the human mind'; for languages are diverse and each language reflects the features of the culture associated with it. This is why attention to 'language' (in the singular) as a guide to 'human cognition' has often led various scholars astray, resulting in the absolutization of the conceptual distinctions and patterns suggested to them by their native language.

Language in the singular — English, French or Japanese — is, as Sapir (1949[1929]) said, a guide to social reality, a guide to culture. It is also a guide to the human psyche as culturally constituted, that is, as shaped not only by innate and universal features of 'human nature' but also by the particular features of historically transmitted 'local' cultures (cf. Shweder 1990). It is only 'languages' in the plural which allow us to see and appreciate both the diversity of cultures and what Franz Boas and others have called 'the psychic unity of mankind'. To study the universal aspects of human cognition (and this applies also to human emotion) we need to pay attention to linguistic universals: it is the shared rather than idiosyncratic features of languages which provide a guide to the workings of the generic human mind.

2. The need for a metalanguage.

We cannot adequately study and describe human cognition without an appropriate metalanguage. All attempts to describe aspects of human cognition in ordinary language — for example, in English — lead inevitably to ethnocentric distortions. For every natural language embodies its own 'naïve picture of the world' (Apresjan 1974, 1992), or its own 'Weltanschauung' (Weisgerber 1939), and this includes its own ethnopsychology. If we do not, at the outset, give serious consideration to the question of an appropriate metalanguage, we get trapped in words such as, for example, *mind*, *emotion*, *self*, and so on — words which are continuously used in the literature as if they were culture-free analytical tools, whereas in fact they are cultural artifacts of one particular language and the intellectual tradition associated with it. It hardly needs to be explained to this audience that, for example, the English *mind* does not mean the same as the French *esprit* or the German *Geist*, or that the English *emotion* does not mean the same as the French *sentiment* or the German *Gefühl* (not to mention, for example, Descartes' *les passions de l'âme* or Wilhelm Wundt's *Gemütsbewegungen*). To recall one striking example, pointed out by the psychiatrist Bruno Bettelheim (1983), important aspects of Freud's doctrine were misunderstood in America, despite his tremendous success and popularity there, simply because his key word *Seele* was rendered in the English

translations of his writings as *mind*. (Not that there was a better word to translate *Seele*; but this does not change the fact that the substitution of the English word *mind* for the German word *Seele* in the English version of Freud's writings led to distortion of his ideas).

As a second example of this kind of distortion I would like to quote from a book by the distinguished American psychologist Ernest Becker (1962: 39), entitled The birth and death of meaning:

The question "What fact is the most basic to an understanding of human motivation?" can be answered with just one word: anxiety. Anxiety is a prime mover of human behaviour, and man will do anything to avoid it. It is what he does to avoid it as a child that is most important, and shapes his entire characterological bent. In fact, one is tempted to coin still another definition, and call man "the anxiety-avoiding animal." Man's quest to avoid anxiety not only explains much about motivation; it explains almost everything. It seems unimaginable that college courses purporting to explain human action can be taught without reference to anxiety. But they are.

Freud, who spent a lifetime trying to uncover the mainsprings of motivation, devoted an entire work to the problem of anxiety.

What is most striking about this passage is Becker's substitution of the English word *anxiety* for the German word *Angst* employed by Freud (1963[1916-17]). Clearly, Becker believed that he was in full agreement with Freud, and indeed that he was only developing Freud's ideas. But in fact, he was talking about something else: anxiety, not *Angst* (for full discussion, see Wierzbicka 1998). As a result, Becker's theory of human motivation and human nature misrepresented Freud.

There are few things quite so misleading as to talk about 'human nature' or 'human cognition' (or 'human emotion') simply in English (or in French or in German), without addressing the central issue of a suitable — non-ethnocentric — metalanguage. To quote a comment made recently by Harré and Gillett (1994:6) concerning what they see as mainstream psychology: 'Mainstream psychologists, with their narrow training and cultural isolation, have not generally been aware of how local their common sense systems are'. Since 'common sense' interpretive systems are linked with 'local languages', the use of any 'local' language as an informal metalanguage is bound to lead to ethnocentric bias.

3. Why artificial formalisms cannot elucidate meaning.

But if an ordinary language such as English, in all its culture-specific richness, is not suitable as a metalanguage for studying conceptual structures, neither are artificial languages, which are necessarily incomprehensible until they are translated into natural language, as I will illustrate in a moment. The only satisfactory solution that I see is the use of a metalanguage based on a highly reduced and standardized natural language — not a full-blown natural language such as ordinary English, or French, which is necessarily culture-specific, and not some artificial technical language, which is necessarily incomprehensible (until translated into some natural language), but a natural language trimmed down to its barest essentials: to a few

dozen basic words, and a limited number of fairly simple grammatical constructions.

As I will discuss shortly, Ray Jackendoff's theoretical position is closer to mine than that of many other well-known theoreticians, for he, too, believes in universal conceptual primitives and conceptual grammar. Since, nonetheless, Jackendoff's position differs from mine in some important aspects, I will use some examples from his writings to explain what is specific about my own creed. Consider, for example, Jackendoff's (1991:41) representation of the meaning of the sentence 'The light flashed until dawn':

The light flashed until dawn =

$$\left[\begin{array}{l} +b \\ [\text{DIM } 1d \text{ DIR}] \\ \\ \text{COMP} \left(\left[\begin{array}{l} -b, +i \\ \text{PL} \end{array} \right] \left(\left[\begin{array}{l} +b \\ \text{DIM } 0d \text{ DIR} \\ \text{LIGHT FLASH} \end{array} \right] \right) \right) \\ \\ \text{Sit} \quad \text{BDBY} + [\text{Time DAWN}] \end{array} \right]$$

In my view, a semantic representation of this kind is not very illuminating; and explanations which tell us that 'b' stands for 'bounded', 'i', for 'individual', 'DIR' for 'direction' and 'DIM' for 'dimensionality', does not make it much clearer. In earlier work, for example in his 1990 book Semantic structures, Jackendoff employed a semantic representation closer to natural language; in effect something half-way between artificial notation and 'quasi-English'. Unfortunately, in his more recent work he has shifted to a more abstract feature-based system (as illustrated above), weakening in the process the comprehensibility and verifiability of his analyses. Only by translating formulae of this kind into natural language can one make them comprehensible, to any extent; and until they are comprehensible they are also unverifiable. Commenting recently on the link between the incomprehensibility of artificial semantic formulae and their unverifiability, my colleague Cliff Goddard (1998) recalled for this purpose the famous analysis of *kill* as 'cause to die', proposed in the 1970s by James McCawley and George Lakoff (*kill* = X CAUSE [Y die]). When many linguists and philosophers objected to this analysis, pointing out that one can cause someone's death without actually killing them, McCawley (1973) proposed that there were two 'abstract' semantic components CAUSE₁ and CAUSE₂, which corresponded, roughly, to direct and indirect causation but which could not be identified with the English word cause or with the English expressions 'directly cause' and 'indirectly cause'. Goddard (1998, 3-13) comments:

... this manoeuvre makes the analysis completely unverifiable. Whenever a counter-example is pointed out, the analyst can just say 'Oh, I didn't mean that. That

component is abstract, you see. It doesn't mean the same as any English word'.

On this point, too, I find myself in full agreement with Harré and Gillett (1994:77-78), who insist that in the study of human thought priority must be given to natural language:

Another important consequence of the second cognitive revolution is the priority that must be given to ordinary languages in defining what are the phenomena for a scientific psychology. We will endeavor as far as possible to present and understand cognition in terms of the ordinary languages through which we think, rather than looking for abstract representations of them. That is radical because it resists the idea that a new, formal calculus must be devised to represent thought. Such calculi lie at the heart of the artificial intelligence project, the methodological principles of Chomsky and the transformational grammarians, and the assumption of formalists of all kinds.

I have been arguing for the priority of natural languages for three decades, in direct opposition to transformational grammarians and their descendants, and the whole 'NSM' approach to the study of language and thought is based on it: 'NSM' stands for a 'Natural Semantic Metalanguage', that is to say a metalanguage based directly on natural language.

4. Conceptual primitives.

As Leibniz said in the passage quoted at the outset, the operations of the human mind are reflected particularly clearly in the meaning of words. But to be able to understand the meaning of words — and the workings of the human mind reflected in them — we must also take note of Leibniz's observation that not everything can be explained, and that the value of our explanations depends crucially on how self-explanatory our basic concepts are — our primitives, that is to say, those concepts that we are not going to try to explain.

If nothing could be comprehended in itself nothing at all could ever be comprehended. Because what can only be comprehended via something else can be comprehended only to the extent to which that other thing can be comprehended, and so on; accordingly, we can say that we have understood something only when we have broken it down into parts which can be understood in themselves. (Leibniz 1903[1704]:430; my translation).

Semantics can have an explanatory value only if it manages to define (or explicate) complex and obscure meanings in terms of simple and self-explanatory ones. If we can understand any utterances at all it is only because these utterances are built, so to speak, out of simple elements which can be understood by themselves.

This basic point, which modern linguistics has largely lost sight of, was made repeatedly in writings on language by the great thinkers of the seventeenth

century such as Descartes, Pascal, and Arnauld. For example, Descartes wrote:

Further I declare that there are certain things which we render more obscure by trying to define them, because, since they are very simple and clear, we cannot know and perceive them better than by themselves. Nay, we must place in the number of those chief errors that can be committed in the sciences, the mistakes committed by those who would try to define what ought only to be conceived, and who cannot distinguish the clear from the obscure, nor discriminate between what, in order to be known, requires and deserves to be defined, from what can be best known by itself. (1931[1701]:324)

In my 1996 book Semantics: Primes and Universals, I illustrated this point with a recent discussion of the acquisition by children of the concept IF:

Two prominent researchers into child language and the authors of a very valuable study on the acquisition of meaning, Lucia French and Katherine Nelson (1985:38), start their discussion of the concept *if* by saying: "it is difficult to provide a precise definition of the word *if*". Then, after some discussion, they conclude: "The fundamental meaning of *if*, in both logic and ordinary language, is one of implication".

As I pointed out commenting on this statement, it reflects two common assumptions: first, that it is possible to define all words — including *if* — and second, that if a word seems difficult to define, one had better reach for a scientific-sounding word of Latin origin (such as *implication*). In my view, these assumptions are not only false, but jointly constitute a stumbling block for semantic analysis. One cannot define all words, because the very idea of 'defining' implies that there is not only something to be defined (a definiendum) but also something to define it with (a definiens, or rather, a set of 'definienses').

As I have argued for three decades (following Leibniz, Descartes, and others), the elements which can be used to define the meaning of words (or any other meanings) cannot be defined themselves; rather, they must be accepted as 'indefinibilia', that is, as semantic primes, in terms of which all complex meanings can be coherently represented. A definition which attempts to explain the simple word 'if' via the complex word 'implication' flies in the face of the basic principle of sound semantic analysis put forward more than two millennia ago by Aristotle (1937:141^a):

First of all, see if he [the analyst] has failed to make the definition through terms that are prior and more intelligible. For the reason why the definition is rendered is to make known the term stated, and we make things known by taking not any random terms, but such as are prior and more intelligible ... accordingly, it is clear that a man who does not define through terms of this kind has not defined at all.

The same applies, in my view, to concepts such as ' \pm Bounded', ' \pm Individual', or ' \pm Dimensionality'. It is not clear how the use of 'abstract primitives' can advance our understanding of the meaning of words or of the operations of the human mind. On the other hand, if we assume that concepts such as IF and BECAUSE, or SOMEONE and SOMETHING, are indeed basic and indefinable, we can build our understanding of other, more complex ideas on this basis. And if someone does not want to accept that the notions of IF or SOMETHING are indeed self-explanatory, I hope it will be agreed, at least, that they are more self-explanatory than, for example, 'implication' or 'boundedness'.

I am not saying this in order to criticize Ray Jackendoff, because, as mentioned earlier, in many ways his vision of language and its relation to thought is similar to my own. For Jackendoff, too, argues that there must be stored 'in our heads a finite number of "pieces of thought or simple concepts" plus a set of patterns for putting them together into more and more complex thoughts'; and he, too, calls these simple concepts 'conceptual primitives' and the patterns that combine them 'conceptual grammar' (Jackendoff 1993:188). Rather, I am seeking to highlight a specific feature of the program pursued by myself and my colleagues (see in particular Goddard 1989, 1994, and 1998), a feature which to us is quite crucial: for us, conceptual primitives have to be hooked on to intuitively comprehensible words in natural languages; and postulated conceptual structures have to be hooked on to intuitively comprehensible sentences in ordinary language.

5. The shared core of all languages as a basis for a semantic metalanguage.

Conceptual analysis needs a semantic metalanguage which can be understood directly via natural language, because only formulae written in some version of natural language can be intuitively comprehensible and therefore testable. As Keith Allan (1986:268) put it, any abstract metalanguage is just 'a degenerate form of a natural language'. To understand the formula we have to mentally undo the 'deformation' and go back to ordinary language. At the same time, it has to be a carefully chosen version of natural language — one stripped of all its culture-specific richness and reduced to those essentials which make it isomorphic to all other natural languages.

The point is essential to my argument and so it needs to be explained in some detail. To start with the vocabulary, most words in any given language are specific to this particular language or to a group of languages, and are not universal. For example, neither English nor French has a word with a meaning corresponding exactly to the meaning of the German word *Angst*. At the same time, evidence suggests that all languages have words with meanings corresponding exactly to the meanings of the English words *good* and *bad*, or *big* and *small*. This suggests that the concepts of 'good' and 'bad' (or 'big' and 'small') can be regarded as universal, and can, therefore, be used as elements of a culture-independent semantic metalanguage.

Proceeding on this assumption, we can still write such words in capital letters, (à la McCawley), as GOOD and BAD or BON and MAUVAIS, to indicate that we are using them as elements of a special semantic metalanguage. At the same time (unlike McCawley) we can identify them with the meanings of ordinary English and French words (*good* and *bad*, *bon* and *mauvais*), and require that semantic formulae including these words be testable via natural language.

Since the words of ordinary language are often polysemous, we will

identify the meanings in question by means of certain 'canonical' sentences such as, for example, 'I did something bad' (*j'ai fait quelque chose de mauvais*) or 'something good happened to me' (*quelque chose de bon m'est arrivé*). Proceeding in this way, we can overcome both the incomprehensibility and unverifiability of technical semantic formulae of the kind used, for example, by both McCawley and Jackendoff, and the ethnocentrism of semantic descriptions using a full-blown natural language such as ordinary English or French.

Whether or not all languages do share a minimum of basic concepts and basic syntactic patterns is an empirical question, and one which my colleagues and I have been pursuing on an empirical basis. The results of these investigations have been reported in our two collective volumes Semantic and lexical universals — theory and empirical findings (Goddard and Wierzbicka, eds. 1994) and Meaning and universal grammar (Goddard and Wierzbicka, eds. Forthcoming).

These results tend to confirm the thrust of the centuries of philosophical postulates about "innate ideas" (Descartes), "the alphabet of human thoughts" (Leibniz), the "mid-point around which all languages revolve" (Humboldt 1903-36, v.4:21-23) and the "psychic unity of humankind" (Boas). Our main conclusion is that all languages do indeed appear to share a common core, both in their lexical repertoire and in their grammar, and that this common core can be used as a basis for a non-arbitrary, and non-ethnocentric metalanguage for the description of language and for the study of human cognition and emotion.

This shared lexical core of all languages, as it emerges from empirical investigations, can be summarized in the form of the following table:

TABLE OF CONCEPTUAL PRIMITIVES AND LEXICAL UNIVERSALS

Substantives	<u>I, YOU, SOMEONE,</u> <u>SOMETHING(THING),</u> <u>PEOPLE, BODY</u>
Determiners	<u>THIS, THE SAME, OTHER</u>
Quantifiers	<u>ONE, TWO, SOME,</u> <u>MANY/MUCH, ALL</u>
Attributes	<u>GOOD, BAD, BIG, SMALL</u>
Mental predicates	<u>THINK, KNOW, WANT,</u> <u>FEEL, SEE, HEAR</u>
Speech	<u>SAY, WORD, TRUE</u>
Actions, events, movements	<u>DO, HAPPEN, MOVE</u>
Existence and possession	<u>THERE IS, HAVE</u>
Life and death	<u>LIVE, DIE</u>
Logical concepts	<u>NOT, MAYBE, CAN,</u> <u>BECAUSE,</u> <u>IF</u>
Time	<u>WHEN(TIME), NOW, AFTER,</u> <u>BEFORE, A LONG TIME,</u> <u>A SHORT TIME,</u> <u>FOR SOME TIME</u>

Space	<u>WHERE(PLACE), HERE,</u> <u>ABOVE, BELOW, FAR,</u> <u>NEAR, SIDE, INSIDE</u>
Intensifier, Augmentor	<u>VERY, MORE</u>
Taxonomy, partonomy	<u>KIND OF, PART OF</u>
Similarity	<u>LIKE</u>

FRENCH VERSION (cf. PEETERS 1994):

Substantives	<u>MOI, TOI,</u> <u>QUELQU'UN(PERSONNE),</u> <u>QUELQUE CHOSE(CHOSE);</u> <u>LES GENS (L'HOMME);</u> <u>CORPS</u>
Determiners	<u>CELUI, LE MÊME, AUTRE</u>
Quantifiers	<u>UN, DEUX, QUELQUE,</u> <u>BEAUCOUP, TOUS</u>
Attributes	<u>BON, MAUVAIS, GRAND,</u> <u>PETIT</u>
Mental predicates	<u>PENSER, SAVOIR, VOULOIR,</u> <u>SENTIR (ÉPROUVER),</u> <u>VOIR, ENTENDRE</u>
Speech	<u>DIRE, MOT, VRAI</u>
Actions, events, and movements	<u>FAIRE, ARRIVER(SE PASSER),</u> <u>BOUGER</u>
Existence and possession	<u>IL Y A, AVOIR</u>
Life and death	<u>VIVRE, MOURIR</u>
Logical concepts	<u>NON(NÉ PAS), PEUT ÊTRE,</u> <u>POUVOIR,</u> <u>À CAUSE DE, SI</u>
Time	<u>QUAND(MOMENT),</u> <u>MAINTENANT, APRÈS,</u> <u>AVANT, LONGTEMPS,</u> <u>QUELQUE TEMPS,</u> <u>PEU DE TEMPS</u>
Space	<u>OÙ(ENDROIT), ICI, SOUS,</u> <u>AU-DESSUS, LOIN, PRÈS,</u> <u>À CÔTÉ, DEDANS(DANS)</u>
Intensifier, augmentor	<u>TRÈS(BEAUCOUP),</u> <u>PLUS(D'AVANTAGE)</u>
Taxonomy, partonomy	<u>ESPÈCE, PARTIE</u>
Similarity	<u>COMME</u>

This is, then, — I would argue — what the 'alphabet of human thoughts'

looks like. All complex meanings, in all conceptual domains, can be represented and explained as configurations of these sixty or so fundamental conceptual building blocks. A fuller discussion and justification of this set is given in my 1996 book Semantics: Primes and Universals, and in the 1994 collective volume edited by Goddard and myself.

The importance of empirical cross-linguistic investigations for establishing the true (non-arbitrary) conceptual primitives cannot be over-emphasized. Leibniz's theory of an 'alphabet of human thoughts' (1903[1704]:435) could be dismissed as utopian because he never proposed anything like a complete list of hypothetical primitives. As one modern commentator wrote, having pointed out the difficulties involved in the proposed search:

In these circumstances it is understandable that Leibniz should consistently avoid the obvious question as to the number and type of fundamental concepts. The approach would be more convincing if one could at least gain some clue as to what the table of fundamental concepts might look like. (Martin 1964:25).

The best clues as to what the table of fundamental concepts might look like come from the study of languages. It is precisely the breadth of empirical cross-linguistic data which makes it possible for contemporary linguistics to succeed where philosophical speculation has failed.

6. Universal grammar.

What applies to the universal 'lexicon ('alphabet') of human thoughts' applies also to the universal 'grammar of human thoughts' manifested in universal syntactic patterns. Empirical evidence suggests that despite the colossal variation in language structures, there is also a common core of 'human understanding' relying not only on some shared or matching lexical items but also on some shared or matching grammatical patterns in which shared lexical items can be used. Arguably, this common core defines a set of 'basic sentences' which can be said in any language, and which can be matched across language boundaries, and the grammar of these basic sentences consists in the possible distribution patterns of the 'atomic elements' (that is, the lexical indefinables). To discover those patterns we have to look at the lexical indefinables themselves, to see what their possibilities of co-occurrence are. Therefore, in searching for universal grammatical patterns, we are looking not for universals of form, but rather for universals of combinability.

For example, the indefinable word *happen* makes sense only if it is put in a certain syntactic frame, such as 'something (good/bad) happened', 'something (good/bad) happened to me', or 'something happened in this place'; and the indefinable word *do* makes sense only if it is put in some frame such as 'someone did something', 'someone did something to someone else', 'someone did something to something with something' (instrumental), or 'someone did something with someone' (comitative). Frames of this kind constitute universal valency options of the primitives in question. In positing the elements HAPPEN and DO as innate and universal conceptual primitives, I am also positing certain innate and universal rules of syntax — not in the sense of some intuitively unverifiable formal syntax à la Chomsky, but in the sense of intuitively verifiable patterns of possible combinations of primitive concepts.

For instance, the meaning of the sentences 'I did something' or 'something happened to me' is intuitively clear to any native speaker of English, and cannot be made any clearer by explanations, or by abstract elaborations. In particular, no explanations in terms of 'agents', 'actors', 'volition', 'action', 'deixis', 'self-reference', 'subjects', 'predicates', 'objects', 'clauses', 'deletions', or any other technical terms and theoretical constructs can bring anyone a millimetre closer to understanding this sentence. On the contrary, it is our understanding of technical terms and theoretical constructs (such as, in particular, 'agency') which has to rest, ultimately, on our intuitive understanding of simple sentences such as 'I did this'. And what applies to English applies also to any other language.

In his influential recent book The Rediscovery of the mind John Searle (1994:242) challenges the idea of universal grammar, pointing out that 'the alleged rules of universal grammar' (as conceived by Chomsky and his followers) are not accessible to consciousness, and consequently are not verifiable. I agree with Searle on this point; but this is precisely where the NSM approach to universal grammar differs from the Chomskyan approach. Searle notes that when challenged on this point, Chomsky and Chomskyans tend to 'invoke that most powerful of philosophical arguments: "What else could it be?" "How else could it work?" Deep unconscious rules satisfy our urge for meaning, and besides, what other theory is there? Any theory is better than none at all' (p. 246).

I submit that the NSM approach to universal grammar does provide an alternative theory. This theory does not postulate any 'deep unconscious rules' governing conceptual structures, for it sees conceptual structures as expressible in natural language (any natural language) and thus accessible to consciousness. For example, the rule of universal grammar which allows the construction of the NSM sentence 'something bad happened to me' is in principle accessible to consciousness and intuitively verifiable (for it is no different from the universally attestable pattern itself).

In this context, I would like to quote the following eloquent passage from George Steiner's (1975) famous book After Babel:

There is room, I submit, for an approach whose bias of interest focuses on languages rather than Language; whose evidence will derive from semantics (with all the implicit stress on meaning) rather than from "pure syntax"; and which will begin with words, difficult as these are to define, rather than with imaginary strings or "pro-verbs" of which there can never be any direct presentation. I question whether any context-free system, however "deep" its location, however formal its modus operandi, will contribute much to our understanding of natural speech and hearing. Investigation has shown that even the most formal rules of grammar must take into account these aspects of semantics and performance which Chomsky would exclude.

7. Is a non-ethnocentric perspective on language, culture and cognition possible?

In conjuring up an imaginary opponent of linguistic and cultural universals, Jackendoff (1993:207) writes:

Anyway, what gives you the idea that you can study other cultures systematically? Any description of another culture is unavoidably ethnocentric, an imposition of your own cultural and theoretical prejudices.

And he replies:

But what about language? Any description of another language is unavoidably ethnocentric too, an imposition of our own linguistic and theoretical prejudices. So what? Any description of anything is inevitably tainted by the point of view of the describer.

To me, this reply appears unnecessarily relativist. I agree that any description of anything is inevitably coloured by the point of view of the describer. I do not think, however, that this is a sufficient reason to abandon the attempt at eliminating ethnocentrism. We must acknowledge the dangers of ethnocentrism, but in confronting these dangers we need more than just good will and, as Jackendoff puts it, 'sensitivity and an awareness of our own fallibility' (p. 208). We need a method. The use of a metalanguage based on empirically attested universals provides such a method.

Jackendoff urges that 'in dealing with another language and culture it is important to respect the point of view of those we are observing and analyzing' (p. 208), that is, of 'native speakers who have been trained as linguists' (p. 207). But this admonition lacks conviction, since it is accompanied by an admission that there is no methodology to combat the researcher's inevitable ethnocentric bias. If the alternative to being ethnocentric in a crude and unconscious way is being ethnocentric in a sensitive and aware way, then to my mind this is not much of a choice.

But I believe that we can do better than that: analysts from any cultural background can learn to look at their languages from the point of view of universals, and to discern within the meaning systems of these languages language-specific configurations of universal human concepts.

Ethnocentrism cannot be overcome without an empirical search for linguistic universals. Reportedly Franz Boas often remarked at the beginning of this century that whatever we postulate as universal may be refuted by the very next language we happen to turn our attention to. And this remark is still valid — not in the sense that there are no true universals but in the sense that these universals can only be found on an empirical basis, not postulated a priori.

On this point I must part company with Harré and Gillett (1994:159). The authors quote with approval Catherine Lutz's work on emotion concepts in the Ifaluk language of Micronesia (as I did in Wierzbicka 1992a) and agree with her (as I also did) that English emotion terms such as *anger*, *fear*, and *sadness* represent cognitive artifacts of Anglo culture rather than universal human concepts. They note (as I also did) with reference to Lutz's work, that these English words do not correspond in meaning to the Ifaluk words *song* (roughly, 'reproach/anger'), *metagu* (roughly, 'fear/shame') or *fago* (roughly, 'love/sadness'). Nonetheless, they conclude, we can still understand the emotional lives of the Ifaluk people on the basis of common human ideas such as dignity and honour. They write: 'The common human themes through which the world of Ifaluk is made intelligible to us are such matters as social differentiation, honour, and dignity.'

But the concepts 'dignity' and 'honour' are also highly culture-specific, no less so that 'anger', 'sadness', and 'fear' or the Ifaluk concepts 'song', 'fago', and 'metagu'. In fact, we can no more understand the Ifaluk people via such European concepts as 'dignity' and 'honour' than we can via the unique Japanese concept of 'amae' (or 'sweet dependence', cf. Doi 1974, 1981; Wierzbicka 1991, 1992a and 1997) or the Japanese/Chinese concept of 'on' (cf. Lebra 1976; Wierzbicka 1991, 1992a and 1997).

On the other hand, we can understand both the Ifaluk and the Japanese via universal human concepts such as 'good' and 'bad', 'do' and 'happen', or 'know' and 'want'. For as I have tried to show in the publications (cited above), even the most culture-specific or indeed unique concepts such as 'amae' and 'on', or 'song', 'fago' and 'metagu', can be portrayed and explained as culture-specific configurations of the same universal human concepts. These universal human concepts, however, must be found, and tested, on an empirical basis, by cross-linguistic semantic investigations.

8. Semantic representations in NSM.

I will illustrate the NSM mode of semantic representation (and conceptual analysis) with partial analysis of some English emotion terms. In view of the limitations of space, I will not try to state the full meaning of the words in question, but will only extract for each of them one component: the thought with reference to which a feeling is defined. (For fuller discussion, see e.g. Wierzbicka 1992b and c.) For example, I would say that the English word *remorse* (as in *I felt remorse*) includes the following semantic component: I thought: 'I did something bad'. Proceeding along similar lines, we can state the cognitive (thought-related) components of several other emotion terms as follows:

I was sad	—>	I thought: something bad happened
I was happy	—>	I thought: something good happened
I was sorry	—>	I thought: something bad happened to someone
I was afraid	—>	I thought: something bad can happen to me
I was envious	—>	I thought: something good happened to someone else (it didn't happen to me; this is bad)
I felt Schadenfreude	—>	I thought: something bad happened to someone else (this is good)

In addition, I would say that each emotion term mentioned above includes the component 'I (this person) felt something because of this', plus one or more further components, which I am not going to discuss here.

What this partial analysis shows is that meaning can be stated in the form of simple and intuitively understandable sentences formulated in natural languages (such as 'something bad happened'), without any artificial grammatical machinery and without any technical terms. Furthermore, it shows that meanings can be stated in words apparently available in all languages, such as 'good' and 'bad', 'someone' and 'something', or 'do' and 'happen', and also, that it can be stated in simple sentences which involve only what appear to be universally available combinations of primitives: 'something good happened', 'I did something bad', 'something bad can happen to me', and so on, and which can, therefore, have equivalents in all

languages.

To highlight the non-arbitrary and universal aspects of the metalanguage used in this analysis I would like to compare it again with the kind of metalanguage used in some alternative approaches to semantic analysis, such as, for example, Jackendoff's. Thus, in his 1990 book *Semantic structures* (where his analytical metalanguage was in fact closer to natural language than in his more recent work) the meaning of some English emotion terms are represented as follows (p. 141):

X pleases Y.	$\left[\text{state AFF}^+ ([X], [Y]) \right]$
X displeases Y.	$\left[\text{state AFF}^- ([X], [Y]) \right]$
Y likes X.	$\left[\text{state REACT}^+ ([Y], [X]) \right]$
Y fears/hates X.	$\left[\text{state REACT}^- ([Y], [X]) \right]$

It seems clear that without glosses in natural language, formulae of this kind are incomprehensible and, consequently, unverifiable, and that glosses such as 'AFF — affected entity' (p. 126) and 'REACT — reaction' (p. 132) do not help very much, given that they, too, rely on artificial signs such as pluses and minuses. As an alternative to such artificial and, in my view, unfalsifiable formulae, I would propose semantic explications formulated (approximately) along the following lines:

- X is pleased. =
 person X thinks something like this:
 something good happened
 I wanted this
 because of this, X feels something good
 like people feel when they think something like this
- X is displeased. =
 person X thinks something like this:
 something bad happened
 because someone did something
 I didn't want this, I wanted something else
 because of this, X feels something bad
 like people feel when they think something like this

There is no time here to try to justify these formulae, though I have done this elsewhere (See Wierzbicka 1992b). But the point is that these formulae are intuitively comprehensible and that they are, therefore, testable (and falsifiable). For example, a potential critic could check whether *displeased* is really less than fully symmetrical with *pleased* and whether it refers necessarily to somebody's action (e.g., could one say 'I was displeased with the weather'?). By contrast, formulae which include artificial symbols such as pluses and minuses or brackets, have no intuitively graspable sense and therefore cannot be similarly tested (not to mention the fact that two verbs as different in meaning as *fear* and *hate* are assigned here the same formula.)

Crucially, as noted earlier, evidence suggests that words such as *feel*, *think*, and *want* (as well as *good* and *bad*, *someone* and *something*, *do* and *happen*, *not* and *because*, *I* and *this*) are all lexical universals. For example, in French we

can match them (in the relevant sense) with the words *sentir, penser, vouloir, bon, mauvais, quelqu'un, quelque chose, faire, arriver, non, à cause de, je, ceci*. By contrast, there are of course countless languages which do not have equivalents of English words such as *affected, entity, state* or *reaction*. To treat such English words as 'universal conceptual primitives' means to impose on other languages English categories that are alien to them.

Of course it could be claimed that the words *affected, entity, state* and *reaction* are used here not as English words but as abstract symbols. But to do this would be to deprive the formulae in question of any link with the empirical reality. Equally well, we could say that *pleased* means '+X+Y' and *displeased* means '+X-Y'.

9. Semantic universals and conceptual structures.

I would argue, then, that meanings can be described and compared in a semantic metalanguage based on the shared core of all natural languages, a metalanguage with as many possible versions as there are natural languages, and therefore accessible through all natural languages but essentially independent of any of them. For example, as mentioned earlier, a particular meaning can be represented in the English version of the Natural Semantic Metalanguage (NSM) as 'something bad happened to me', and in the French version, as 'quelque chose de mauvais m'est arrivé', and since the English NSM and the French NSM are mutually isomorphic, nothing hinges on the choice of this or that version: both are intuitively comprehensible and both are free — in their semantics — of culture-specific features.

The question arises: why should all languages, with their tremendous diversity, be reducible to mutually isomorphic mini-languages, with the same mini-lexicon and the same mini-grammar? Why should simple sentences such as 'something bad happened to me' be directly translatable, without any addition or subtraction of meaning, into language after language — given that usually more than 99% of sentences in any text that comes up for translation into many other languages do present problems and do require addition or subtraction of meaning?

My hypothesis is that all these mini-languages, which provide lexical and grammatical primitives for semantic analysis of utterances and texts within any language, are isomorphic because they are all language-specific variants of the same innate and universal 'lingua mentalis', the language of the human mind (cf. Wierzbicka 1980). In my view, true insights into the working of the human mind will come neither from pure speculation about language and thought in general nor from the invention of new formalisms, but from the empirical study of languages — languages conceived of not as autonomous formal systems, but immensely complex, culturally-shaped and constantly changing tools for creating and expressing meaning.

REFERENCES

- Allan, Keith. 1986. *Linguistic Meaning*. London: Routledge, Kegan Paul.
 Apresjan, Jurij D. 1974. *Leksičeskaja semantika*. Moscow: Nauka.
 Apresjan, Jurij D. 1992. *Lexical semantics: User's guide to contemporary Russian vocabulary*. Ann Arbor: Karoma.
 Aristotle. 1937. *Categories*. In W.D. Ross (ed.) *The works of Aristotle*. Vol. 1. Oxford: Clarendon Press.

- Becker, Ernest. 1962. *The birth and death of meaning: A perspective in psychiatry and anthropology*. New York: Free Press of Glencoe.
- Bettelheim, Bruno. 1983. *Freud and man's soul*. London: Hogarth Press.
- Descartes, René. 1931[1701]. *The search after truth by the light of nature*. In *The philosophical works of Descartes*. Vol 1: 305-327. York: Free Press of Glencoe.
- Doi, Takeo. 1974. *Amoe*: A key concept for understanding Japanese personality structure. In: Takie S. Lebra and William P. Lebra (eds.). *Japanese culture and behavior*. Honolulu: University of Hawaii Press. 145-154.
- Doi, Takeo. 1981. *The anatomy of dependence*. Tokyo: Kodansha.
- French, Lucia and Katherine Nelson. 1985. *Young children's knowledge of relational terms: Some Ifs, Or and Buts*. New York: Springer Verlag.
- Freud, Sigmund. 1963[1916-17]. *Introductory lectures on psycho-analysis (part III)*. London: Hogarth Press.
- Goddard, Cliff. 1989. *Issues in Natural Semantic Metalanguage*. *Quaderni di Semantica*, 10(1): 51-64.
- Goddard, Cliff. 1994. *Semantic theory and semantic universals*. In Goddard and Wierzbicka (eds.), 7-29.
- Goddard, Cliff. 1998. *Semantic Analysis*. Oxford: Oxford University Press.
- Goddard, Cliff and Anna Wierzbicka (eds.). 1994. *Semantic and lexical universals: Theory and empirical findings*. Amsterdam: John Benjamins.
- Goddard, Cliff and Anna Wierzbicka, (eds). *Forthcoming. Meaning and universal grammar*.
- Harré, Rom and Grant Gillett. 1994. *The discursive mind*. London: Sage.
- Humboldt, Carl Wilhelm von. (1903-36). *Wilhelm von Humboldts Werke*. 17 volumes. Ed. Albert Leitzmann. Berlin: B. Behr.
- Jackendoff, Ray. 1990. *Semantic structures*. Cambridge, MA.: MIT Press.
- Jackendoff, Ray. 1991 *Parts and boundaries*. In Beth Levin and Steven Pinker (eds), *Lexical and conceptual semantics*. Cambridge MA: Blackwell. 9-47.
- Jackendoff, Ray. 1993. *Patterns in the mind: Language and human nature*. New York: Harvester Wheatsheaf.
- Lebra, Takie. S. 1976. *Japanese patterns of behavior*. Honolulu: The University Press of Hawaii.
- Leibniz, Gottfried Wilhelm. 1903[1704]. *Table de définitions*. In Louis Couturat (ed.). *Opusculs et fragments inédits de Leibniz*. Paris: Presses Universitaires de France. 437-510. (Reprinted 1961, Hildesheim: Georg Olms.)
- Leibniz, Gottfried Wilhelm. 1981[1703]. *New essays on human understanding*. Peter Remnant and Jonathan Bennett (trans.). Cambridge: Cambridge University Press.
- Lutz, Catherine. 1990. *Unnatural emotions*. Chicago: University of Chicago Press.
- Martin, Gottfried. 1964. *Leibniz; Logic and metaphysics*. K. J. Northcott and P. G. Lucas (trans.). Manchester: Manchester University Press.
- McCawley, James D. 1973. *Grammar and meaning: Papers on syntactic and semantic topics*. New York: Academic Press.
- Peeters, Bert. 1994. *Semantic and lexical universals in French*. In Goddard and Wierzbicka (eds.). 423-442.
- Sapir, Edward. 1949[1929]. *Status of Linguistics as a science*. In *Selected writings of Edward Sapir in language, culture and personality*. David Mandelbaum (ed.) Berkeley: University of California Press. 160-166.
- Searle, John. R. 1994. *The rediscovery of the mind*. Cambridge MA: MIT Press.

- Shweder, Richard A. 1990. Cultural psychology: What is it? In J.W.Stigler, R.A.Shweder, and G.Herd (eds.). *Cultural psychology: Essays on comparative human development*. Cambridge: Cambridge University Press. 1-46.
- Steiner, George. 1975. *After Babel*. London: Oxford University Press.
- Weisgerber, Leo. 1939. *Die volkhaften Kräfte der Muttersprache*. Frankfurt a.M.
- Wierzbicka, Anna. 1980. *Lingua mentalis: The semantics of natural language*. Sydney: Academic.
- Wierzbicka, Anna. 1991. Japanese key words and core cultural values. *Language in Society* 20: 333-85.
- Wierzbicka, Anna. 1992a. *Semantics, culture and cognition: Universal human concepts in culture-specific configurations*. New York: Oxford University Press.
- Wierzbicka, Anna. 1992b. Defining emotion concepts. *Cognitive Science*. 16: 539-81.
- Wierzbicka, Anna. 1992c. Talking about emotions: Semantics, culture and cognition. *Cognition and Emotion* (Special issue on basic emotions). 6(3-4): 289-319.
- Wierzbicka, Anna. 1996. *Semantics, Primes and Universals*. Oxford: Oxford University Press.
- Wierzbicka, Anna. 1997. *Understanding cultures through their key words*. New York: Oxford University Press.
- Wierzbicka, Anna. 1998. Angst. *Cultural Psychology*. 4(1):11-39.

**PARASESSION:
PHONETICS AND
PHONOLOGICAL UNIVERSALS**

Testing Opposing Phonetic Structural Principles: Polarization and Gestural Economy

Victoria. B. Anderson
University of California, Los Angeles

1. Introduction. This study examines an aspect of the relationship between phonetics and phonology: the influence of phonological patterns on phonetic implementation. Specifically, the matter under focus here is whether the number of phonological contrasts along a given phonetic continuum affects the articulation of those contrasts. The larger goal is to contribute to our understanding of the ways in which physical and abstract representations of language sounds are connected. The approach taken to consider this question is the use of instrumental articulatory evidence from a language which employs varying numbers of contrasts in different phonotactic positions, and so is suited for such an analysis. We take as specific starting points for the investigation two desirable but opposing structural principles of articulation: *polarization* and *gestural economy*.

1.1. Polarization. The original purpose of the polarization principle (Keating 1984) was to explain the differing behavior of segments belonging to the same phonetic categories, in different languages. In treating the implementation of Voice Onset Time (VOT) across languages, Keating noticed that firstly, three and only three VOT categories seemed to exist across languages which had been described: {voiced}, {voiceless unaspirated}, and {voiceless aspirated}. Secondly, adjacent categories were always employed by a given language. Thus, Keating suggested that along a given phonetic continuum, phonological oppositions are instantiated by a fixed number of phonetic categories. Moreover, phonological oppositions can be mapped to different phonetic categories in different languages. Keating observed that English and Polish word-initial stops both employ two of the three contrastive types along the VOT continuum, and both make use of the {voiceless unaspirated} category. However, in Polish the opposing stop is in the {voiced} category, whereas in English the opposing stop is in the {voiceless unaspirated} category, as shown schematically for labials in Figure 1.

Voice Onset Time continuum----->

Phonetic Category	{voiced}	{voiceless unaspirated}	{voiceless aspirated}
Polish	b	--->p	
English		b<---	p ^h

Figure 1: Schematized VOT in Polish and English word-initial labial stops

Crucially, polarization states that once phonetic categories have been chosen to instantiate the phonological oppositions in a language, actual phonetic values will be widely separated within those categories, so as to maximize the difference between the contrastive segments. (In this respect, polarization is a subset of *dispersion theory*, Lindblom 1986, 1990; Flemming 1997, to appear.) Thus, the stops in the {voiceless unaspirated} category in English and Polish are not identical, but are displaced with respect to each other, as shown in Figure 1. This has the effect of maximizing the difference with the other member of the contrastive pair; {voiced} in the case of Polish, and {voiceless aspirated} in the case of English.

Note that polarization implicitly places a higher value on perceptual ease in distinguishing among contrasts than on articulatory ease in producing them, in the sense that articulations will be polarized even if this means more extreme

articulations, or means storing more than one motor implementation program for a contrast.

1.2. Gestural Economy. Another possible view of the relationship between phonetics and phonology involves a theory of articulation called *gestural economy* (Maddieson 1996.) The ideas of gestural economy emerge from concepts entertained in Lindblom and Maddieson, 1988. Gestural economy implicitly acknowledges that there must be a balance between articulatory and perceptual ease. Maddieson notes that it is necessary and desirable to maintain sufficient perceptual distinctiveness among contrastive segments. However, given sufficient distinctiveness, the desiderata of gestural economy are two-fold. First, languages place high value on the use of efficient gestures; those gestures which show “economy of kind”. Secondly, languages place high value on re-use of these efficient gestures, so that speakers need not internalize more motor programs than necessary to convey the phonological oppositions of their language. This principle constitutes “economy of number.”

Maddieson presents instrumental evidence for *economy of kind* from Sele and Ewe, two languages of West Africa. Sele employs a single labiodental fricative which does not contrast with another labial fricative. Ewe, on the other hand, places a labiodental fricative in contrast with a bilabial fricative. On the continuum of upper lip height, Maddieson establishes instrumentally that in the Sele labiodental, the upper lip is in a neutral position, neither lowered nor raised. Elsewhere, Maddieson (1984) gives evidence that labiodental fricatives are frequent members of consonant inventories, and proposes that this is because they require “precise positioning of only one active articulator” and “a relatively small movement” while remaining acoustically distinct from other fricatives (Maddieson 1996).

In Ewe, the bilabial fricative shows upper lip lowering as both lips approach each other for the constriction. However, even though it contrasts with the bilabial, the Ewe labiodental fricative is not appreciably different from the Sele labiodental. (The distribution of these segments is schematized in Figure 2.) Maddieson argues that this is because /f/ is qualitatively economical. In Ewe a more efficient segment is used over a potentially more distinctive segment; gestural economy outweighs maximal distinctiveness. Recall that under polarization, one would expect that the Sele and Ewe labiodentals would differ; the Ewe labiodental would involve a higher lip position, in order to maximize its difference from the bilabial.

Upper lip height----->

Phonetic Category	{low}	{neutral}	{raised}
Sele		f	
Ewe	φ	f-->f̥	

Figure 2: Schematized upper lip height positions for Sele and Ewe labial fricatives. (Outlined symbol shows position under polarization.)

Maddieson cites the re-use of efficient places of articulation across different manners to illustrate *economy of number*. The fact that languages very commonly have stops, nasals and laterals reoccurring at the same places of articulation is, in Maddieson’s view, an indication that efficient component gestures are being re-used to minimize the number of articulatory motor programs needed for phonological contrasts. However, here Maddieson’s evidence comes from the UCLA Phonological Segment Inventory Database (Maddieson 1984) in which segments are categorized *phonologically*; phonetic detail is purposely underplayed. Thus, it remains an empirical question for the majority of languages whether stops and

nasals for instance, are actually articulated identically, or whether they are simply categorized together phonologically.

	p	t	k
√	m	n	ŋ
*	ɱ	ɳ	ɴ

Figure 3: Economy of number vs use of different gestures for stops and nasals

2. Predictions. The issue under focus in this study is that polarization and gestural economy make different predictions about the impact of phonological structure on phonetic implementation. Specifically, in a single language in which different numbers of contrasts operate in different environments, polarization predicts that phonetic realization will be affected by the different numbers of contrasts, while gestural economy predicts that segments will not polarize when contrasts are added.

The language examined here was Western Arrernte, an Arandic language of Central Australia, which makes use of four contrastive coronal places of articulation in the stops, nasals, laterals and prestopped nasals. Coronals are those sounds made by an articulation of the tongue tip and/or blade with the upper surface of the vocal tract. Common auditorily-based descriptions of tongue configurations for these four places of articulation are shown in Figure 4, and involve two tip articulations (apicals) contrasting in place, as well as two blade articulations, (laminals) contrasting in place. Intervocally, the four-way contrast holds. However, in word-initial position, there is no place contrast between the apicals.

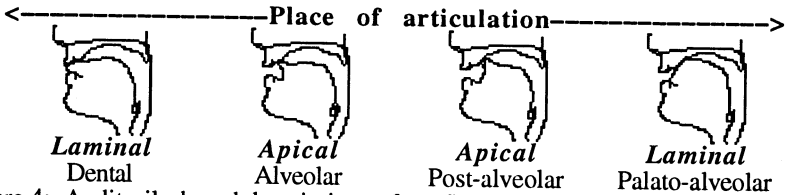


Figure 4: Auditorily-based descriptions of configurations for Australian coronals

Fig. 5 shows examples for stops. "T" is used to refer to the non-contrastive apical.

	Bilabial	Coronal places of articulation				Dorsal Velar
		Laminar Dental	Apical Alveolar	Apical Post- alveolar	Laminar Palato- alveolar	
Stops	p	t	t	t	t	k
Word-initially	pəṯə <i>pouch (n)</i>	ṯəmə <i>grind (vt)</i>	Təpə <i>back (n)</i>		ṯənə <i>friend (n)</i>	kəpə <i>firestick (n)</i>
Between n vowels	mapə <i>many (n)</i>	aṯə <i>I (pr, tr.)</i>	latə <i>today (n)</i>	kwaṯə <i>egg (n)</i>	kwaṯə <i>water (n)</i>	makə <i>elbow (n)</i>

Figure 5: W. Arrernte stops, word-initially and between vowels

2.1. Predictions of Polarization. The different predictions that the two principles make are as follows. Under a polarization scenario, the two contrastive apicals will be widely distributed on the place continuum with respect to the noncontrastive apical, in order to maximize distinctions between them. The advantage here is that listeners economize on perceptual effort in recognizing the contrastive segments.

The disadvantage, however, is that speakers must control three different motor programs for apicals, which violates economy of number.

Place dimension ----->

Non-contrastive	T
Contrastive	t <----- • -----> ‡

Figure 6: Arrernte Polarization Scenario--contrastive vs non-contrastive apicals

A second scenario involving polarization compares stops at each place of articulation with their corresponding nasals. On the assumption that nasals are less perceptually robust than stops because they do not have as many acoustic cues, it may be that nasals are articulated more divergently on the place continuum than stops, in order to maximize perceptual distances among them. Anecdotal evidence for such a possibility has been observed in Malayalam, a Dravidian language of India which, like Arrernte, employs four contrastive oppositions for tongue-palate articulations in both stops and nasals. Ladefoged and Maddieson (1996) observe that for a few speakers of Malayalam, dental nasals are articulated as interdental, while dentals stops are articulated as post-dentals. For these speakers, dental nasals are articulated further forward on the place continuum than dental stops. In Figure 7, positions for the other stops and nasals have been hypothesized, although only the dentals are explicitly mentioned by Ladefoged and Maddieson.

Place of articulation continuum----->

Phonetic Category	{dental}	{alveolar}	{postalveolar}	{palatal}
Stops	ʈ	ʈ̚	ʈ̠̚	ʈ̠̞̚
Nasals	ɳ	ɳ̠	ɳ̠̠	ɳ̠̠̞

Figure 7: Schematized places of articulation in Malayalam

2.2. Predictions of Gestural Economy. Gestural economy of number dictates that an articulation will be re-used in both contrastive and non-contrastive environments. That is, an articulation representing the non-contrastive apical will reappear as one of the contrastive apicals, as shown in Figure 8.

Place dimension ----->

Non-contrastive	T
Contrastive	t ----- ‡

or

Non-contrastive	T
Contrastive	t ‡

Figure 8: Arrernte Gestural Economy Scenario--contrastive vs non-contrastive apicals

The advantage here is that speakers economize on motor programming; only two articulations exist for apicals. However, the disadvantage is that the contrastive articulations are not so widely divergent on the place continuum, which violates maximal distinctiveness.

Gestural economy also dictates that nasals will differ from stops in the position of the velum, but will be identical in terms of the gestures used to achieve the oral configuration.

Place of articulation continuum----->

Phonetic Category	{dental}	{alveolar}	{postalveolar}	{palatal}
Stops	t̪	t	t̠	tʃ
Nasals	ɲ̪	ɲ	ɲ̠	ɲ̟

Figure 9: Schematized coronal places of articulation in Arrernte: gestural economy

3. Method. In order to empirically investigate how W. Arrernte resolves these conflicting structural principles, direct instrumental articulatory data for tongue-palate contact patterns was collected using a method called static palatography, described in detail in Ladefoged (1997). Briefly, a speaker's tongue is painted with a non-toxic marking material, in this case a mixture of olive oil and digestive charcoal. The speaker utters a word containing the segment of interest, and inserts a mirror into the mouth to reflect the contact area on the palate. The contact area is videotaped (see Figure 10) and the speaker rinses his or her mouth with water.

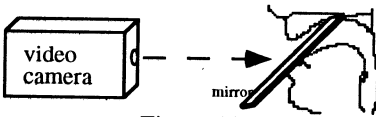


Figure 10

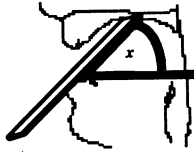
3.1. Data Set. To the extent possible, it is important to choose words containing only low vowels and labials in addition to the coronal of interest, in order to avoid confounding effects of other tongue-palate contacts. The data set used for this experiment is shown in Figure 11.¹

	Laminal Dental	Apical Alveolar	Apical Post-alveolar	Laminal Palato-alveolar
<i>Stops</i> Between Vowels	'pəʃə <i>pouch (n)</i>	'matə(ɹə) <i>cloud (n)</i>	'pəʃə <i>rock (n)</i>	'petəmə <i>is coming (vi)</i>
<i>Stops</i> Word-initially	't̪əmə <i>grind (vt)</i>	'Təpə <i>back (n)</i>		'tapə <i>grub (n)</i>
<i>Nasals</i> Between Vowels	ip'məŋə <i>grandmother (n)</i>	'manə <i>money (n)</i>	'məŋə <i>veg. food (n)</i>	'mpənə <i>crumb (n)</i>
<i>Nasals</i> Word-initially	'ŋəmə <i>rain is falling (vt)</i>	'Nəmə <i>is sitting (vi)</i>		't̪nəmə <i>is falling (vi)</i>

Figure 11: Data set for W. Arrernte static palatography

3.2. Image Processing.

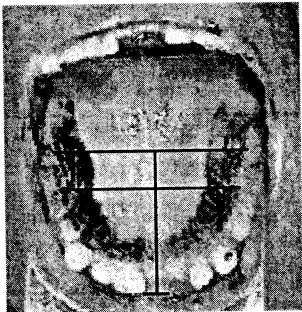
For each token, the relevant video clip was digitized into a computer file from which a high-quality still frame was chosen for analysis. The still frame was corrected for angular distortion caused by slight differences in the angle at which the mirror was placed in the mouth relative to the occlusal plane (angle x , left.) For each still frame, vertical and horizontal planes were corrected independently, based on actual millimeter measurements from lifesize palatal plaster casts.



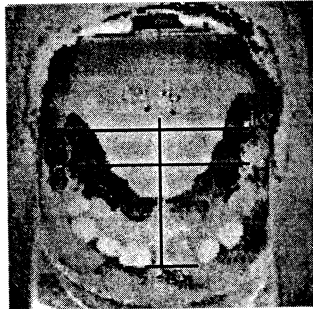
To obtain the horizontal correction factor, a horizontal calibration measure was taken at a line just forward of the left and right first molars (line "h", Figure 13.) This measure was taken on the lifesize plaster cast as well as on the still frame in question, and the ratio of the two was used as the correction factor.

To obtain the vertical correction factor, a vertical calibration measure was taken at a line drawn from the front edge of the front incisors to the horizontal calibration line (line "v", Figure 13.) Again, this measure was made on both the lifesize plaster cast and on each still frame, to produce a correction factor for each vertical measurement on a given still frame.

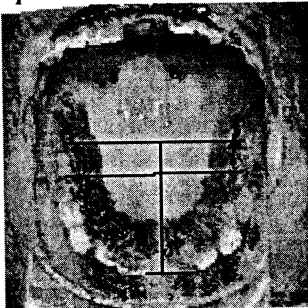
4. Analysis. Figure 12 shows, for one speaker, a sample still frame for each of the four categories of contrastive stop. In each case, upper teeth are shown at the top of the photograph, and are also reflected at the bottom.



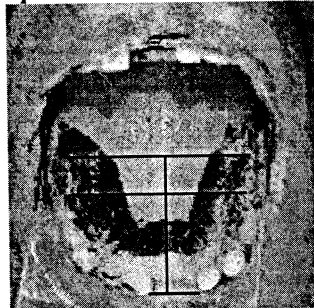
Apical Alveolar



Apical Post-alveolar



Laminal Dental

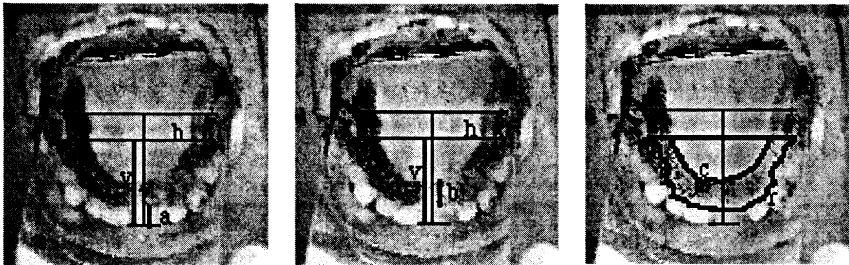


Laminal Palato-alveolar

Figure 12: Tokens of the medial stops for one speaker

After examining still frames for several speakers, measurements were chosen that would yield fruitful ways of differentiating among the five place categories.

4.1. Frontmost Contact. Comparative inspection of frames showed that *frontmost contact* on the palate in the midline was likely to yield differences among categories. Indeed, this metric has been used before by Butcher (to appear), Dart (1991) and others, as a index of place of articulation.² Note that in Figure 12 the laminal dental shows contact on the surfaces of the incisors. The apical alveolar begins further back, clear of the teeth in the case of this token, and frontmost contact for the apical postalveolar and laminal palatoalveolar begin appreciably further back than the apical alveolar. Frontmost contact was quantified by measuring from a line drawn at the base of the teeth back to the front edge of the contact pattern, in the midline (line a, Figure 13).



Frontmost Contact

Length of Contact

Palatality Index

Figure 13: Measurements shown on non-contrastive apical token

4.2. Length of Contact. The apical/laminal distinction itself leads us to expect differences between apicals and laminals based on length of contact from front to back in the midline. However, there may be significant differences among each of the place categories, or indeed between stops and nasals in a single category, in addition to the apical/laminal distinction. Observe that the laminal dental and laminal palatoalveolar in Figure 12 are very broad in their contact area from front to back, although it is not possible to tell from these tokens which may be broader or whether there is a significant difference between them. The apicals, as expected, are much narrower in midline contact. Like frontmost contact, this metric has been used before (Butcher, to appear), and has expected acoustic correlates.³ Length of contact was measured by drawing a line from the front to rear of the contact in the midline (line b, Figure 13.)

4.3. Palatality Index. Having considered the space in front of the constriction and the length of the constriction itself, we considered differences in the amount of empty space that remained behind the constriction, which is affected by raising of the sides of the tongue body and how far back the midline constriction extends. The more contact on the palate, the smaller this space will be. Here the clearest difference is between the apical alveolar and the laminal palatal, both because of the rear point of midline contact, and the extent of lateral contact on the sides of the palate. Since this space reflects cavity size behind the constriction, we again expect acoustic concomitants.⁴ To quantify the palatality index, an area bounded by the horizontal calibration line and the rear contact line was measured. (This area is shown as the inner outlined area "c" in Figure 13.) Analogous to the linear measurements described above, these area measurements were corrected by reference to a calibration area. In each case the calibration area was defined by the

horizontal calibration line and the juncture of the teeth with the gumline. (This area is the outer reference outlined area “r” in Figure 13.)

Incidentally, Figure 13 a, b, and c show a token of the non-contrastive apical for the same speaker, for purposes of comparison with Figure 12. Recall that we wish to categorize non-contrastive stops and nasals in terms of the three measurements described above, to determine how or if they differ from the contrastive apical categories. If the non-contrastive apical has characteristics of an alveolar or a postalveolar, gestural economy is indicated. If it has intermediate characteristics, this would indicate polarization of the contrastive apicals with respect to the non-contrastive apical.

5. Results

5.1. Stops vs Nasals. Let us first consider results of the three metrics for stops and nasals at each place category. Statistical analysis of variance (ANOVA) showed that stops and nasals at any given place of articulation were indistinguishable from each other along all of the measures taken here.

Recall the possible Malayalam scenario outlined in section 2.1. Investigators suggested that in order to make weaker nasals more perceptually robust, speakers might articulate them more divergently on the place continuum. However, at least for W. Arrernte, this is found not to be the case. First, perception tests of the stops and nasals with native listeners show that the nasals are as perceptually robust as the stops (Anderson 1997). Second, the manner distinction does not interact with place distinctions in any of our measures of tongue-palate contact. Thus we have the scenario shown in Figure 9. Gestural economy of number is suggested by these results. Though the position of the velum differs for stops and nasals, the same oral gesture is used in both. In the following discussion of results for contrastive and non-contrastive apicals, we will collapse stops and nasals together.

5.2. Contrastive vs Non-contrastive Apicals. Before comparing initial apicals with medial ones, we must first determine whether such a comparison is appropriate. For example, it may be that the segments simply involve positional differences, whether or not they involve different numbers of contrasts. (See Keating et.al., to appear, regarding fortition of initial segments.) Luckily, the laminals provide a way to address this question, because they involve the same numbers of contrasts in initial and medial positions. Results of ANOVA on laminal initials and medials show that initials are not strengthened or weakened vis-a-vis medial segments. Therefore, we will take the comparison between non-contrastive and contrastive apicals to be an appropriate one, and assume that results are not an artifact of position. Moreover, results for initial and medial laminals will be shown together. Let us now consider each measure in turn, comparing non-contrastive and contrastive apicals.

Results for the five place categories were compared, both as raw measurements in millimeters and as proportional measurements reflected as a percentage of the appropriate calibration measurement. Proportional measurements take into account the different mouth sizes of speakers. Quite surprisingly, the use of either raw or proportional measurements yielded the same significance results. Proportional measurements showed slightly greater significance values, and will be reported here. In each case, the summarized results below reflect data from 171 tokens ([7 stops + 7 nasals] x 6 speakers x approx. 2 tokens each.) Statistical significance depended on a p value of $\leq .05$.

5.2.1. Frontmost Contact.

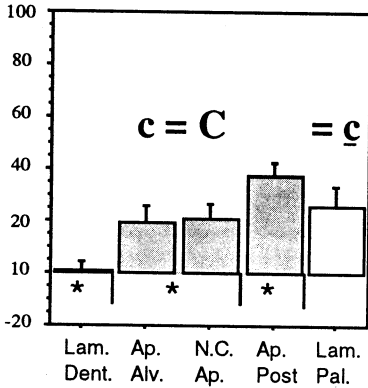


Figure 14

categories are not statistically differentiable; i.e. they behave as a group. Both are statistically differentiable from the postalveolar, which begins significantly further back from the alveolar and non-contrastive segments.⁵

5.2.2. Contact Length

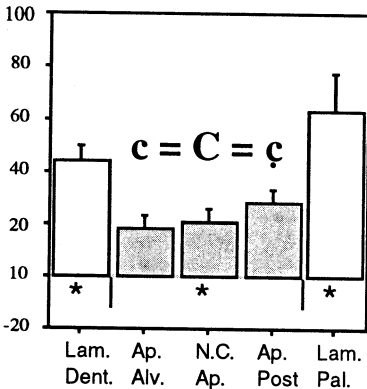


Figure 15

polarization or gestural economy.

Figure 14 shows results for frontmost contact on the palate in the midline, summed over the six speakers. The x-axis shows the five place categories, with non-contrastive apicals placed between the contrastive ones for ease of comparison. Apical categories are shown with shaded bars. Laminals are included for completeness, and are shown unshaded. Measurements on the y-axis are reported as percent of the vertical calibration line. Lines above bars reflect one standard deviation.

As expected, we find that laminal dentals have the lowest percent values, and thus the furthest frontmost contact of the five segment categories. The non-contrastive and alveolar apical categories group together, which is as expected. As expected, we find that laminal dentals have the lowest percent values, and thus the furthest frontmost contact of the five segment categories. The non-contrastive and alveolar apical categories group together, which is as expected. Results are shown in Figure 15. Again, place categories are shown on the abscissa; contact length as a percentage of the vertical calibration line is shown on the ordinate. As expected, we find laminals to be much broader in contact than apicals; these segments have been correctly described in auditory descriptions. However, the laminals do not act as a group. Laminal palatoalveolars are statistically significantly broader than laminal dentals. On the other hand, none of the apicals differ statistically from each other, as we hypothesized would be the case at first glance. Thus, since even the two contrastive apical categories group together, this measure provides no insight into either

5.2.3. Palatality Index.

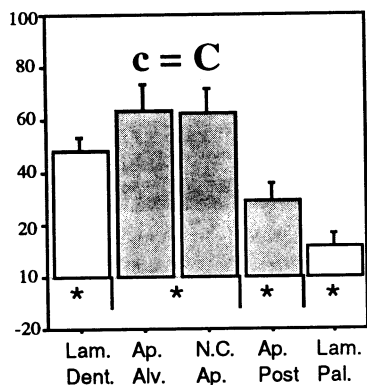


Figure 16

In Figure 16, the x-axis is as above; the y-axis shows percent of *uncontacted* space as a percentage of the calibration area described in 4.3 above. Since a large amount of contact with the palate means a small cavity size, we expect an inverse relationship between the percentage value reported and amount of contact on the palate.

Laminal palatoalveolars have far and away the most contact with the palate, and thus the smallest uncontacted area. This category is quite significantly differentiable from all others, and reflects both the laminality of the midline contact and lateral tongue raising behind the midline constriction.

Alveolar and non-contrastive apicals do not differ at all on this measure. They behave as a group, but are significantly differentiable from every other place category.

6. Summary and Conclusion. In this study we have examined evidence for the phonetic principles of polarization and gestural economy, by quantitatively characterizing the articulation of segments in which we might expect phonological structure to affect phonetic realization of segments. We focused on the coronal stops and nasals in Western Arrernte, which employ different numbers of contrasts initially and medially. Results for three quantitative articulatory measures were shown. Results for *contact length* were compatible with either polarization or gestural economy, and so did not bear on the issue of how these opposing principles may compromise or be weighted. Results for *frontmost contact* and *palatality index* indicate that in terms of both the front and back cavities, alveolar and non-contrastive apicals are identical, which is evidence for re-use of gestures and bears out the predictions of gestural economy of number. Moreover, the fact that the alveolar rather than the postalveolar was used in the greater number of contexts points to gestural economy of kind: use efficient (in this case *less displaced*) gestures.

We also showed that stops and nasals of a given place category do not differ from each other along any of these measures, which indicates economy of number. This result, added to the fact from perceptual work that Arrernte nasals can be differentiated just as well as the stops also gives credence to the idea that "sufficient perceptual contrast" rather than contrast-maximizing principles may be at work.

Polarization and gestural economy are both desirable structural principles of articulatory phonetic implementation. However, in this study we have accrued evidence only for the latter, in both "the number and the nature of the gestures used" (Maddieson 1996).

¹Two cases require discussion. In the word for cloud, /^hmatəɹə/, speakers were instructed to omit the final syllable, in order to avoid potential confounding effects of the /ɹ/ on the resulting pattern. Speakers participating in the study were consulted to make sure they were satisfied that their articulation and the resulting portion of the word sounded reasonably accurate. Also, since a word completely lacking other coronals or high vowels could not be found for the laminal palatoalveolar word-initial nasal /^hn/, a prestopped nasal /^htn/ was used.

²Moreover, since this measure reflects the size of the cavity in front of the constriction, it can be taken to be associated with acoustic correlates such as spectral shape.

³Since this measure is associated with the mass of the articulator, we would expect extent of midline contact to relate directly to relative voice onset time, amount of friction at the burst release, and duration of formant transitions.

⁴For example, transition loci at the onset of the following vowel.

⁵As an aside, the frontmost edge of the apical postalveolar is further back than that of the laminal palatoalveolar. And though the palatoalveolar is statistically distinct from the laminal dental, it is in the realm of the alveolar/non-contrastive apical group, which is unexpected. However, it is harder to maintain that frontmost contact is the optimal measure of intended "place" or "target" in the case of such a broad contact.

References

- Anderson, V. 1997. The perception of coronals in Western Arrernte. *Proceedings of the Fifth European Conference on Speech Communication and Technology*, vol. 1, 389-392.
- Butcher, A. To appear. *The phonetics of Australian languages*. Oxford: Oxford University Press.
- Dart, S. 1991. Articulatory and acoustic properties of apical and laminal articulations. *UCLA Working Papers in Phonetics* 79.
- Flemming, E. 1997. Phonetic optimization: Compromise in speech production. *University of Maryland Working Papers in Linguistics* 5.
- Flemming, E. To Appear. *Phonetic detail in phonology: Towards a unified account of assimilation and coarticulation*. Southwest workshop in optimality theory: Features in OT. SWOT 1. Coyote papers.
- Keating, P. 1984. Phonetic and phonological representations of consonant voicing. *Language* 60. 286-319.
- Keating, P., T. Cho, C. Fougerson, C. Hsu.. To appear. Domain-initial articulatory strengthening in four languages. *Labphon* 6.
- Lindblom, B. 1986. Phonetic universals in vowel systems. *Experimental Phonology*, ed. by J. Ohala and J. Jaeger. Orlando: Academic Press.
- Lindblom, B. 1990. *Phonetic content in phonology*. PERILUS 11.
- Lindblom, B. and I. Maddieson. 1988. *Phonetic universals in consonant systems. Language, speech and mind: Studies in honor of Victoria A. Fromkin*, ed. by L. Hyman and C. Li. London: Routledge.
- Ladefoged, P. 1997. Instrumental techniques for linguistic phonetic fieldwork. *Handbook of phonetic sciences*, ed. by W. Hardcastle and J. Laver, 137-166. Blackwells: Oxford.
- Ladefoged, P. and I. Maddieson. 1996. *The sounds of the world's languages*. Cambridge, MA: Blackwells.
- Maddieson, I. 1984. *Patterns of sounds*. New York: Cambridge University Press.
- Maddieson, I. 1996. *Gestural economy*. *UCLA Working Papers in Phonetics* 94.

Perceptual Confusions and Phonological Change: How Confused is the Listener?

Patrice Speeter Beddor
University of Michigan

and

Rena Arens Krakow
Temple University

1. Introduction

It has long been known that phonological systems are shaped, in part, by auditory factors. A crucial aspect of this shaping is the need to preserve perceptual distinctiveness among the contrastive speech sounds of a language. However, as is evident from daily communicative interactions, flawless productions of these speech sounds do not always occur and the perceptual mechanisms that enable listeners to maintain the distinctions are far from perfect. Hence, perceptual confusions arise more than occasionally. Because the communicative process is a dynamic one, with variations in the speech of a linguistic community having the potential to become a systematic property of that community's linguistic system, such confusions may lead to sound changes and consequently modifications of the set of phonological contrasts.

Acoustic and perceptual experimental evidence offers insight into the conditions leading to, and the consequences of, failure to maintain sufficient auditory distinctiveness in sound systems. Nasalization is a particularly interesting property to study in this regard. Perceptual evidence has shown, for example, that place distinctions in consonants are less salient when the consonants are nasal, than when they are oral (Mohr and Wang 1968, Hura et al. 1992). These distinctions are particularly non-salient in syllable-final position (Repp and Svastikula 1988, Ohala 1990, Beddor and Evans-Romaine 1995), which is compatible with articulatory evidence showing that nasal consonants have weaker oral constrictions in final than in initial position (Krakow 1989) and with acoustic evidence of a less abrupt change in spectral structure in VN than in NV syllables (Seitz et al. 1990, Manuel 1991). It is not only consonants that become more confusable when nasalized. For vowels, height differences are less distinct for nasal vowels than for their oral counterparts (Mohr and Wang 1968, Wright 1986) and perceived height may shift as a consequence of vowel nasalization (Krakow et al. 1988, Beddor and Hawkins 1990, Kingston 1991). The perceptual height effects are in keeping with acoustic evidence that nasal coupling, like tongue/jaw height, has its primary spectral effects on the first formant region of the vowel spectrum (Hawkins and Stevens 1985, Maeda 1993). Consequently, it is possible for the spectral information in that region to be ambiguous or misleading to the listener.

Clearly related to these experimental findings is an array of phonological facts. Perhaps most evident is the cross-linguistic frequency with which nasal (N) consonants in VNC clusters lose their distinct place constriction and take on that of the following oral (C) consonant. More relevant to the focus of the present paper is that, in most languages with distinctive vowel nasalization, these vowels have their origin in VN sequences in which the final consonant constriction disappeared while the coarticulatory effect of that nasal remained on the preceding vowel. That relatively non-salient place information in final nasal consonants contributes to consonant loss is supported by historical evidence that, prior to loss, place distinctions in final nasal consonants sometimes merge (Chen 1973, Tuttle 1991; cf. Hajek 1997). In addition, changes in vowel height due to vowel nasalization are not uncommon in the world's languages (Bhat 1975, Ruhlen 1978, Beddor 1982).

All of these phonological phenomena can be viewed as instances in which listeners fail to attribute a spectral property to the correct coarticulatory source.

2. A Complex Picture

But what might lead listeners to fail to attribute a spectral property to its coarticulatory source? Any attempt to link experimental evidence of relatively non-salient, and hence presumably unstable, perceptual distinctions with phonological change has to face the question of how an earlier stage of that sound system could tolerate the distinctions to begin with. In a now-classic paper, Ohala (1981) proposed that listeners are normally able to disambiguate at least certain types of perceptual similarities between speech sounds by applying 'reconstructive rules' that are based on listeners' language-specific knowledge of coarticulatory organization; sound changes may occur when listeners either fail to apply, or overapply, these rules. Under this view, normal perceptual processing would be particularly likely to break down if listeners failed to detect the coarticulatory source — that is, the conditioning environment — for the spectral property (Ohala 1981, 1986, 1993).

Conversely, listener confusions should be minimized when the acoustic signal provides clearly audible contextual information. Listeners who detect the source of the spectral property in question should be able to use their knowledge of coarticulatory structure to adjust for the variation and correctly identify the intended utterance. Perceptual studies that manipulate the consonantal context for vowel nasalization provide evidence that supports this approach. For example, Kawasaki (1986) found that attenuation of the nasal consonants in natural NVN utterances enhanced perceived vowel nasality. The American English listeners in that study were more likely to perceive a nasal vowel as oral when it was in a clearly audible nasal consonant context. This is consistent with the view that, when listeners detected the nasal consonant, they attributed the effects of nasalization on the vowel spectrum to the nasal consonant rather than to the vowel itself. Focusing on the height rather than on the nasality of nasal vowels, Krakow et al. (1988) found that American English listeners made accurate vowel height judgments of synthetic nasal vowels in an oral consonant context ([b_d]), but made vowel height mistakes on such vowels in a nasal consonant context ([b_nd]) (cf. Kingston and Macmillan 1995). One interpretation of this outcome is that listeners used the coarticulatory nasal context to attribute correctly the relevant aspects of the vowel's low-frequency spectral properties to nasality rather than tongue/jaw height.

More generally, experimental evidence (involving a variety of speech sounds and contexts), indicating that listeners attribute spectral effects of coarticulation to the context rather than the target segment, has led to the hypothesis that listeners adjust or 'compensate' for coarticulatory influences in identifying and discriminating speech sounds. We believe, consistent with Ohala's account and with our own account in previous work (e.g., Beddor et al. 1986, Krakow et al. 1988), that there is a link between failure to compensate for coarticulatory effects and perceptual confusions that give rise to sound change. At the same time, not all of the phonological and experimental data on nasal vowels are easily reconciled within this approach. Phonologically, we would expect that distinctive vowel nasalization would not occur in contexts in which the vowel is followed by a nasal consonant (because listeners should attribute the nasalization to context). Although such contrasts are rare, Hajek (1997) cites a number of languages with synchronic contrasts between [vN] and [ʔN]. We would also expect that contextual vowel nasalization would not give rise to vowel height shifts (again, because the context

should enable listeners to attribute the spectral properties of the first formant region to nasalization rather than tongue height), but the phonological literature provides considerable cross-linguistic evidence of height shifts in contextually nasalized vowels (Beddor 1982). Experimentally, there is some evidence that the mechanism responsible for perceptual compensation for coarticulatory nasalization is more complex than that suggested by the data of Kawasaki (1986) and Krakow et al. (1988). In earlier work related to the current study (Krakow and Beddor 1991), we found that listeners asked to judge oral and nasal vowels in different contexts did not respond in a way indicating they consistently compensated for the coarticulatory effects of a nasal context. We expect that a better understanding of the types of adjustments listeners make — or fail to make — when presented with contextual information relevant to a perceptual decision will lead to a better understanding of the role of perceptual compensation in the structure of phonological systems. The two experiments described below were designed to tease apart some of these issues as related to vowel nasalization.

3. Rating Nasality

In the first experiment, listeners were asked to rate the relative nasality of vowels in a pair of syllables. The original stimuli were taken from a recording of a male native speaker of American English who produced multiple tokens of bed, men, bode, and moan in isolation. Two tokens of each of the 4 words, matched as closely as possible for vocalic pitch and loudness, were selected for use in the perceptual tests. Vowels were excised from their original contexts and were (a) spliced into the consonantal context of the other token of that word type (e.g. the vowel of bed₁ was spliced into the consonantal context of bed₂), (b) spliced into the consonantal context of the 'opposite' word type (e.g. the oral vowel of bed₁ was spliced into the nasal context of men₁), or (c) left excised from context (e.g. isolated e). (Excising did not affect the integrity of flanking nasal consonants: nasal consonant duration was sufficiently long that, even when vowel nasality was removed, [m] and [n] were unambiguously nasal.)

The splicing method yielded, for each vowel set (/e/ or /o/), 2 tokens of each of 6 syllable types, as in Table 1. Within a given vowel set, all possible pairings of stimuli (including identity, included as a control condition) were created. To illustrate, pairs for the /e/ set are shown in Table 2; the order of the pair members was counterbalanced in the actual test sequences. (For certain pair types in the /e/ set, the duration of spliced vowels was also manipulated. Lax /e/ — but not tense /o/ — was consistently shorter in nasal [m_n] than in oral [b_d] contexts. So that contextual appropriateness of vowel length would not be a factor in listeners' judgments, oral vowels were shortened to match nasal vowel duration in [mēn] - [men] pairs and nasal vowels were lengthened to match oral vowel duration in [bed]-[bēd] and [e]-[ē] pairs.

We presented the rating test to 16 native speakers of American English who were asked to decide for each pair which member had the more nasal vowel, or whether the pairs had equally nasal / non-nasal members. The test was blocked into /e/ and /o/ sets, with half of the listeners responding first to the /e/ set and the other half responding first to the /o/ set.

/e/ set:	[bed]	[bēd]	[men]	[mēn]	[e]	[ē]
/o/ set:	[bod]	[bōd]	[mon]	[mōn]	[o]	[ō]

Table 1. Stimulus syllable types

Same-context pairs:

[bed]-[bed]	[mẽn]-[mẽn]	[e]-[e]
[bed]-[bẽd]	[men]-[mẽn]	[e]-[ẽ]
[bẽd]-[bẽd]	[men]-[men]	[ẽ]-[ẽ]

Cross-context pairs:

[bed]-[mẽn]	[bed]-[e]	[mẽn]-[ẽ]
[bed]-[men]	[bed]-[ẽ]	[mẽn]-[e]
[bẽd]-[men]	[bẽd]-[ẽ]	[men]-[ẽ]
[bẽd]-[mẽn]	[bẽd]-[e]	[men]-[e]

Table 2. Rating pairs for the /e/ set

In presenting the results, we evaluate them relative to the *expected* results based on a compensation hypothesis which states that, in English (a language typically analyzed as having non-distinctive vowel nasalization), listeners attribute the effects of nasalization on the vowel spectrum to a flanking nasal consonant rather than to the vowel itself. The hypothesis predicts that listeners should make mistakes judging the nasality of pairs with vowels in nasal consonant contexts (except under the identity condition), but they should do relatively well judging vowel nasality in oral contexts and isolation. Fig. 1 gives the pooled percent correct responses of same-context pairs whose vowels differed in nasality, with /e/ results in the left panel and /o/ results in the right. (Results for the control — i.e., identity — pairs are not presented, but averaged 83% correct across the 6 conditions.) Clearly, the general prediction is upheld: listeners' responses hover around 80% correct (n.b. close to control-condition accuracy) in the non-nasal contexts, but drop to around 40% correct in the nasal context. In addition, the /e/ and /o/ sets elicited similar patterns of responses. For this reason, results for the two vowel sets are collapsed for the remainder of the comparisons.

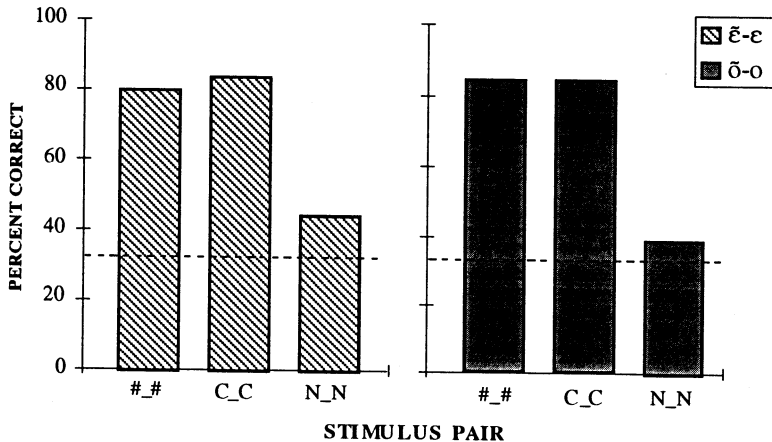


Fig. 1. Pooled percent correct rating responses to oral-nasal vowel pairs in identical contexts (striped bars for /e/ and solid for /o/). Chance performance (horizontal dashed line) is 33%.

<u>Rating pair</u>	<u>Correct response</u>	<u>Predicted (correct) response</u>
[bed]-[e]	'same'	'same'
[bēd]-[ē]	'same'	'same'
[bed]-[ē]	[ē]	[ē]
[bēd]-[e]	[bēd]	[bēd]

Table 3. Predicted rating responses to cross-context comparisons involving non-nasal contexts

The cross-context comparisons provide additional insights into what occurs when vowels in pairs have the same, or different, nasality. The specific predictions for cross-context comparisons involving *non-nasal* contexts, where listeners should continue to perform accurately, are illustrated in Table 3 for the /e/ set. The obtained results for these comparisons, shown in Fig. 2, support these predictions: as with the same-context comparisons, listeners' cross-context performance on vowels in non-nasal contexts remain at roughly 80% correct.

Turning to vowels in *nasal* contexts, the working hypothesis predicts that listeners should perform more poorly on cross-context comparisons in which one vowel is in a nasal context and the other is not. Because contextual information should lead listeners to compensate for nasality in the former case but not the latter, they should choose the wrong outcome for pairs involving [mēn] or [mōn]. We adopt, for the sake of exposition, a strong version of the compensation hypothesis which claims that American English listeners will attribute all of the nasalization on a vowel to a nasal context, if present. The resulting predictions are illustrated in Table 4 for [mēn]. The obtained results in Fig. 3 show that listeners have considerable difficulty judging the nasality of cross-context comparisons that involve [NŃN]: the near-chance performance on these nasal context pairs is a strikingly different pattern of results from that seen in Fig. 2 for non-nasal contexts.

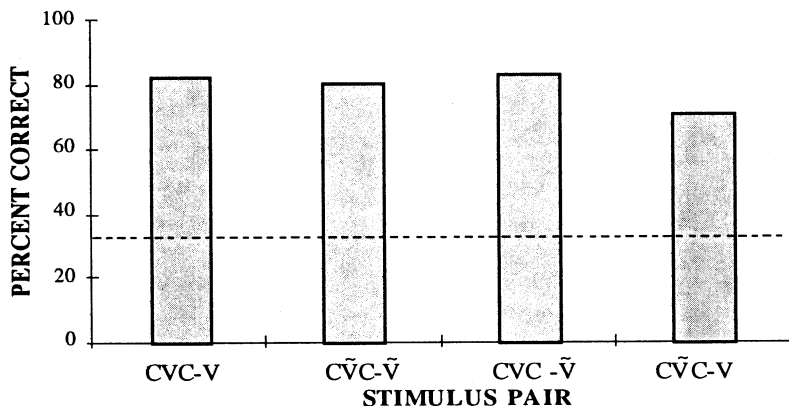


Fig. 2. Pooled percent correct rating responses to cross-context comparisons involving non-nasal contexts. Chance performance (horizontal dashed line) is 33%.

<u>Rating pair</u>	<u>Correct response</u>	<u>Predicted (incorrect) response</u>
[mēn]-[e]	[mēn]	'same'
[mēn]-[bed]	[mēn]	'same'
[mēn]-[ē]	'same'	[ē]
[mēn]-[bēd]	'same'	[bēd]

Table 4. Predicted rating responses to cross-context comparisons with [mēn]

What about *oral* vowels in *nasal* consonant contexts? Here we are less certain what to predict because we can imagine making a case for each of two conflicting hypotheses, both of which draw on the relation between production and perception. First, it seems reasonable to predict that listeners would accurately identify a vowel as oral in the coarticulatorily inappropriate context [N_N], much as we predicted them to be accurate at identifying a vowel as nasal in the coarticulatorily inappropriate context [C_C]. In such cases, lacking contextual support for presence ([CVC]) or absence ([NVN]) of vowel nasalization is thought to render vowel nasality (or lack of nasality) perceptually obvious. Second, we might predict that listeners would be inaccurate at identifying a vowel as oral in a nasal context, parallel to the prediction that a nasal vowel will sound oral in a nasal consonant context (due to compensation for the contextual effect). Perhaps, for the [NVN] stimuli, listeners attempt to compensate for the expected effects of the nasal context, even though those effects are not present. Such compensation would make the vowel sound different than an oral vowel in an oral context or a nasal vowel in a nasal context (and the quality of such a vowel might be reported to be 'nasal' in the same way that individuals tend to describe the hyponasal speech of individuals with colds as 'nasal'). The first hypothesis predicts performance on pairs with [NVN] to be as accurate as performance on pairs with [CVC]; the second predicts [NVN] performance to be as poor as [N̄VN]. Interestingly, the results in Fig. 4 show that accuracy on pairs with [NVN] usually falls in between that of the other two types of pairs.

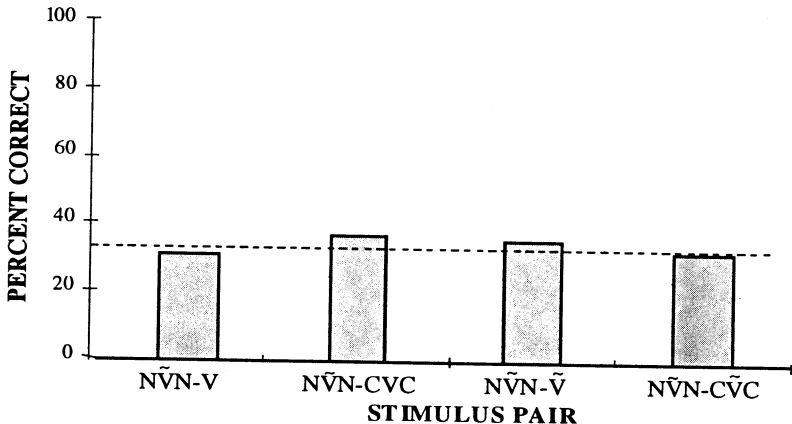


Fig. 3. Pooled percent correct rating responses to cross-context comparisons involving nasal vowels in nasal contexts. Chance performance (horizontal dashed line) is 33%.

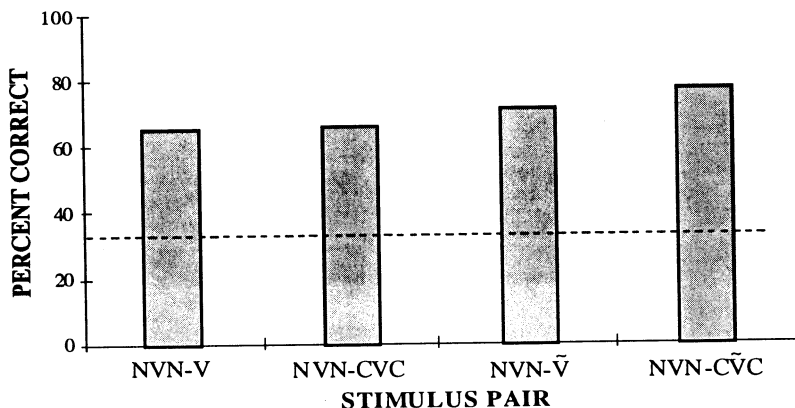


Fig. 4. Pooled percent correct rating responses to cross-context comparisons involving oral vowels in nasal contexts. Chance performance (horizontal dashed line) is 33%.

The general pattern of results is consistent with an account which suggests that listeners attribute vowel nasalization in a nasal consonant context to the context. Specifically, listeners are least likely to give the correct answer when pairs involve nasal vowels in nasal contexts (average: 34% correct). Listeners are increasingly likely to give correct answers in cross-context pairs as we progress from matches involving oral vowels in nasal contexts (average: 70% correct), to nasal vowels in oral contexts (average: 75% correct), to oral vowels in oral contexts (average: 84% correct).

However, a complex picture emerges when we consider the detailed pattern of responses shown by analysis not of percent correct responses, but rather of percent *choice* — that is, how often listeners chose each of the 3 options in a given pair type. Fig. 5 presents the 3-way response breakdown for the 4 cross-context pairs with a nasal vowel in a nasal consonant context from Fig. 3. Recall from the predictions above that the two left-most pairs, involving oral-to-nasal vowel pairings, were predicted to elicit ‘same’ responses. ‘Same’ was, indeed, the preferred choice for these two. On the other hand, Fig. 5 shows that when listeners chose one of the other options, they were more likely to make the correct nasality choice. That listeners were able to make this distinction suggests that the strong version of the compensation hypothesis is not correct.

The two right-most pairs in Fig. 5 are those [ÑVN] pairs predicted not to elicit ‘same’ responses (i.e., nasal-to-nasal vowel pairs in which one pair member was contextually appropriate and the other was not). Although ‘same’ was, in fact, not listeners’ most common choice, listeners did select the (correct) ‘same’ option approximately one-third of the time, indicating that compensation did not consistently lead them to make an incorrect nasality assessment.

Overall, performance on the rating task supports a modified version of the compensation hypothesis. When asked to judge the relative nasality of two vowels, listeners perform most poorly when one of the vowels is a nasal vowel in a nasal context, which is what we would expect if perceptual compensation leads listeners

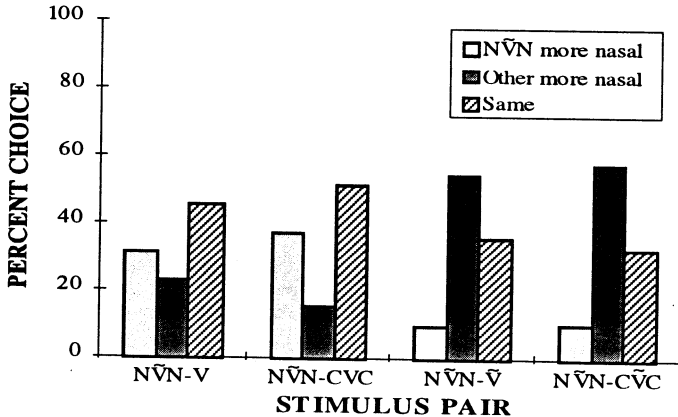


Fig. 5. Pooled percent choice responses to cross-context comparisons with nasal vowels in nasal contexts: frequency of listeners' choice of each of the 3 response options.

to attribute contextual vowel nasalization to context. At the same time, listeners do not respond in a way which indicates they usually hear nasal vowels in nasal contexts as oral, suggesting that perhaps the nasal context leads listeners to compensate *partially* for the effects of nasalization on the vowel spectrum. However, it may be premature to reach this conclusion solely on the basis of listeners' choices in a rating task, which requires listeners to make metalinguistic judgments of relative nasality. It is possible that the nature of the task interferes with our ability to determine the conditions under which oral and nasal vowels sound similar or different.

4. Discriminating Nasality

To address these issues further, we supplemented the rating task with a discrimination task which does not require listeners to make explicit nasality assessments. The same stimuli were arranged into a 4IAX discrimination paradigm in which each trial consisted of 2 stimulus pairs. In one pair, the vowels differed in nasality; in the other pair, the vowels had the same nasality and were, in fact, acoustically identical. (To preserve acoustic identity, the length manipulation applied to some stimulus pairs in the rating task was not used here.) To illustrate, discrimination trials for the /e/ set are given in Table 5. As before, the tasks include the /e/ and /o/ sets in separate blocks. (Note that not all possible trials were included: for trials involving vowels in consonantal contexts, all trials involved one pair in which vowel nasality was coarticulatorily appropriate. Consequently, we excluded trials such as [bēd]-[mēn]/[bēd]-[men] or [bed]-[bēd]/[bēd]-[bēd].)

The discrimination task was presented to the same 16 American English-speaking listeners who participated in the rating task. Listeners were told to select, for each trial, the pair whose vowels were more different. (Listeners took the discrimination test before the rating test, so as to avoid potential influence of nasality ratings on discrimination judgments.)

Same-context trials:

[bed]-[bēd] / [bed]-[bed]
[men]-[mēn] / [mēn]-[mēn]

[ε]-[ē] / [ε]-[ε]
[ε]-[ē] / [ē]-[ē]

Cross-context trials:

[bed]-[ē] / [bed]-[ε]	[men]-[ē] / [men]-[ε]	[bed]-[mēn] / [bed]-[men]
[bēd]-[ē] / [bed]-[ε]	[men]-[ē] / [mēn]-[ē]	[bed]-[mēn] / [bēd]-[mēn]
[bed]-[ē] / [bēd]-[ē]	[mēn]-[ε] / [men]-[ε]	
[bēd]-[ε] / [bēd]-[ē]	[mēn]-[ε] / [mēn]-[ē]	

Table 5. Discrimination trials for the /e/ set

Applying the same working hypothesis of perceptual compensation to this task again leads to the general prediction that listeners will perform more poorly on cross-context trials involving nasal vowels in nasal contexts than on other types of trials. (Accurate performance was expected, and obtained, on all same-context trials: mean performance on each of the 4 types of same-context trials was above 95% correct.) As seen in the results presented in Fig. 6, the prediction of good performance on cross-context trials with vowels in non-nasal contexts was upheld, with performance ranging from 85-91% correct across the 4 conditions.

Predicting discrimination of vowels in nasal contexts, beyond the general expectation that such trials should be harder to discriminate than those represented in Fig. 6, is again straightforward only for those trials with [N̄VN] that do not also include [NVN]. Given the possible groupings of our stimuli, this means that only 2 of the 6 trials with [N_N] yield clear predictions, as in Table 6.

The results for trials with [N_N] - isolated vowel pairings (corresponding to the first 4 rows of Table 6) are given in Fig. 7. Overall performance is poorer than that on the [C_C] - isolated vowel pairings in Fig. 6, dropping by an average of 16% and consistent with the general prediction. Comparing performance within the set of [N_N] - isolated vowel trials, Fig. 7 shows that the two trial types involving

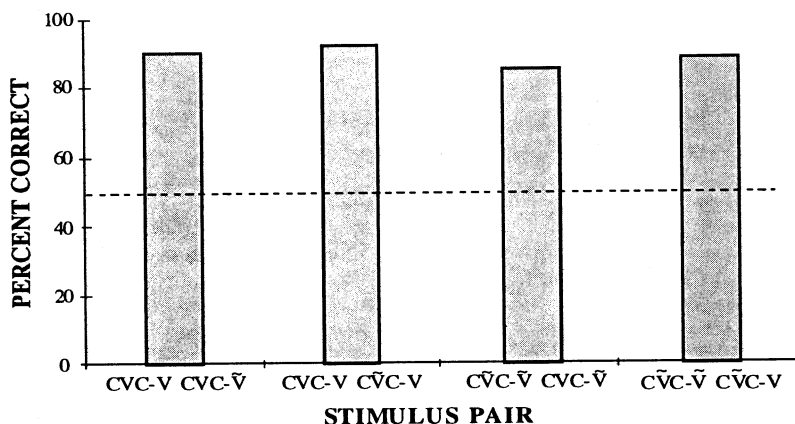


Fig. 6. Pooled percent correct discrimination responses to [C_C] - isolated vowel comparisons. Chance performance (horizontal dashed line) is 50%.

<u>Trial type</u>	<u>Correct choice</u>	<u>Predicted choice</u>
[men]-[ē] / [men]-[e]	[mēn]-[e]	?
[men]-[ē] / [mēn]-[ē]	[men]-[ē]	?
[mēn]-[e] / [men]-[e]	[mēn]-[e]	?
[mēn]-[e] / [mēn]-[ē]	[mēn]-[e]	[mēn]-[ē]
[bed]-[mēn] / [bed]-[men]	[bed]-[mēn]	?
[bed]-[mēn] / [bēd]-[mēn]	[bed]-[mēn]	[bēd]-[mēn]

Table 6. Predicted discrimination choices for cross-context comparisons with [mēn]

[N̄VN] - [V] pairs (the two right-most bars) were more difficult than trials not involving this comparison. It is interesting to note that this is the only pairing involving the [N_N] context and an isolated vowel in which the nasality of both vowels is contextually appropriate.

Turning our attention to discrimination of the consonantal context comparisons, we see in Fig. 8 that the contextually appropriate pairing of [N̄VN] and [CVC] is like that of [N̄VN] and [V] in that it hindered performance. However, the trials with [CVC] proved to be even more difficult. Discriminability of these trials was at chance regardless of whether the contextually appropriate vowels ([N̄VN] - [CVC]) were compared with acoustically identical nasal ([N̄VN] - [C̄VC]) or oral ([NVN] - [CVC]) vowels.

5. Interpreting the Results

Overall, the rating and discrimination results exhibit similar response patterns. Vowel judgments in the two experimental paradigms showed the same relative order of difficulty for the different contextual conditions. Judgments of vowels in nasal contexts were more difficult (i.e., less accurate) than those of vowels in non-nasal contexts (oral contexts or isolation). Most difficult of all were judgments of nasal vowels in nasal contexts. Listeners responding to the rating task performed at

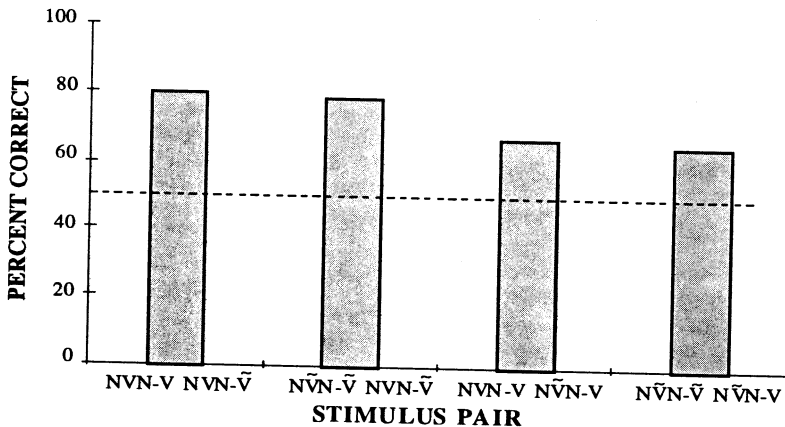


Fig. 7. Pooled percent correct discrimination responses to [N_N] - isolated vowel comparisons. Chance performance (horizontal dashed line) is 50%.

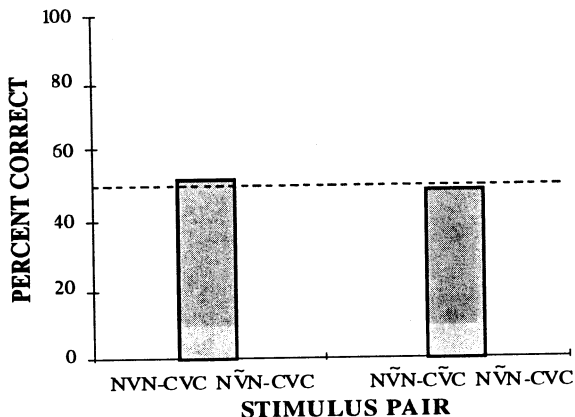


Fig. 8. Pooled percent correct discrimination responses to [N_N] - [C_C] comparisons. Chance performance (horizontal dashed line) is 50%.

chance level when judging the nasality of [N̄VN] relative to [CVC], [V], and [̄V] (Fig. 3), and close to chance level relative to [NVN] (Fig. 1). Along similar lines, discrimination scores were at chance level for trials involving [N̄VN] - [CVC] comparisons (Fig. 8) and were next lowest for [N̄VN] - [V] comparisons (compare the right-most bars of Fig. 7 to the remainder of Fig. 7 and Fig. 6). These findings are generally consistent with a compensation hypothesis: when a target speech sound is in a coarticulatorily appropriate context, listeners' knowledge of coarticulatory organization lets them attribute the consequences of overlapping articulation on the target's spectrum to the coarticulatory source. Perceptual compensation is thought to *facilitate* speech perception under usual circumstances involving natural communicative settings. However, under laboratory conditions requiring listeners to make nasality judgments or to discriminate nasality, attributing spectral properties of the target to the context will give rise to listener *mistakes*: listeners attribute vowel nasality in a [N̄VN] sequence to the nasal context and judge that vowel as less nasal than the same vowel in non-nasal contexts. This general outcome of the rating and discrimination tests is consistent with Kawasaki's (1986) findings using an experimental paradigm in which flanking nasal consonants were gradually attenuated and with our earlier (1991) findings using a matching paradigm.

Importantly, however, when listeners judge [̄V] as less nasal in nasal contexts than in oral contexts or isolation, this does not mean that they hear the [N̄VN] vowel as *oral*. Indeed, the rating and discrimination results argue strongly against this interpretation. For example, in the rating task, if listeners heard contextually nasalized vowels as oral, they should have systematically chosen the incorrect option (in comparison with vowels in [C_C] context or isolation), but their performance was around chance rather than substantially below chance. In the discrimination task, if listeners heard the vowel in [N̄VN] as oral, they should have systematically selected acoustically identical vowels in [N̄VN] - [̄V] and [N̄VN] - [C̄VC] pairs as more different than non-identical but coarticulatorily appropriate vowels in

[N̩V] - [V] and [N̩N] - [CVC] pairs. But here again, their performance was at or above chance.

Our findings clearly indicate that American English listeners do not - at least not consistently - hear contextually nasalized vowels as oral. At first blush, the data might seem to lend themselves to the interpretation that the chance and near-chance performance on some pairs involving [N̩N] stimuli (in both the rating and cross-context discrimination tasks) simply means that listeners are not certain about the nasality of such vowels and are therefore responding inconsistently. However, when we consider the entire dataset, we get a somewhat different view. For example, consider again the rating task comparisons involving [N̩N] paired with [CVC] or [V]. The detailed results (Fig. 5) showed that subjects selected [N̩N] as the 'more nasal' item considerably more frequently than they selected the oral alternative. Furthermore, performance on the discrimination task was around 70% correct for comparisons involving the contextually appropriate pair [N̩N] - [V]. Taken together, the results from the two experiments are most compatible with a view of partial compensation, that is, that listeners attribute some, but not all, of vowel nasalization in an [N̩N] sequence to the context.

6. Implications for Phonological Change

What are the implications of these findings for an account of phonological change which attributes (at least some of) the perceptual confusions leading to change to listeners' failure to compensate for coarticulatory effects? The account offered earlier in this paper (section 2), based largely on work by Ohala (1981, 1986, 1993 and others), predicts the occurrence of perceptual confusions when the conditioning environment for the spectral property at issue is not detected: in the absence of a coarticulatory source for that property, listeners analyze the property (e.g. nasalization) as an inherent characteristic of the target sound (e.g. vowel) rather than due to its contextual source (nasal consonant). On the other hand, clearly audible contextual information provides the requisite source; by enabling listeners to compensate perceptually for coarticulatory influences, source detection should block, or substantially reduce the likelihood of, confusion.

The perceptual data presented here point toward the need to modify this account. These findings indicate that unambiguous contextual information need not disambiguate the intended target sound. Indeed, listeners' inconsistency in judging the nasality of vowels in nasal contexts in some respects suggests that presence of the coarticulatory source leads to confusion rather than disambiguation. As just argued, however, patterns in the data indicate that listeners were not so confused in their responses as they were intermediate (nasal vowels in nasal contexts sounded less nasal than in oral contexts, but less oral than oral vowels). This pattern of partial compensation is not entirely surprising. If we assume that the listener weighs in all available information in making perceptual decisions (e.g. Hawkins 1995), then contextual information is but one factor in these decisions. Sometimes this information is not enough to disambiguate the acoustic signal. This may be particularly true of a property like contextual nasalization in American English, which can vary as a function of a variety of non-segmental factors such as syllable position and stress (Schourup 1973; see also Krakow 1993) as well as speaking rate (Kent et al. 1974, Bell-Berti and Krakow 1991, Krakow submitted), and also exhibits idiolectal and dialectal variation. Furthermore, the extent of contextual nasalization can vary as a function of vowel height (Clumeck 1976; this finding is

consistent with vowel height effects on velar height, see Bell-Berti 1993 for a review).

Viewed in this way, one role that perception should play in phonological change would involve not so much a breakdown or failure of normal perceptual processes, but rather a shift in the relative weighting of factors contributing to the linguistic percept. When contextual information enables listeners to compensate only partially for coarticulatory effects, the spectral property is perceived as belonging to both the context and the target sound. If, indeed, this hypothesis of partial compensation is correct, then perceptually motivated sound changes in which the conditioning environment is *not* lost or particularly weak should be expected. For example, partial compensation for vowel nasalization in nasal consonant contexts leads, in principle, to at least two predictions. One prediction is that, since our findings indicate that nasalization can be perceived as belonging to a vowel even when the final consonant is clearly evident to the listener, distinctive nasal vowels could emerge in VN sequences without concomitant loss of the final nasal. A second prediction is that, if vowels in nasal contexts are perceived as (partially) nasalized, listeners should sometimes fail to disambiguate fully the spectral contribution of nasalization and tongue/jaw position in both contextual and distinctive nasal vowels, fostering vowel height shifts in both.

Reconsidering the phonological information provided in sections 1 and 2 in light of these predictions, we find that the evidence is consistent with the second prediction. Although the magnitude of the height shifts appears, in some cases, to be greater in distinctive than in contextual nasal vowels, there is no obvious difference in the frequency of occurrence (Beddor 1982). The situation corresponding to the first prediction is relatively rare: although there are languages exhibiting [VN] - [ʋN] contrasts, the usual historical development for contrastive nasalization involves nasal consonant loss. We expect this is linked to the fact that nasalization before final nasal consonants tends to be heavier than after initial nasals and that final consonants, including nasals, tend to have weaker oral tract constrictions than initial consonants (Krakow 1989, submitted). These factors conspire to make a final nasal consonant less perceptible than an initial nasal, especially in normal speaking (i.e., casual speech) conditions in which consonant constrictions can be further weakened (Krakow submitted) and anticipatory velar lowering more extensive (Bell-Berti and Krakow 1991).

What is our answer to the question posed in the title of the paper: how confused is the listener? The finding that listeners may partially compensate for a spectral property on a target sound when provided with salient contextual information suggests that listeners may be both more and less uncertain of the articulatory events that gave rise to the acoustic signal than previously thought. Listeners are less uncertain in that they can perceive a non-contrastive property as present on a target sound even when the property's coarticulatory source is also perceived. On the other hand, listeners' detection of the coarticulatory source need not fully disambiguate the acoustic signal; that is, such detection is apparently not a sufficient condition for unambiguous perception. We believe that perceptual ambiguity or partial compensation of the kind observed here is due largely to the apparent difficulty of perceptually resolving the spatial magnitude or temporal extent of gestural overlap among segments - not an unexpected difficulty in view of the considerable range of possible variation in these aspects of coarticulation under normal speaking conditions.

Acknowledgments

The research reported here was supported in part by NSF Grant SBR 9319597 to Patrice Beddor and NIH Grant NS-131617 to Haskins Laboratories. We wish to thank James Harnsberger for assistance in data collection, and the University of Michigan phonetics-phonology discussion group and participants in the BLS 24 Parasession for helpful comments.

References

- Beddor, P. S. 1982. Phonological and phonetic effects of nasalization on vowel height. Doctoral dissertation, University of Minnesota, Minneapolis. (Reproduced by Indiana University Linguistics Club, 1983).
- , and D. Evans-Romaine. 1995. Listener misperceptions as a source of phonological assimilations: a study of syllable-final nasals. *Rivista di Linguistica* 7:1.145-74.
- , and S. Hawkins. 1990. The influence of spectral prominence on perceived vowel quality. *J. Acous. Soc. of Am.* 87.2684-2704.
- , R. A. Krakow, and L. M. Goldstein. 1986. Perceptual constraints and phonological change: a study of nasal vowel height. *Phonology Yearbook* 3.197-217.
- Bell-Berti, F. 1993. Understanding velic motor control: studies of segmental context. Nasals, nasalization, and the velum, ed. by M. K. Huffman and R. A. Krakow, 63-85. Orlando, FL: Academic Press.
- , and R. A. Krakow. 1991. Anticipatory velar lowering: A coproduction account. *J. Acous. Soc. of Am.* 90.113-23.
- Bhat, D. 1975. Two studies on nasalization. Nasálfest: papers from a symposium on nasals and nasalization, ed. by C. Ferguson, L. Hyman, and J. J. Ohala, 27-48. Stanford: Stanford University.
- Chen, M. 1973. Nasals and nasalization in the history of Chinese. Doctoral dissertation, University of California, Berkeley.
- Clumbeck, H. 1976. Patterns of soft palate movements. *J. of Phonetics* 4.337-51.
- Hajek, J. 1997. Universals of sound change in nasalization. Boston: Blackwell.
- Hawkins, S. and K. N. Stevens. 1985. Acoustic and perceptual correlates of the non-nasal—nasal distinction for vowels. *J. Acous. Soc. of Am.* 77.1560-75.
- Hawkins, S. 1995. Arguments for a non-segmental view of speech perception. Proceedings of the 13th International Congress of Phonetic Sciences, vol. 3.18-25. Stockholm.
- Hura, S. L., B. Lindblom, and R. L. Diehl. 1992. On the role of perception in shaping phonological assimilation rules. *Language and Speech* 35.59-72.
- Kawasaki, H. 1986. Phonetic explanation for phonological universals: the case of distinctive vowel nasalization. Experimental phonology, ed. by J. J. Ohala and J. J. Jaeger, 81-103. Orlando, FL: Academic.
- Kent, R. D., P.J. Carney, and L. R. Severeid. 1974. Velar movement and timing: evaluation of a model of binary control. *J. of Speech and Hearing Research* 17. 470-88.
- Kingston, J. 1991. Integrating articulations in the perception of vowel height. *Phonetica* 48.149-79.
- , and N. A. Macmillan. 1995. Integrality of nasalization and F_1 in vowels in isolation and before oral and nasal consonants: a detection-theoretic application of the Garner paradigm. *J. Acous. Soc. of Am.* 97.1261-85.

- Krakow, R. A. 1989. The articulatory organization of syllables: a kinematic analysis of labial and velar gestures. Doctoral dissertation, Yale University, New Haven, CT.
- . 1993. Nonsegmental influences on velum movement patterns: syllables, sentences, stress, and speaking rate. Nasals, nasalization, and the velum, ed. by M. Huffman and R. Krakow, 87-113. New York: Academic Press.
- . Submitted. Articulatory organization of syllables: a review.
- , and P. S. Beddor. 1991. Coarticulation and the perception of nasality. Proceedings of the 12th International Congress of Phonetic Sciences, vol. 5.38-41. Aix-en-Provence.
- , P. S. Beddor, L. M. Goldstein and C. A. Fowler. 1988. Coarticulatory influences on the perceived height of nasal vowels. J. Acous. Soc. of Am. 83.1146-58.
- Maeda, S. 1993. Acoustics of vowel nasalization and articulatory shifts in French nasal vowels. Phonetics and phonology: nasals, nasalization and the velum, ed. by M. K. Huffman and R. A. Krakow, 147-67. Orlando, FL: Academic Press.
- Manuel, S. Y. 1991. Some phonetic bases for the relative malleability of syllable-final versus syllable-initial consonants. Proceedings of the 12th International Congress of Phonetic Sciences, vol. 5.118-21. Aix-en-Provence.
- Mohr, B. and W. S.-Y. Wang. 1968. Perceptual distance and the specification of phonological features. *Phonetica* 18.31-45.
- Ohala, J. J. 1981. The listener as a source of sound change. Papers from the parasession on language and behavior, ed. by C. S. Masek, R. A. Hendrick, and M. F. Miller, 178-203. Chicago: Chicago Linguistic Society.
- . 1986. Phonological evidence for top-down processing in speech perception. Invariance and variability of speech processes, ed. by J. S. Perkell and D. H. Klatt, 386-401. Hillsdale, NJ: Erlbaum.
- . 1990. The phonetics and phonology of aspects of assimilation. Papers in Laboratory Phonology I: between the grammar and physics of speech, ed. by J. Kingston and M. Beckman, 258-75. Cambridge, UK: Cambridge University Press.
- . 1993. Coarticulation and phonology. *Language and Speech* 36.155-70.
- Repp, B. H. and K. Svastikula. 1988. Perception of the [m]-[n] distinction in VC syllables. *J. of the Acoustical Society of America* 83.237-47.
- Ruhlen, M. 1978. Nasal vowels. Universals of human language, vol. II: Phonology, ed. by J. Greenberg, C. Ferguson, and E. Moravcsik, 203-42. Stanford: Stanford University Press.
- Schourup, L. 1973. A cross-language study of vowel nasalization. *Ohio State Working Papers in Linguistics* 15.190-221.
- Seitz, P., M. McCormick, I. Watson, and R. A. Bladon. 1990. Relational spectral features for place of articulation in nasal consonants. *J. Acous. Soc. of Am.* 87.351-58.
- Tuttle, E. F. 1991. Nasalization in Northern Italy: syllabic constraints and strength scales as developmental parameters. *Rivista di Linguistica* 3.23-92.
- Wright, J. T. 1986. The behavior of nasalized vowels in the perceptual vowel space. Experimental phonology, ed. by J. J. Ohala and J. J. Jaeger, 45-67. Orlando, FL: Academic Press.

Functional Bases of Phonological Universals: A Connectionist Approach

Marc F. Joanisse and Mark S. Seidenberg
University of Southern California¹

1 Introduction

The phonological inventories of the world's languages exhibit non-random patterns relating to how phonemes are grouped into natural classes, the types of processes targeting these classes, the direction of historical change, and the frequency with which phonemes occur and co-occur in languages. The present study investigated an important class of regularities concerning the distributions of vowels. Although the human vocal apparatus can produce many possible vowels, a large proportion of languages only use between 4 and 8 of them. In addition, languages with a given number of vowels tend to use similar sets of vowels. For example, most five-vowel languages employ the set [i e a o u]; a handful use similar inventories such as [i e a ə u]; and many possible sets of vowels are not observed at all, such as [e y œ æ u]².

The standard approach within the generative tradition has been to view these phenomena in terms of the concept of markedness: vowel features are organized into a markedness hierarchy, such that vowels incorporating more marked features are less suitable in an inventory (Chomsky & Halle 1968; Clements 1985). The major drawback to this approach is the lack of criteria independent of mere frequency of occurrence for determining which vowels or features are "marked."

Our research addresses the hypothesis that vowel inventory patterns reflect functional constraints related to perception and production; specifically, languages tend to maximize distances between vowels. We focus here on the tendency to maximize the acoustic distances between constituent vowels, although featural or gestural distance may also be relevant. Inventories involving acoustically well-dispersed vowels are easier to both acquire and process because they are easier to discriminate, creating a tendency for languages to recruit such inventories. In contrast, less acoustically dispersed vowel inventories are more difficult to acquire and process because of the greater probability of misperceiving a constituent vowel, leading to languages shifting away from such inventories. On this view, inventories such as [i I e e a] do not occur because they involve smaller distances than other, attested vowel sets (Jakobson 1941).

A similar idea has been explored using mathematical models (Liljencrants & Lindblom 1972; Lindblom 1986; Boë, Schwartz, & Valée 1994) in which the acous-

tic distances of vowels are expressed as a function of the Euclidean distance between their corresponding formant frequencies. Such models predict perceptually optimal inventories by maximizing the distances between all vowels in a given set. Although this approach can account for a considerable amount of data about the distributions of vowels, it is limited in several respects. First, it does not represent the variability with which vowels are produced; as we shall see, this variability can play a role in the frequency with which a vowel occurs in an inventory. Second, this approach is not a model of why this type of distance maximization occurs; we tie this optimization to constraints related to learning and processing. Finally, this approach does not easily allow the integration of other types of factors used to differentiate vowels, such as nasalization, diphthongization and vowel length.

Connectionist Models of Phoneme Acquisition

The present work builds on Lindblom's approach by situating it in a theory of how phonemic inventories are acquired and processed. In Joanisse & Seidenberg (1997) we described research in which connectionist models are trained to recognize the phonemes of pseudo-languages consisting of different vowel inventories. The approach is based on the premise that, like humans, connectionist models are not equally predisposed to learning and recognizing all types of patterns. By varying the characteristics of these vowel inventories and assessing the models' capacities to learn them, we generate predictions about the relative suitability of inventories. Inventories that are easier to learn are predicted to be more likely to occur in the languages of the world. Inventories that cannot be learned are predicted to not occur.

The general structure of the models used in the present research is illustrated in Figure 1. The ellipses represent groups of artificial neurons that encode information as patterns of activation. Lines between layers of units represent sets of connections through which activation is passed, in the direction of the arrows. The middle, so-called hidden, layer of units is used to enhance the representational capacity of the network³.

The model is trained to recognize speech sounds based on a set of examples. The input to the model is the spectrographic representation of a given speech item. Identification results from passing activation through two layers of connections to the output layer. Learning proceeds through the adjustment of connection weights using the backpropagation learning algorithm (Rumelhart, Hinton, & Williams 1986); over the course of training, the model adjusts these weights in ways that facilitate accurate identification of the training stimuli and generalization to novel examples. The relative difficulty of learning the training set is reflected in the rate at which training proceeds, asymptotic levels of performance, and the number of items a trained network misclassifies.

We used this model to explore two types of vowel inventory phenomena. The

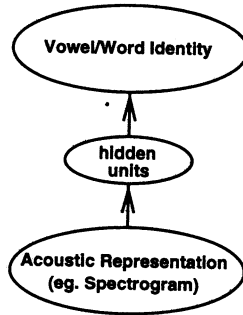


Figure 1: General structure of the models used in the present paper.

first relates to the preference of four-vowel languages to choose front mid vowels over back mid vowels, which we propose is related to differences in these vowels' variability. In the second set of simulations, we explore the interaction between the number of vowel quality contrasts in a language, and its tendency to use contrastive length. It is shown that both factors affect ease of learning and processing in the model, providing an explanation for why certain inventories are preferred.

2 Experiment 1: Front-Back Asymmetries

The first set of simulations concerns the observation that, in four vowel languages, there is a greater tendency for languages to use a front mid vowel than a back mid vowel, as illustrated in Figure 2. This is an interesting asymmetry, given that the dispersion characteristics of the two inventories are approximately equal. A simple dispersion model as in Liljencrants & Lindblom (1972) predicts there to be little difference in the occurrence of the two inventory types, but this is contradicted by the empirical data.

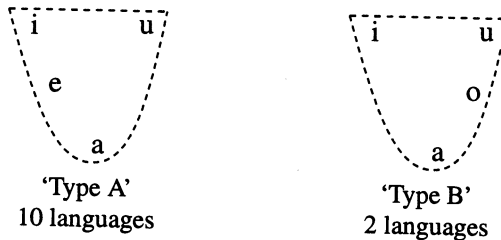


Figure 2: Asymmetry of back vowels over front vowels. Note that the relative dispersion of the two sets is roughly the same.

One explanation for this asymmetry derives from differences between front and back vowels with respect to their tendency to vary in production. As Figure 3 illus-

trates, there is a difference between nonlow front and back vowels in their tendency to overlap with each other. Speakers are able to produce the vowel /i/ with a better degree of precision, by stiffening the genioglossus muscle and propping the tongue laterally against the dental ridge (Beckman, Jung, Lee, de Jong, Krishnamurthy, Ahalt, & Cohen 1995). This is not possible for back vowels however, which leads to a greater F1 variability for /u/ than for /i/, causing more overlap for the /u/ – /o/ contrast than for the complementary /i/ – /e/ contrast.

This type of explanation is consistent with Stevens' Quantal Theory (Stevens 1989), in which a phoneme's distinctiveness is affected by nonlinearities in the relationship between the vocal tract and acoustics: phonemes with quantal articulations are those which can be produced with greater precision as a result of such nonlinearities. In the present account, a vowel inventory's frequency can be explained in part as the result of the overlap of its constituent vowels – inventories which minimize this overlap have more discriminable vowels, and are more likely to occur.

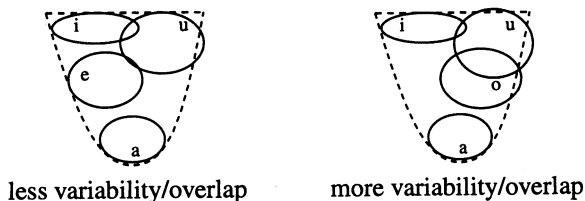


Figure 3: Schematization of vowel variability. Note the greater tendency for non-low back vowels to overlap, compared to their front counterparts.

Method & Stimuli

To test the hypothesis that production variability has an effect on vowel inventory frequencies, we trained networks on the invented vocabularies of one of two pseudolanguages. These pseudolanguages differed only with respect to their vowel inventories: the artificial vocabularies consisted of all CV combinations of the consonant set [p t k b d g] and either the vowel set [i e a u] (Language A) or [i a o u] (Language B). Training stimuli were devised by extracting CV syllables from the DARPA TIMIT Acoustic-Phonetic Continuous Speech Corpus of spoken English. Raw waveforms were transformed into fourteen 8 ms frames of 96 spectral coefficients (bandwidth: 62.5 Hz, frequency range: 0-6000 Hz) using the Fast Fourier transformation. The result was a set of 1,344 spectral coefficients for each CV syllable in the training set.

Using waveforms drawn from the TIMIT database allowed us to obtain many instances of each syllable type (in most cases, greater than 25 of each type) as spoken by many speakers of both sexes and of many American regional dialects. Although networks were not explicitly trained to identify vowels and consonants,

the identification task at hand would force these networks to learn to identify the consonant and vowel components of the input set in order to perform the task at optimal levels. This is consistent with what we believe to occur in children as they learn to identify the words in their target language. It is hypothesized that the greater degree of overlap between nonlow back vowels compared to nonlow front vowels will cause slower learning rates and poorer generalization in models trained on Language B, compared to Language A.

Network training consisted of 100,000 training trials. During each training trial, the model was presented with the spectral coefficients of a CV syllable randomly drawn from the training set, and trained to identify this syllable by activating the correct output unit. There were 24 output units, each corresponding to a different CV syllable type. All types were presented an equal number of times over the course of training.

Results and Discussion

Three different networks were trained on either Language A or B, for a total of six networks; averaged results for all six are reported here. To assess overall learning of training items, networks were tested at intervals of 10,000 training trials on the complete training set. This was done by presenting all items to the network, and calculating the resulting Sum Squared Error which is a function of the overall correct and incorrect activation of all output units (Rumelhart, Hinton, & Williams 1986). Results are plotted in Figure 4. Higher mean error rates for models trained on Language B indicate that these networks had more difficulty learning the training set consisting of the vowel set [i a o u] compared to those trained on [i e a u].

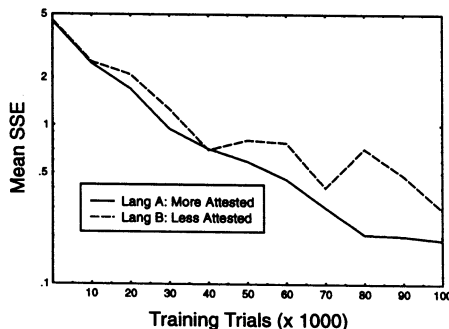


Figure 4: Sum Squared Error rates averaged over all networks trained on two pseudolanguages. Lower rates for Language A suggest less difficulty in learning the training set.

This result is consistent with the hypothesis that the learnability of a vowel

inventory affects its frequency of occurrence in the world's languages. In addition, these networks were capable of learning both pseudolanguages, consistent with the observation that both inventories are attested, although with different frequencies. Neither inventory is completely unlearnable within this architecture (unlike more extreme cases investigated in Joanisse & Seidenberg (1997), such as [i I e ε æ]). Given sufficient training, networks trained on either inventory can achieve near-perfect performance in categorizing items in the training set.

To further investigate how these inventories differ in their degree of suitability, we also tested networks on their ability to generalize to novel stimuli. This type of testing is comparable to the type of task confronting an adult language user, who must recognize novel tokens of familiar words in the course of everyday language processing. This was done by presenting fully trained networks with instances of CV syllables randomly withheld from the training set. Analyses of these errors indicated that networks trained on Language B misclassified novel vowels more frequently than Language A (Lang A: 71% correct; Lang. B: 65% correct).

Results from these simulations support the hypothesis that vowel systems are functionally optimized in ways that maximize the discriminability of their constituent vowels because of its effects on learning. In addition, these simulations suggest that accounting for these phenomena turns on examining stimuli that realistically represent variability in terms of production and overlap.

3 Experiment 2: Length and Quality Interactions

So far we have only considered vowel inventory tendencies in terms of quality contrasts related to formant frequencies. However, many languages use other types of cues to contrast vowels. Here we consider how one such cue, vowel length, interacts with spectral (formant-based) cues. Maddieson observes that 'The probability of length being part of the vowel system increases with the number of vowel quality contrasts' (Maddieson 1984:129). As such, 12.5% of languages with 4 to 6 vowel qualities and 24.7% of languages with 7-8 vowel qualities incorporate length contrasts, compared to 53.8% of languages with 10 or more vowel qualities. However, these data are incomplete, owing to the nature of the Maddieson (1984) database which does not consider contrastive length in cases where all vowel qualities participate in length contrasts.

We propose that this pattern is not accidental, and that a more extensive survey of length in vowel systems would reveal a similar pattern for languages not included in the Maddieson database. This pattern is attributed to the weak contrastiveness of durational cues for vowels, compared to spectral cues, owing to the degree of variability intrinsic to vowel length. Vowel length appears to be a useful cue in disambiguating various language contexts, for example in determining the voicing of an adjacent consonant (Chen 1970), rate of speech (Magen & Blumstein 1993) and a lexical boundary (Davis, Marslen-Wilson, & Gaskell 1997). Given this tendency

for vowels to vary in duration, it is plausible that adding a durational contrast would be dispreferred.

This analysis is additionally supported by the observation that length contrasts are often observed to accompany close quality contrasts (Maddieson 1984). This suggests that duration is a useful secondary cue to differentiating spectrally similar vowels, though on its own, it may prove to be less useful than a spectral contrast.

Connectionist models provide a way to explore these phenomena, because of their capacity to exploit multiple, simultaneous, probabilistic regularities in the service of learning to perform a task. The relative usefulness of cues can be assessed in terms of the model's ability to reliably learn and use them. In the present simulations, connectionist networks were trained to acquire vowel inventories that use spectral or length contrasts to different degrees. It was predicted that differences between network learning and generalization would reflect known facts about the occurrence of these cues in vowel systems of different sizes.

Method & Stimuli

Networks were trained on one of three vowel inventories which use length and spectral contrasts to different degrees. As Table 1 shows, these inventories seek to double the number of contrasts in the familiar [i e a o u] set by adding spectral contrasts, length contrasts, or both. For clarity, these inventories are also plotted in Appendix A.

<i>inventory</i>	<i>vowel set</i>
1 - quality only	[i ʌ e ε æ a ɔ o u]
2 - length	[i iː e eː a aː o oː u uː]
3 - both length and quality	[iː ʌ eː ε æː a ɔ oː u uː]

Table 1: Vowel inventories used in the present study. All seek to double the base [i e a o u] inventory, though in different ways.

Given the hypothesized differences in the contrastiveness of length and spectral cues, we predicted that networks would have more difficulty learning Inventory 1 compared to 2 and 3, because only spectral contrasts are used, and that Inventory 3 would be easier to acquire than Inventory 2 due to the interaction of duration and spectral cues in maintaining contrast in such a relatively large and crowded inventory. This would be consistent with the observation that a contrast like /i/ – /I/ is dispreferred compared to /i/ – /e/, though it might be more common than /i/ – /iː/. Ultimately, /iː/ – /I/ seems to represent a good compromise for languages with a crowded vowel space.

The architecture of the model was similar to those used in the previous simulations. Input consisted of 2040 spectral coefficients encoding the acoustic representation of a vowel. A total of 40 vowels of each type were devised for training

purposes. Training sets consisted of synthetic vowels, made to be highly realistic by using formant means and variances drawn from observed data in the TIMIT database and data published in Beckman et al. (1995). Contrastive vowel length was simulated by creating long vowels with a mean duration 1.66 times longer than short vowels, and with a standard deviation of 0.5 for the long vowels, and 0.33 for the short vowels⁴. In this model, vowel length was implemented by varying the number of non-empty frames presented to the input layer, such that longer vowels had more ‘filled’ frames, and shorter vowels had more empty frames.

Model training proceeded similarly to the previous experiment, though in the present simulations, the training task was simply to identify the vowel presented on the input by activating the appropriate output node. There were 10 output nodes for each network, each corresponding to a separate vowel identity; in cases where similar vowel qualities were differentiated by contrastive length, separate output nodes were used for the long and short versions of the two vowels. Three networks of each type were trained for 250,000 training trials. To compensate for the relatively simple nature of this task, 15 hidden units were used in each simulation.

Results

Sum Squared Error means are plotted in Figure 5. These results indicate that networks trained on Inventory 3 were learning the training set slightly better than those trained on Inventory 2, and much better than those trained on Inventory 1. To further assess model performance, fully trained networks were tested on a set of 10 novel vowels of each type. The mean percent of correctly generalized items was 62.3% for Inventory 1, 86.7% for Inventory 2, and 91% for Inventory 3.

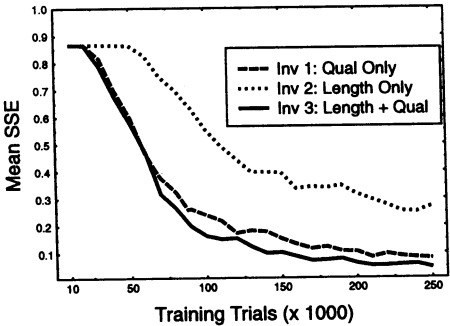


Figure 5: Mean Sum Squared Error rates for the three network types in this experiment. Lower values indicate better performance.

Summary

The present simulations support the hypothesis that contrastive vowel length represents a weaker cue than typical spectral differences for discriminating vowels, and affects languages' tendency to recruit contrasts which rely solely on length distinctions. This explains the paucity of languages with smaller inventories (3-8 vowel quality contrasts) that use vowel length distinctions. And while frequency data is missing for languages in which all vowel qualities participate in length contrasts, these results predict that the facts should not be different for such languages. Finally, these results are consistent with observations that length contrasts tend to accompany smaller quality contrasts (e.g. /i:/ -/I/).

4 Discussion

This work has explored the idea that the vowel inventory preferences of the world's languages result from their functional optimization. The simulations described here demonstrate the utility of connectionist networks in exploring this type of hypothesis. The performance of these models is easy to assess, and allows us to directly compare results to empirical facts about the world's languages. Here we have explored only a few of the possible uses of this approach. Additional applications include testing the contrastiveness of other vowel cues, such as diphthongization and nasality. This type of approach could also be useful in assessing the role of discriminability in consonant inventory frequencies. For example languages' preference of velar and alveolar stops over palatal stops might also be a function of these phonemes' discriminability.

The results of such simulations can serve to explain phonological patterns based on articulatory, acoustic and computational constraints. The network discovers constraints in the course of learning to perform a task. In contrast to other approaches such as Optimality Theory (Prince & Smolensky 1997), constraints do not have to be specified in advance; they emerge in the course of acquisition given the nature of the architecture, the characteristics of the input, and the task being performed.

Notes

¹ Address correspondence to Marc Joanisse, USC Neuroscience Program, University Park, Los Angeles CA 90089-2520 (email: marcj@gizmo.usc.edu).

² See chapter 8 of Maddieson (1984) for an overview of facts concerning vowel inventories.

³ See Elman, Bates, Johnson, Karmiloff-Smith, Parisi, & Plunkett (1996) for a discussion of how connectionist models work.

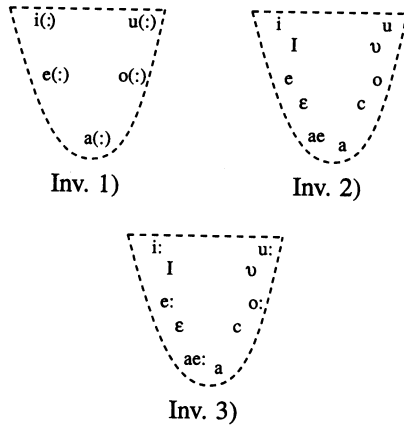
⁴Variabilities are estimates drawn from recordings of Finnish speakers, where lengths of vowels spoken in similar consonantal contexts – but differing prosodic contexts – were measured and compared. They represent a best guess at durational contrast and variability for a language with reliable length contrast, although languages will tend to vary along these parameters.

References

- Beckman, M. E., T.-P. Jung, S. Lee, K. de Jong, A. K. Krishnamurthy, S. C. Ahalt, & K. B. Cohen (1995). Variability in the production of quantal vowels revisited. *Journal of the Acoustical Society of America* 97, 471–90.
- Boë, L.-J., J.-L. Schwartz, & N. Valée (1994). The prediction of vowel systems: Perceptual contrast and stability. In E. Keller (Ed.), *Fundamentals of speech synthesis and speech recognition*, pp. 185–213. John Wiley & Sons.
- Chen, M. (1970). Vowel length as a function of the voicing of the consonant environment. *Phonetica* 22, 129–159.
- Chomsky, N. & M. Halle (1968). *The Sound Pattern of English*. MIT Press.
- Clements, G. (1985). The geometry of phonological features. *Phonology Yearbook* 2, 225–52.
- Davis, M. H., W. D. Marslen-Wilson, & M. G. Gaskell (1997). Ambiguity and competition in lexical segmentation. In *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society*, Volume 19, Mahwah, NJ, pp. 167–172. Laurence Erlbaum.
- Elman, J. L., E. A. Bates, M. H. Johnson, A. Karmiloff-Smith, D. Parisi, & K. Plunkett (1996). *Rethinking Innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.
- Jakobson, R. (1941). Kindersprache, aphasie und allgemeine lautgesetze. In *Selected writings I*, pp. 328–401. The Hague: Mouton.
- Joanisse, M. F. & M. S. Seidenberg (1997). [i e a u] and sometimes [o]: Perceptual and computational constraints on vowel inventories. In *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society*, Volume 19, Mahwah, NJ, pp. 331–336. Laurence Erlbaum.
- Liljencrants, J. & B. Lindblom (1972). Numerical simulation of vowel quality systems: The role of perceptual contrast. *Journal of Phonetics* 48(4), 839–862.
- Lindblom, B. (1986). Phonetic universals in vowel systems. In J. Ohala (Ed.), *Experimental Phonology*. Academic Press.
- Maddieson, I. (1984). *Patterns of Sounds*. Cambridge: Cambridge University Press.

- Magen, H. S. & S. E. Blumstein (1993). Effects of speaking rate on the vowel length distinction in Korean. *Journal of Phonetics* 21, 387–409.
- Prince, A. & P. Smolensky (1997). Optimality: From neural networks to Universal Grammar. *Science* 275, 1604–1610.
- Rumelhart, D., G. Hinton, & R. Williams (1986). Learning internal representations by error propagation. In D. Rumelhart and J. McClelland (Eds.), *Parallel distributed processing, vol. 1*. Cambridge, MA: MIT Press.
- Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics* 17, 3–45.

Appendix 1: Inventories used in Experiment 2



Perception in Optimality Theory: The Frugality of the Base

Paul Lassettre & Patricia Donegan
University of Hawai'i at Manoa

A complete theory of phonology must model perception as well as production. Children's accurate phonological perceptions and inaccurate phonological productions would seem to pose a problem for Optimality Theory (OT, Prince & Smolensky 1993), where we expect a single grammar for both, yet Smolensky (1996) gives an account of the asymmetry. In his model, structural well-formedness and faithfulness constraints evaluate surface representations, but only faithfulness constraints evaluate underlying representations. However, the OT account of perception seems to predict that adult phonological perception will be highly accurate – that adults will perceive the surface representation. Any adult who has ever studied an unfamiliar language knows that this is not true: distinctions that are perfectly obvious to speakers of one language can seem impossibly *subtle* to others. So, although there is much to recommend Smolensky's proposal, we believe it requires some modification. In our model of perception, not only the faithfulness constraints but also a certain class of structural well-formedness constraints evaluate underlying representations. The result is underlying representations that look very similar to traditional phonemic representations.

1. An OT model of perception.

Smolensky (1996) presents a model that accounts for the widely-noted discrepancy between linguistic production and perception in young children, who are able to hear (and remember) differences between sounds that they cannot pronounce distinctly. The model uses a single grammar for both production and perception, and it builds on previous work such as Prince & Smolensky (1993: 175-196) and Gnanadesikan (1995). It makes use of two representations, UR and SR. Production is modeled as mapping a given UR onto an unknown SR, and perception is modeled as mapping a given SR onto an unknown UR. The discrepancy between children's production and perception corresponds to a mapping that is more accurate from SR to UR than from UR to SR.

An OT grammar evaluates candidates for whichever representation is to be discovered. Two types of constraints are used: STRUCTURAL CONSTRAINTS (SCs) and FAITHFULNESS CONSTRAINTS (FCs). SCs evaluate the well-formedness of SRs. FCs evaluate URs and SRs with respect to each other; cases where UR and SR differ are less optimal than where UR and SR are alike. In the OT model, the evaluation function does not treat UR and SR equally: UR is evaluated only in terms of its faithfulness to SR, while SR is evaluated with respect to both its faithfulness to UR and its inherent well-formedness. In the early stages of acquisition, SCs are assumed to outrank FCs (Tesar & Smolensky 1996). The unequal

ranking of the two types of constraints and their inequality in evaluating the two representations account for the discrepancy between perception and production.

Tableau 1 shows the grammar of a child who pronounces *fish* as [fis] but rejects that pronunciation by adults; the same child pronounces *kiss* as [kis] and accepts that pronunciation by adults. This is our interpretation of Smolensky's model, using our own notation: candidates are shown as ordered pairs, in which the known member of the pair is fixed and the unknown varies freely (although we show only two variants). Faithfulness is UR-SR correspondence (McCarthy & Prince 1994, McCarthy 1995): MAX /feature/ is violated if the specified feature in UR lacks an identical correspondent in SR, and DEP [feature] is violated if the specified feature in SR lacks an identical correspondent in UR.

Perception: given [fiʃ], [kis]

	*C [+pal]	MAX /+pal/	DEP [+pal]		*C [+pal]	MAX /+pal/	DEP [+pal]
a.</fis/, [fiʃ]>	*		*!	☞c.</kis/, [kis]>			
☞b.</fiʃ/, [fiʃ]>	*			d.</kiʃ/, [kis]>		*!	

Production: given /fiʃ/, /kis/

	*C [+pal]	MAX /+pal/	DEP [+pal]		*C [+pal]	MAX /+pal/	DEP [+pal]
☞e.</fiʃ/, [fis]>		*		☞g.</kis/, [kis]>			
f.</fiʃ/, [fiʃ]>	*!			h.</kis/, [kiʃ]>	*!		*

Tableau 1. After Smolensky: accurate perception, inaccurate production. *C[+pal]: Consonants are non-palatal. MAX /+pal/: Palatality in UR corresponds to palatality in SR. DEP[+pal]: Palatality in SR corresponds to palatality in UR.

The production tableaux show cases where the URs /fiʃ/ and /kis/ are the given information. Candidates are ordered pairs of UR and SR in which the SR varies. The candidate set contains all possible SRs but just one UR. The highest ranked of the three constraints shown is a structural constraint, which evaluates the SR. The harmonic ordering of the candidate set is therefore affected by the well-formedness of the SR, and the winning candidates are [kis] and the unfaithful but well-formed [fis].

In the perception tableaux, the two SRs [fiʃ] and [kis] are given. As above, the candidates are ordered pairs, but now it is the SR that is constant. The candidate set includes all possible URs but just one SR. All SRs incur equal violations of the structural constraints, and therefore the well-formedness of the SR does not affect the harmonic ordering of the candidate set. FCs, although ranked below SCs, are the only constraints that affect the harmonic ordering and the winning candidate is the one where the UR and the SR are most alike.

The child's ability to hear the difference between [ʃ] and [s] is reflected in the model by the faithful mapping from [fiʃ] and [kis] to /fiʃ/ and /kis/. The same child's inability to pronounce /ʃ/ distinctly from /s/ is shown by the same grammar's mapping /fiʃ/ onto [fis].

2. Perception and richness of the base.

‘Richness of the base’ is the phrase by which Prince & Smolensky (1993: 191) refer to the basic assumption that the candidate set is unlimited. The implication is that systematic gaps and omissions in a speaker’s outputs cannot be explained merely by gaps and omissions in the input, but must be explained by the phonology itself (in OT, this is accomplished by the evaluation function). Gaps and omissions in a language’s lexicon reflect the surface patterns, not the other way around.¹ ‘Richness of the base’ is a reiteration of the claim that the ‘locus of explanatory action’ (Prince & Smolensky 1993:3) is, in OT, the ranked constraint hierarchy.

For example, Standard Hawaiian has no closed syllables. The explanation for this is not that Standard Hawaiian has no underlying strings that could be parsed into closed syllables. Rather, Standard Hawaiian phonology is such that every possible underlying string, including strings that contain such things as consonant clusters and final Cs, will be parsed into surface open syllables. The explanation lies in the constraint ranking, not in the inputs. Any lack of underlying consonant clusters and final consonants is a projection of the surface pattern.

As a further example, Korean has no surface contrast between [l] and [r]. The distribution of these liquids is predictable: [l] occurs syllable finally and [r] occurs elsewhere. The phonology of Korean disallows [l] and [r] in complementary surface structure environments. The lack of an underlying contrast between the two is due to their surface neutralization.

However, as Smolensky (1996) notes, the neutralization of Korean [l] and [r] is not just neutralization in production. Korean speakers also neutralize [l] and [r] in perception, tending to perceive English *bear* and *bail*, for example, as homophones.² The OT model with faithful mapping from SR to UR predicts that these sounds will be perceptually distinct for Korean speakers as [ʃ] and [s] are for the child acquiring English.

Smolensky proposes (in a footnote), that this effect is due to lexical retrieval: in addition to URs and SRs, there are underspecified LEXICAL REPRESENTATIONS (LRs). Features that are not used distinctively in a language are absent from the LR. In lexical retrieval, a UR is matched to an LR. Because of underspecification, some distinct URs may match the same LR. Failure to perceive a distinction is reflected in this model as matching a single LR. Smolensky suggests this is the case with [l] and [r] in Korean: Korean speakers hear *bear* and *bail* as homophones because Korean speakers cannot match them with distinct LRs (Tableau 2).

	Struct	Faith		Struct	Faith
a.</ber/, [ber]>	*		☞ c.</ber/, [bel]>	*	*!
☞ b.</bel/, [ber]>	*	*!	d.</bel/, [bel]>	*	

Lexical retrieval: /ber/ => lbelL, /bel/ => lbelL

Tableau 2. Korean: Faithful URs, homophony at lexical retrieval.

This model maintains the principle of 'richness of the base', but it does so indirectly. Perception of homophony (gaps and omissions in the perceptual pattern) is explained by gaps and omissions (underspecified features) in the LR. But gaps and omissions in the LR are ultimately explained in terms of the constraint hierarchy, by way of patterns in SRs.

3. Nasal vowels in American English.

A model of perception that depends on underspecified LRs does not account for all perceptual phenomena. The model described above predicts, for example, that the LRs of many American English dialects would have to include nasalized vowels. But speakers of such dialects do not perceive the large number of contrasts entailed by this analysis, or at least not more reliably than Korean speakers distinguish [r] and [l].

American English speakers ordinarily pronounce *say*, *sate*, *sane*, and *saint* as [seɪ], [seɪt], [sɛ̃n], and [sɛ̃t]. There does not exist any fifth lexical item ordinarily pronounced [sɛ̃], [seɪn], [seɪnt], or [sɛ̃nt] that is not homophonous with *say*, *sate*, *sane*, or *saint*. No such lexical item is possible. The phonology determines how nasalized vowels are distributed on the surface, by mapping the eight logically possible URs, /seɪ/, /seɪt/, /seɪn/, /seɪnt/, /sɛ̃/, /sɛ̃t/, /sɛ̃n/, and /sɛ̃nt/ onto the four attested SRs, [seɪ], [seɪt], [sɛ̃n], and [sɛ̃t].

Some SRs differ only by vowel nasalization, but are not neutralized in perception: *sate* [seɪt] and *saint* [sɛ̃t], *debt* [det] and *dent* [dɛ̃t], *lip* [lɪp] and *limp* [lɪp], *duck* [dʌk] and *dunk* [dʌ̃k], and so on. URs are faithful to SRs, and LRs are not distinct from URs. Vowel nasalization is specified in LR or it is underspecified; there is no other option. If vowel nasalization were underspecified in LR, *debt* and *dent* and similar pairs would be perceived as homophones (Tableau 3). Therefore, vowel nasalization must be specified in LR to account for these minimal pairs.

	*V N [-nas]	*NT	DEP [+nas]	*V _[+nas]	*MAX /C/	MAX /+nas/
a. </det/, [det]>						
b. </dɛ̃t/, [det]>						*!
c. </det/, [dɛ̃t]>			*!	*		
d. </dɛ̃t/, [dɛ̃t]>				*		
e. </dɛ̃nt/, [dɛ̃t]>				*	*!	
f. </dent/, [dɛ̃t]>				*	*!	

Lexical retrieval: /det/ => |dɛ̃_[Ønas]t| /dɛ̃t/ => |dɛ̃_[Ønas]t|

Tableau 3. English perception 1: [det], [dɛ̃t] as homophones, with accurate URs, underspecified LRs. *V[-nas]N: Non-nasal vowels do not occur before nasals. *NT: Nasal consonants do not occur before homorganic voiceless stops. DEP [+nas]: Nasality in SR must be attributable to nasality in UR. *V[+nas]: Nasalized vowels do not occur. MAX /C/: A consonant in UR corresponds to a consonant in SR. MAX /+nas/: Nasality in UR corresponds to nasality in SR.

But if this OT model were correct, American English speakers would reliably distinguish the eight SRs, [sej̥], [sej̥t̚], [sej̥n̚], [sej̥nt̚], [sej̥̃], [sej̥̃t̚], [sej̥̃n̚], and [sej̥̃nt̚], whether or not they can produce them. We do not believe American English speakers can distinguish all eight any more reliably than Korean speakers can distinguish pairs like *bear* and *bail*. No scheme of lexical underspecification will condition a realistic pattern of perceptual mergers and distinctions.

The standard OT model predicts that adult perception will be highly accurate, more accurate than adult production. Smolensky's model, with its underspecified LR, attempts to account for cases where this prediction is incorrect. But the underspecified LR model cannot account for more complicated examples of the same type, such as the American English perception of vowel nasalization, so it does not solve the problem.

4. Systematically unfaithful URs.

A modification to the way some structural constraints evaluate representations can resolve the nasalization dilemma. We recognize two classes of structural constraints, which can be distinguished formally and functionally. One class of constraints should evaluate URs as well as SRs, while the other should continue to evaluate only SRs.

One class we call SEGMENTAL WELL-FORMEDNESS (SEGWF) constraints, although we might equally well term them Fortitive or Enhancement or Paradigmatic constraints (cf. Stevens & Keyser 1989, Pulleyblank 1997). Formally, these constraints refer to the occurrence or simultaneous co-occurrence of phonological features. (1) shows our notation (cf. Lassette 1995).³

- (1) a. *[feature] b. * feature;
 feature_k

Empirical evidence for SEGWF constraints comes from context-free phonological phenomena, including crosslinguistic patterning of phoneme inventories, unconditioned historical sound changes, and stylistic variation observed in various kinds of hyperarticulated speech. SEGWF constraints have a functional basis in perception, though they may have articulatory motivation, too.

The other class we call SEQUENTIAL WELL-FORMEDNESS (SEQWF) constraints, although they might be called Lenitive constraints. Formally, these constraints refer to sequences of features or feature combinations (2). Empirical evidence for these constraints comes from context-sensitive phonological phenomena, such as assimilation and reduction, and 'conditioned' historical sound changes. SEQWF constraints, as their name implies, have a functional basis in the articulation of sequences, typically reducing the number or magnitude of differences between gestures that follow one another.⁴

- (2) *[feature_i] [feature_k]

SEGWF constraints evaluate URs, while SEQWF constraints evaluate SRs only. The result is that in perception URs are not fully faithful to given SRs, although the mappings from SR to UR and from UR to SR are still asymmetrical. Perception of contrast is modeled as mapping of distinct SRs onto distinct URs. Perception of homophony, however, is modeled as mapping of distinct SRs directly onto the same UR.

Tableau 4 shows how this resolves the English vowel nasalization problem. $*V_{[-nas]}N$ and $*NT$ are SEQWF constraints and do not evaluate in perception. The only SEGWF is $*V_{[+nas]}$, but its application in perception is sufficient to change the optimal UR for [dēt] from /dēt/ to /dent/, given the ranking $DEP_{[+nas]} > *V_{[+nas]} > MAX_{/C/}$. Thus [det] and [dēt] are associated with distinct URs, without allowing underlying / \tilde{v} / in English. This constraint ranking will lead to all surface occurrences of stressed [\tilde{v}] being perceived as /VN/ by English speakers.⁵

	$*V_{[-nas]}N$	$*NT$	$DEP_{[+nas]}$	$*V_{[+nas]}$	$*MAX_{/C/}$	$MAX_{/+nas/}$
a. </det/, [det]>						
b. </dēt/, [det]>						*!

c. </det/, [dēt]>			*!	*		
d. </dēt/, [dēt]>				**!		

Tableau 4. English perception 2: [det] as /det/, [dēt] as /dent/, with unfaithful URs.

In fact, not all surface occurrences of [\tilde{v}] have to be perceived as /VN/. Because of another sequential constraint, $*NV_{[-nas]}$, which dominates $*V_{[+nas]}$, unstressed syllables of the form VNT and VT are homophonous after a nasal consonant, as in *approximate* (adj.) and *approximant*, both of which are usually [ə'praksĩmĩt]. Tableau 5 suggests that the nasalized vowel here will be perceived, other things being equal (i.e., without certain morphological or orthographic knowledge – cf. Stampe 1987) as a nonnasalized vowel without a following nasal consonant. Its nasality can be attributed to the preceding nasal. $*V_{[+nas]}$, $Max_{/C/}$, and $Max_{/+nas/}$ are unviolated. (See Tableau 5.)

	$*V_{[-nas]}N$	$*NV_{[-nas]}$	$*NT$	$Dep_{[+nas]}$	$*V_{[+nas]}$	$Max_{/C/}$	$Max_{/+nas/}$
e. </mæt/, -[mĩt]>							
f. </mænt/, -[mĩt]>						*	

Tableau 5: English perception of unstressed -[mĩt].

Perception of the final syllable as including a postvocalic nasal (as /VNT/) requires that $Max_{/C/}$ be violated. This is of course possible, since $Max_{/C/}$ may be ranked below morphological or orthographic constraints. Spelling errors like *dominate* (adj.) for *dominant* and *pregnate* or *pregnit* for *pregnant* (but, we

predict, not *litigate* for *litigant*) give evidence that this (mis)perception does occur. Thus, the revised model is capable of handling production and perception of nasalized vowels in various environments in English simply.

The Korean case is resolved in the same way as the English one. *Bear* [ber] and *bail* [bel] are mapped directly onto the same UR (Tableau 6). As in Smolensky’s model, perception of homophony is mapping onto the same representation. Unlike that model, however, this model derives that mapping directly from an optimality- theoretic constraint hierarchy evaluating a candidate set. Korean phonology neutralizes liquids in both perception and production.

	Structure	SEQWF	SEGWF
a.</ber/, [ber]>	*	**!	
♣b.</bel/, [ber]>	*	*	
c.</ber/, [bel]>	*	*!	
♣d.</bel/, [bel]>	*		

Tableau 6. Korean perception: homophony at UR.

5. On children’s perceptions.

Returning to the question with which we started, that of children’s perceptions, we note that the revised model allows the possibility that children learning English may at first be unable to perceive the difference between [ʃ] and [s]. If structural constraints outrank faithfulness constraints initially, children will need to rerank constraints in order to perceive the distinction. Now, there is little if any evidence of systematic misperceptions (like perceptual merger of [ʃ] and [s]) in children. It seems, rather, that children are able to perceive virtually all of the phonetic distinctions used in languages from a very early age (Eimas et al. 1971, Trehub 1976, etc.). This ability changes somewhat during the babbling period, but at the onset of speech children still appear to be able to perceive all the phonemic distinctions of the ambient language (Werker & Pegg 1992, cf. Smith 1973.).

Yet it is not necessary to rule out the possibility of perceptual mergers in the early period of acquisition, especially since data on linguistic perception at this period is notoriously hard to come by. In Tableau 6, examples a. - d. show a possible (though unattested) stage of acquisition where the dominant SEGWF constraint *C_[+pal] determines perception. Examples e. - h. in the same tableau show a later stage, where the faithfulness constraint, DEF_[+pal], determines perception; note that at this point the faithfulness constraint MAX _{/+pal/} is still outranked by *C_[+pal], so *C_[+pal] determines the child’s productions.

The model presented here is a model of linguistic perception, but we assume that both representations, SR as well as UR, are percepts. The surface representation has already undergone significant cognitive processing, and it is not a direct representation of the acoustic wave. Rather, it is a representation of what is heard. It may be that the youngest children studied perceive all possible

phonetic distinctions because they treat linguistic stimuli simply as sounds, so that there would be at this point no difference between 'speech perception' and non-speech auditory perception. This would mean that the youngest infants respond simply to differences in SR.

Earlier Perception: given [fiʃ], [kɪs]

	*C [+pal]	MAX /+pal/	DEP [+pal]		*C [+pal]	MAX /+pal/	DEP [+pal]
a.</fis/, [fiʃ]>	*		*!	c.</kis/, [kɪs]>			
b.</fiʃ/, [fiʃ]>	**!			d.</kɪʃ/, [kɪs]>	*!	*	

Later Perception: given [fiʃ], [kɪs]

	DEP [+pal]	*C [+pal]	MAX /+pal/		DEP [+pal]	*C [+pal]	MAX /+pal/
e.</fis/, [fiʃ]>	*!	*		g.</kis/, [kɪs]>			
f.</fiʃ/, [fiʃ]>		**		h.</kɪʃ/, [kɪs]>		*!	*

Tableau 7. Earlier and later linguistic perception.

Only when the experience of using their own vocal tracts causes the constraints on their phonetic abilities to become part of their mental processing of speech would 'speech perception' in children begin to be different from non-speech auditory perception.⁶ When the child begins to perceive in terms of a smaller set of categories, limited by SEGWF constraints, which are presumed to characterize the speaker's intentions – then we can say that the child begins to perceive in terms of URs. This change does not involve a loss of auditory sensitivity or ability, of course: adults may be enabled to hear in terms of SRs under certain conditions (Werker & Tees 1984, Best et al. 1988). Under such conditions, Korean speakers may become able to judge that *bear* and *bail* do not sound exactly alike, even though they may remain unable to apprehend the difference linguistically.

6. Variability across languages.

Gaps, omissions, and other patterns in URs are not mere projections of surface patterns. Differences among languages with respect to URs, as well as SRs, can be attributed to different constraint rankings. For example, consider the constraints:

*VN Nonnasal vowels do not occur before nasal consonants (SEQWF).

*V_[+nas] Nasalized vowels do not occur (SEGWF).

MAX /nas/ Nasality in UR corresponds to nasality in SR.

DEP [nas] Nasality in SR corresponds to nasality in UR.

These constraints, ranked differently, account for vowel nasalization in different languages. In English: *VN >> *V_[+nas] >> DEP[nas]. As a result, nasalized vowels occur before nasal consonants in SRs, but the SEGWF constraint, *V_[+nas], ensures that vowels in UR are all nonnasal. In French, the faithfulness

constraints MAX/nas/ and DEP[nas] >> *VN and *V_[+nas]. With the faithfulness constraints most highly ranked, both nasalized and nonnasalized vowels occur in URs, and they appear as nasalized and nonnasalized vowels, respectively, in SRs. In Hindi, *VN >> DEP[nas] >> *V_[+nas], MAX /nas/. As a result of this ranking, both nasalized and nonnasalized vowels appear in URs, and some nonnasalized UR vowels are nasalized in SRs.

7. Conclusion.

The input to perception must be just as unlimited as the input to production: 'richness of the base' is bidirectional. Gaps and omissions in perception must be explained by the phonological system itself, as the result of the interaction of faithfulness and perceptually motivated structural constraints.

Notes.

¹ The phrase is new, but the concept is not (e.g., Andersen 1973, Stampe 1987). In a phonological theory which attempts to encompass phonological acquisition and change, things could hardly be otherwise.

² The Korean difficulty applies also to *bear* and *bell*, an alternative minimal pair.

³ They may be expressed as constraints on licensing or feature geometry.

⁴ Our definition of the simultaneous-sequential distinction is an approximation. We have not addressed prosodically defined positional limitations of occurrence, like *[-son +voi] in final position (German) or *[-cont +asp] in 'minor' (unstressed) syllables (Khmer). Nor have we addressed the question of diphthongization (of vowels or consonants). Diphthongization addresses simultaneity: it seems to avoid the simultaneity of a combination – not one of the given features. But diphthongization, being dissimilative, is not sequence-optimizing; it optimizes individual phonetic properties at the expense of the integrity of the segment.

⁵ Children's spontaneous spellings of words like *ant* and *dent* as AT, DAT (Read 1975: 54ff) might be thought to suggest that children at first interpret the nasalized-vowel-plus-voiceless stop [ṽt] as such, /ṽt/. Spellings like AD, AGRE for *and*, *angry* which are not pronounced with [ṽd], suggest that this is not the only possible explanation of such spellings. (See Read for possibilities.)

⁶ Whether or not this requires the kind of abstraction from phonetic abilities (viewed as construction of phonological constraints) implied in Hayes' (1996) proposal of 'inductive grounding' is not at issue here (cf. Donegan 1995, where constraints are more directly phonetic).

References

- Andersen, Henning. 1973. Abductive and deductive change. *Language* 49:765-793.
- Best, C.T., G.W. McRoberts, & N. M. Sithole. 1988. The phonological basis of perceptual loss for non-native contrasts: Maintenance of discrimination among Zulu clicks by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14:345-60.
- Donegan, Patricia. 1995. The innateness of phonemic perception. WECOL 7, ed. by Vida Samiian, 59-69. Fresno, CA: Western Conference on Linguistics.
- Eimas, Peter D., E.R. Siqueland, P. Jusczyk, & J. Vigorito. 1971. Speech perception in infants. *Science* 171:303-306.
- Gnanadesikan, Amalia. 1995. Markedness and faithfulness constraints in child phonology. ROA. University of Massachusetts, Amherst.
- Hayes, Bruce. 1996. Phonetically driven phonology: The role of Optimality Theory and inductive grounding. ROA. UCLA.
- Lassette, Paul. 1995. Process effects in non-processual phonology. University of Hawai'i Working Papers in Linguistics 27:37-54.
- McCarthy, John. 1995. Extensions of faithfulness: Rotuman revisited. ROA. University of Massachusetts, Amherst.
- McCarthy, John, & Alan Prince. 1994. Faithfulness and reduplicative identity. ROA. University of Massachusetts, Amherst, and Rutgers University.
- Prince, Alan, & Paul Smolensky. 1993. Optimality Theory: Constraint interaction in generative grammar. ms. Rutgers University & University of Colorado at Boulder.
- Pulleyblank, Douglas. 1997. Optimality Theory and features. *Optimality Theory: An Overview*, ed. by D. Archangeli & D. T. Langendoen, 59-101. Malden MA: Blackwell.
- Read, Charles. 1975. Children's Categorization of Speech Sounds in English. Urbana: National Council of Teachers of English.
- Smith, Neilson V. 1973. *The Acquisition of Phonology*. Cambridge: Cambridge University Press.
- Smolensky, Paul. 1996. On the comprehension/production dilemma in child language. *Linguistic Inquiry* 27:720-731.
- Stampe, David. 1987. On phonological representations. *Phonologica* 1984, ed. by W. Dressler, H.C. Luschutzky, O.E. Pfeiffer, and J.R. Rennison, 287-300. London: Cambridge University Press.
- Stevens, Kenneth, and Samuel J. Keyser. 1989. Primary features and their enhancement in consonants. *Language* 65:81-106.
- Tesar, Bruce, and Paul Smolensky. 1996. Learnability in Optimality Theory (short version). ROA. Johns Hopkins University.
- Trehub, S.E. 1976. The discrimination of foreign speech contrasts by infants and adults. *Child Development* 47:466-472.
- Werker, J.F. & J.E. Pegg. 1992. Infant speech perception and phonological acquisition. *Phonological Development: Models, Research, Implications*, ed. by C.A. Ferguson, L. Menn, & C. Stoel-Gammon, 285-311. Timonium, MD: York Press.
- Werker, Janet F. & R.C. Tees. 1984. Phonetic and phonemic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America* 75:1866-1878.

Stress effects on CV coarticulation and their implications for phonology

Joo-Kyeong Lee
University of Illinois at Urbana-Champaign

1. Introduction

In this paper, I report on an experiment conducted to investigate the degree of coarticulatory effects of word-level stress on CV structures. The goal of this experiment is to determine if there is a difference in the degree of coarticulation between stressed and unstressed syllables at the word level.

Many phonetic studies have examined the acoustic and articulatory differences between stressed and unstressed vowels. Acoustically, a stressed vowel is typically identified by a higher fundamental frequency, greater intensity, greater duration and a different formant structure than an unstressed vowel (Brown Jr. & McGlone 1974, Engstrand 1988, Kent & Netsell 1971, etc.). Articulatorily, a stressed vowel involves a relative lowering of the jaw for low vowels, a greater opening of the lips and a movement of the tongue body to a more precise target position (De Jong et. al. 1993, De Jong 1995, Kent & Netsell 1971, Summers 1987, etc.). This paper investigates how such different manifestations of a stress contrast influence CV coarticulation by comparing the acoustic correlates of gestural overlap within stressed versus unstressed syllables. Specifically, this work explores how tongue body gestures of a vowel influence the constriction location in the vocal tract during the production of stop consonants by examining F2 variability across varying vowel contexts.

Most previous phonetic studies employed sentence stress, i.e. accent, and compared an accented syllable with an unaccented syllable (Engstrand 1988, De Jong et. al. 1993, De Jong 1995, Kent & Netsell 1971, and Summers 1987). As discussed in De Jong et al. (1993) and De Jong (1995), accented vowels, which have sentential or emphatic stress, are hyperarticulated: bigger jaw displacement for low vowel /a/ and more tongue retraction for back vowel /u/. They interpret the hyperarticulation as an enhancement of the phonemic distinctiveness of a segment in an accented syllable by reducing intergestural overlap. Therefore, they argue that accented syllables involve less gestural overlap, hence less coarticulation.

It has not been considered how lexical-level stress properties influence CV coarticulation, but on analogy with phrasal stress or accent, we might expect to find greater coarticulation between a consonant and a vowel in unstressed syllables within the word. If the absence of lexical stress conditions coarticulation, we might expect such coarticulation to develop into categorical place assimilation, and so expect an increased level of assimilation in stressless contexts. However, a review of the phonological literature on CV place assimilation does not support this hypothesis; there are very few examples of CV place assimilation only in stressless contexts.¹

2. Assumptions and Hypotheses

I examine gestural coarticulation of a consonant affected by vowel gestures under the assumptions proposed by Recasens (1991). He claims that vowels are less vulnerable to coarticulation because they are produced by means of global vocal tract shapes which require articulatory control upon the entire tongue body configuration. On the other hand, consonants involve only local constrictions which leave other articulatory regions free to coarticulate. Therefore, we would

expect a greater effect of a vowel on an adjacent consonant than that of a consonant on a vowel.²

A stop consonant has reliable place cues, the locus and F2 transitions into a flanking vowel. However, when a stop is articulatorily influenced by a flanking vowel, the locus is variant across the vowel contexts, as discussed by Kewley-Port (1982). Therefore, I hypothesize that the greater coarticulation in unstressed syllables will be manifested in a greater variance in F2 values in the vocalic regions of the syllables. Figure 1 displays the hypothetical F2 variation ranges of unstressed and stressed syllables across varying vowel contexts. More specifically, the range will be defined by the F2 of vowels in unstressed syllables: a greater range may result from various F2 values of the vowels as a consequence of a vowel's considerable coarticulatory effects. However, the range will be centered on a stop consonant locus in stressed syllables; the range may be an implementation of the approximate value of a stop locus as a result of less gestural overlap with adjacent vowels.

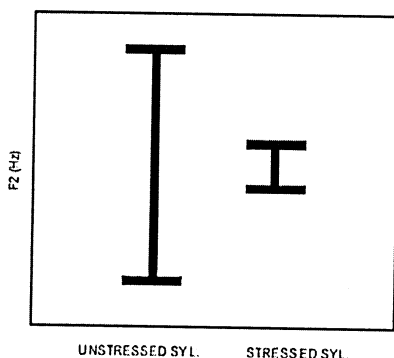


Figure 1. Hypothetical F2 variation range in unstressed and stressed syllables.

As discussed by Lindblom (1963) and Engstrand (1988), the transition duration is shorter in fast speech. It has been also claimed that a shorter transition results from a greater degree of overlap between a stop and a vowel (Browman & Goldstein 1989 and De Jong et al. 1993). Therefore, I hypothesize that unstressed syllables will involve a shorter duration of transition EITHER due to a greater degree of gestural overlap between a vowel and a consonant OR due to the intrinsically shorter duration of segments in unstressed syllables.

3. Experiment

3.1 Data

Nonsense Spanish words, /CVpo/ with /b, d, g/ and stressed and unstressed vowels, /i, e, a, o, u/ were recorded 6 times by three Spanish native speakers. Spanish was chosen because unstressed vowels are not categorically reduced in Spanish, allowing for a meaningful comparison of F2 values in stressed and unstressed syllables. English unstressed vowels are, by contrast, fully reduced in many contexts, it would not be appropriate to compare consonant and vowel coarticulation under stress contrasts. Only voiced stops were employed in the

experiment since the voiced stops show more distinct formant transitions than voiceless stops (Kewley-Port 1982). The /CVpo/ tokens were recorded in 8-bit and 8KHz onto a Sparc station with the Sony F-VX30 microphone. The stimuli were read in a carrier sentence as follows:

A: ¿escriben _____ para mi? 'Do they write _____ for me?'
 B: No, DICEN _____ para ti. 'No, they say _____ for you.'

The underlined part in B is the only source of data. Since a nuclear accent is assigned to an object noun in B, which is a target word in the sentence, the nuclear accent is intentionally put on DICEN so that the following object noun may not be affected by the phrasal stress. Pitch tracks show a flat low f0 in the target noun, while the preceding nuclear-accented DICEN displays HL.

3.2 Procedure

First, F2 values were measured at three points using Entropic's Waves program: at consonant release (hereafter C-release), at the F2 onset and at the steady state of a vowel. All measurements were carried out from spectra, compared with the corresponding waveform and spectrogram. A time window was used for sampling the spectrum to 6.5 ms. Since no noise-like jagged (burst) portion is seen for bilabial stops before the periodic waveforms of a vowel, the F2 values could not be measured at consonant release in /bVpo/ tokens. F2 variability was compared between stressed and unstressed syllables at two points, at C-release as well as at the onset of F2, although C-release alone can display sufficient information about vowels' coarticulatory effects on the preceding stops, since the C-release is farther from a vowel in timing. However, due to the fact that bilabial stops do not involve the C-release portion in Spanish, F2 variability solely at C-release would not show a general tendency for the three stop consonants associated with vowel's influences.

Second, since the stressed and unstressed vowels are qualitatively different - but not categorically different as in English - and exhibit slightly different F2 values, e.g. between [e] and [é], I compared the F2 values at two points, at consonant release and at the beginning of F2 transition with the F2 values of the vowel steady state. This measurement is to identify the extent to which vowels influence consonant place of articulation in association with stress.

Third, I calculated the transition duration of F2: the time duration from the F2 onset to the initiation point of the steady vowel. However, I factored out syllables involving the alveolar stop with /o, u/ vowels and the syllables involving the velar stop with /a, o, u/ vowels, because those syllables do not show a distinct plateau of F2 regions due to the rapid tongue movement from high F2 to the low F2 of the following bilabial stop.

3.3 Results

Figure 2 shows a comparison of F2 onset range between stressed and unstressed syllables for speaker 1. Since I acquired exactly the same results for speakers 2 and 3, I present the results of only speaker 1.

The three panels in Figure 2 exhibit the F2 onset range involved in the syllables with bilabial, alveolar and velar stop consonants, respectively. The range of F2 onset values is slightly, but not significantly, more variant in unstressed than in stressed syllables with a bilabial and alveolar stops as obtained from Levene's test for equality of variance ($p = 0.05$). For syllables containing a velar stop, this does not seem to be the case: the F2 range seems to be more variant even though the

variance is not significantly different. Moreover, F values in the equality of variance test are all below 1.

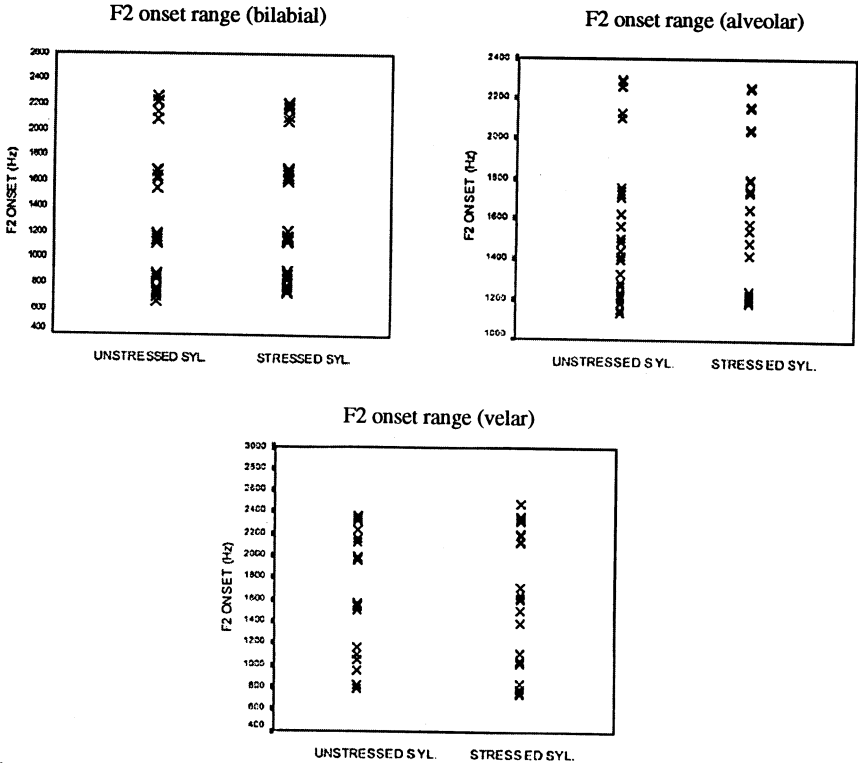


Figure 2 The range of F2 onset; the variability of F2 onset values is not statistically significant in stressed and unstressed syllables ($p = 0.05$) and all F values are below 1.

Figure 3 displays the F2 range at the alveolar and velar stop release in stressed and unstressed syllables. Again, there is no distinctively visible burst release in the bilabial voiced stop of /bVpo/ tokens in Spanish, and the F2 range at the bilabial consonant release cannot be shown in Figure 2. The left panel shows a comparison of the F2 in syllables with an alveolar stop, in which the F2 values are more variant in unstressed syllables than in stressed syllables. For the figure at the right, which represents the F2 range in syllables with a velar stop, this does not seem to be the case. However, the F2 range of unstressed syllables might be more variable if we exclude the highest value of F2 at around 2700 Hz in a stressed syllable, which could be a wrong measurement or an error. According to Levene's test, the variance of F2 values at stop release is not statistically different for unstressed and stressed syllables with both alveolar and velar stops ($p = 0.05$).

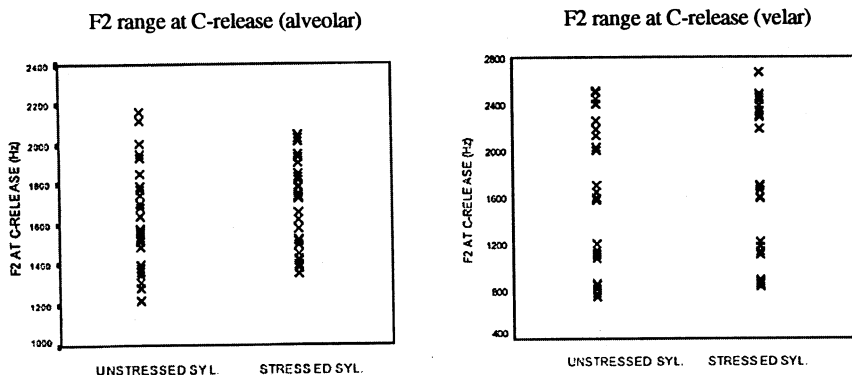


Figure 3 The range of F2 at consonant release; no significant difference in variability of F2 values is found between stressed and unstressed syllables ($p = 0.05$) and all F values are below 1.

Comparable results were obtained from the other two speakers in that the variation range of F2 at the consonant release and at the onset of transition is not statistically different between stressed and unstressed syllables ($p = 0.05$): the F values are all below 1.

The results indicate that the F2 values at the consonant release and at the onset of transition vary across the five vowel contexts regardless of stress. Therefore, the hypothesis that the range will be more variant in unstressed syllables than in stressed syllables is not supported. In other words, this test shows that there is no appreciable difference in stress effects on coarticulation between a vowel and a stop consonant in Spanish. It seems that stress information does not play a significant role in the degree of coarticulation between a vowel and a preceding stop consonant. Therefore, a possible interpretation could be made regarding temporal coordination of CV structures such that the overlap phasing between consonant and vowel gestures is not affected by word level stress.

Figures 4 and 5 show the correlation of a vowel's F2 values with F2 onset and with F2 at the stop release for speaker 1. Figure 4 exhibits the correlation between the onset of F2 and the F2 at the steady state of vowels in stressed and unstressed syllables. Correlation coefficients are 0.958 for unstressed syllables and 0.937 for stressed syllables ($p < 0.0005$), and the correlation both in stressed and unstressed syllables is statistically significant ($p = 0.01$). Therefore, F2 values between a vowel and the onset of transition are strongly correlated both in stressed and unstressed syllables.

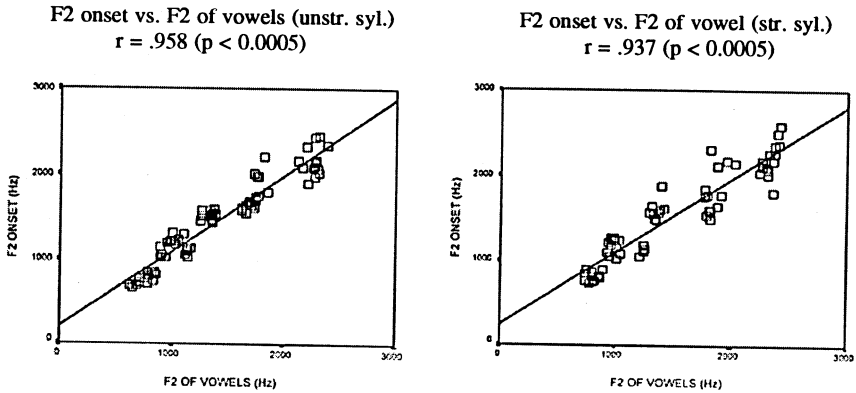


Figure 4 Correlation between F2 onset and F2 at the steady state of vowels in stressed and unstressed syllables. Correlation coefficients (r) are given ($p < 0.0005$), and the correlation both in stressed and unstressed syllables is statistically significant at the 0.01 level (2-tailed).

Figure 5 illustrates the correlation between F2 values at stop consonant release and F2 at the steady state of vowels. Although the correlation coefficient is higher in unstressed syllables, which is 0.924, than in stressed syllables, which is 0.866, both correlation values are statistically significant ($p = 0.01$). Consequently, this result also indicates a strong correlation between F2 at the steady state of a vowel and the consonant release.

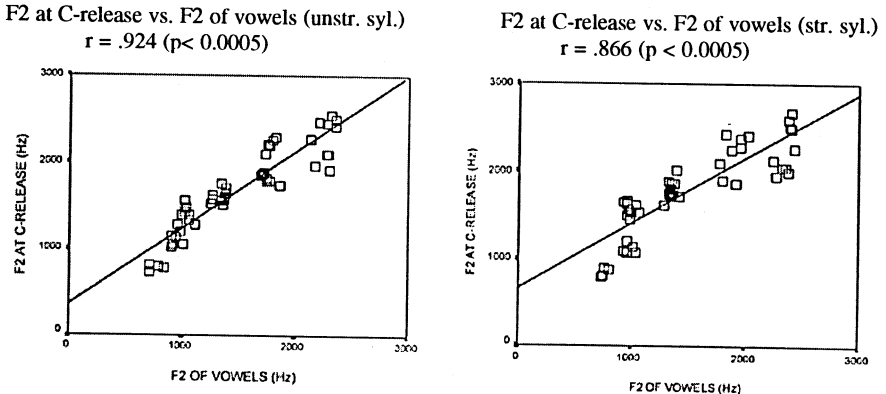


Figure 5 Correlation between F2 at stop release and F2 at the steady state of vowels in stressed and unstressed syllables. Correlation coefficients (r) are given ($p < 0.0005$), and the correlation both in stressed and unstressed syllables is statistically significant at the 0.01 level (2-tailed).

The other two speakers show the same correlation result as in Table 1 and all the correlation values are significant ($p = 0.01$). Therefore, a strong correlation

between a vowel and onset of F2 and between a vowel and consonant release is supported.

	F2 onset vs. F2 of vowels ($p < 0.0005$)	
	unstressed syllables	stressed syllables
speaker 2	$r = 0.899$	$r = 0.874$
speaker 3	$r = 0.904$	$r = 0.866$
	F2 at C-release vs. F2 of vowels ($p < 0.0005$)	
	unstressed syllables	stressed syllables
speaker 2	$r = 0.796$	$r = 0.786$
speaker 3	$r = 0.835$	$r = 0.794$

Table 1 Correlation (speakers 2 & 3)

The correlation results indicate that there is no considerable difference in the degree of a vowel's influence on a preceding consonant's articulation for stressed and unstressed syllables. In other words, the results do not indicate that a stressed vowel is less likely to coarticulate with a consonant than an unstressed vowel due to its hyperarticulation or its achieving a more accurate target position in Spanish. Since F2 values at the consonant release and at the onset of transition are strongly correlated with vowels' F2 values both in stressed and unstressed syllables, the difference in the degree of coarticulation associated with lexical stress is not found in the present work. Therefore, word-level stress would not be a factor affecting gestural overlap between consonant and vowel gestures.

Mean values of F2 transition duration are calculated between the F2 onset and the initiation point of the steady state of vowels (excluding /a, o, u/ vowels in the syllables with a velar stop and /o, a/ in the syllables with an alveolar stop as mentioned in 3.2). Figure 6 indicates this transition duration of speaker 1 in unstressed and stressed syllables. The mean duration of stressed syllables is significantly longer than that of unstressed syllables ($p < 0.0005$). The other two speakers show the same result as speaker 1 in that the mean duration of a transition is significantly longer in stressed syllables than in unstressed syllables ($p < 0.0005$). This may be either (1) because more overlap between two gestures manifests short-timing transition, given the inter-target transition indicates the distance of two gestural targets or (2) because the short duration of an unstressed vowel brings about short transition. Since there is no distinct difference of CV gestural overlap invoked by lexical stress, the hypothesis (2) is supported. In other words, the short transitional duration in an unstressed syllable is not the manifestation of greater gestural overlap, but that of the intrinsically shorter duration of a segment in an unstressed context.

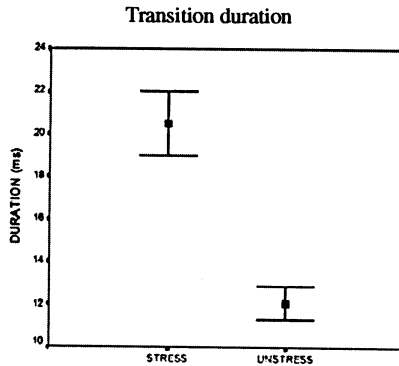


Figure 6 Mean duration of the F2 transition in stressed and unstressed syllables. Stressed syllables show 20.70ms, and unstressed syllables, 12.05ms. The two values are significantly different ($p < 0.0005$). The error bars represent standard error of means.

4. Discussion

I have presented some acoustic correlates of gestural overlap between a stop and a vowel associated with a stress contrast at the word level. I examined Spanish which does not involve a categorical reduction of an unstressed vowel and investigated the range of F2 variance at the consonant release and at the onset of F2 transition. The results show that F2 values are very slightly more variable, but not significantly, for unstressed syllables than for stressed syllables. The F2 range of a velar stop context is much greater than that for other stop contexts, but there is still almost no difference of the F2 range between stressed and unstressed syllables. In other words, word-level stress of a vowel does not seem to have much influence on the articulation of a preceding consonant, which implies that coarticulation between a vowel and a consonant is not significantly sensitive to word-level stress. If we take Browman & Goldstein's claim (1989) that coarticulation results from gestural overlap, the present work can be interpreted as saying that there is no considerable difference in the temporal phasing of consonant and vowel gestures in CV structures for stressed and unstressed syllables. Therefore, the results of this experiment do not support the hypothesis of greater gestural overlap in syllables that do not receive word-level stress.

The correlation results illustrate that consonant place cues are strongly correlated with a vowel's articulation, and also this significant relationship exists regardless of lexical stress. This finding also implies that vowel gestures extend to consonant articulation, which arises from gestural overlap, but the overlap in CV structures is found to be independent of stress.

The transition duration has been shown to be much longer in stressed syllables than in unstressed syllables. The shorter duration would simply result from an inherent characteristics of unstressed vowels. In other words, since the total duration of unstressed vowel is shorter, its transition should be shorter even if the same overlap phase is executed as in stressed syllables. This also implies that transition duration does not necessarily reflect on the degree of gestural overlap.

Previous studies explore coarticulation in accented and unaccented syllables with a phrase-level stress contrast, and find that there is greater gestural overlap in unaccented syllables. On the other hand, the present study investigates a

comparison of CV coarticulation between stressed and unstressed syllables with a word-level stress contrast, and show the result that there is no significant difference of gestural overlap. These findings seem to be consistent with Netsell's claim: "the three peripheral systems (pulmonary, laryngeal and supraglottal) may or may not all participate in stress production, depending on the degree of the stress contrast" (cited in Kent & Netsell 1971: 42).

I suggest a general principle governing the stress degree system such that the low-level stress contrast may not be perceptually as great as the high-level stress contrast, and it might be due to the fact that low-level stress is to magnify the perceptual function of stressed syllables only through a single word.³ In other words, perceptually less contrastive positions are more likely to be produced in favor of articulatory economy rather than perceptual salience, which may result in coarticulation both in stressed and unstressed syllables.

Accented syllables, on the other hand, should enhance great perceptual prominence so as to attract a listener's attention throughout the whole sentence. The high-level stress contrast might be carried out by a maximal differentiation in pulmonic pressure, articulatory/muscle strength, etc. In other words, perceptually salient accented syllables may tolerate articulatory cost, for example, by virtue of hyperarticulation: thus, coarticulation does not occur. Perceptually less salient unaccented syllables, however, may take advantage of articulatory economy; thus, coarticulation takes place. This might be only true in languages like Spanish where a low-level stress contrast is not qualitatively distinct along the segmental dimension; low-level stress shows a different contribution to gestural overlap from high-level stress.

5. Phonological implications and conclusion

If phonetic details are effectively reflected in phonological patterns (Ohala 1993), my results supplement the implication for phonology that consonant place assimilation affected by an adjacent vowel such as palatalization, labialization, and velarization should not be conditioned by lexical stress. My finding is consistent with the observation that the absence of stress is not a conditioning factor for phonological place assimilation as exemplified in languages like Akan (Schachter 1969, Boadi 1988), Czech (Kucera 1973), Eastern Ojibwa (Bloomfield 1957), Hungarian (Siptar 1994), Kashmiri (Kelkar & Trisal 1964), Konkani (Misra 1989), Nupe (Smith 1967), Polish (Rubach 1981), Russian (Jones & Ward 1969), Sarcee (Cook 1978), Slovak (Rubach 1993), Swazi (Ponelis 1974), etc. However, to the best of my knowledge, there are very few languages showing examples of CV place assimilation in a stressless environment.

Now, I argue that the slight, not significant, difference in coarticulation between stressed and unstressed syllables may not give rise to a perceptually detectable difference which could develop into CV place assimilation only conditioned by stressless contexts, especially when a language does not show contrastive values between a stressed and an unstressed vowel such as a full reduction of the unstressed vowel. It might be possible that if a language involves a qualitative distinction between a stressed and an unstressed vowel as in English where unstressed vowels are fully reduced to schwa, the language may tend to maximize the contrast; perceptually salient stressed syllables are governed by perceptual salience maximization while less salient unstressed syllables are produced by satisfying articulatory economy requirements and, the CV sequences are modified by coarticulation. In fact, English is a language where palatalization is conditioned by lexical stress.⁴

Notes

* I would like to thank Jennifer Cole, Molly Homer, Jose Hualde and Khalil Iskarous for their helpful suggestions. I am especially grateful to Daniel Silverman for his valuable comments and discussion. All remaining errors are my responsibility, of course.

¹ CV place assimilation denotes such processes as palatalization, labialization and velarization of a consonant conditioned by an adjacent vowel.

² I may have to exclude a gestural coordination between a consonant involving tongue body gesture as a primary articulator (e.g. a pharyngealized consonant) and a vowel since both the consonant and the vowel are configured by a global vocal tract and might be competitively coarticulating each other.

³ According to the grid representation of intonation (Lieberman & Pierrehumbert, 1984), lexical stress attracts two grids, and a nuclear accent, which is assigned to emphatic stress word or sentence stress, is represented with four grids. In this sense, I term an accent "high-level stress" and lexical stress "low-level stress."

⁴ For example, 'cónstitute' [tu] – 'constituent' [tSu] (See Borowsky 1986). I am not sure, however, that the languages which I listed above all do not show a categorical reduction of unstressed syllables.

References

- Bloomfield, Leonard. 1957. Eastern Ojibwa: grammatical sketch, texts and word list. Ann Arbor, MI: University of Michigan Press.
- Boadi, Lawrence. 1988. Problems of palatalization in Akan. *Journal of West African Languages* 18. 3-16.
- Borowsky, Toni J. 1986. Topics in the lexical phonology of English. Amherst, MA: University of Massachusetts dissertation.
- Browman, Catherine and Louis Goldstein. 1989. Tiers in articulatory phonology with some implications for casual speech, *Laboratory Phonology I: between the Grammar and Physics of Speech*, ed. by John Kingston and Mary Beckman, 341-76. Cambridge: Cambridge University Press.
- Brown Jr., W. S. and Robert McGlone 1974. Aerodynamic and acoustic study of stress in sentence production. *Journal of the Acoustical Society of America* 56: 971-74.
- Cook, Eung-Do 1978. Palatalization and related rules in Sarcee. *Linguistic Studies of Native Canada*, ed. by Eung-Do Cook and Jonathan Kaye, 19-35. Vancouver: University of British Columbia.
- De Jong, Kenneth. 1995. The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America* 97: 491-504.
- De Jong, Kenneth, Mary Beckman, and Jan Edwards. 1993. The interplay between prosodic structure and coarticulation. *Language and Speech* 36: 197-212.
- Engstrand, Olle. 1988. Articulatory correlates of stress and speaking rate in Swedish CV utterances. *Journal of the Acoustical Society of America* 83: 1863-75.
- Jones, Daniel and Dennis Ward. 1969. *The phonetics of Russian*. Cambridge: Cambridge University Press.
- Kelkar, Ashok and Pran Trisal. 1964. Kashmiri word phonology: a first sketch. *Anthropological Linguistics* 16: 13-22.

- Kent, Ray and Ronald Netsell. 1971. Effects of stress contrasts on certain articulatory parameters. *Phonetica* 24: 23-44.
- Kewley-Port, Diane. 1982. Measurements of formant transition in naturally produced stop consonant-vowel syllables. *Journal of the Acoustical Society of America* 72: 379-89.
- Kucera, Henry. 1973. Language variability, rule interdependency and the grammar of Czech. *Linguistic Inquiry* 6: 499-521.
- Lee, Joo-Kyeong. 1997. The asymmetry of C/V coarticulation in CV and VC structures and its implications in phonology. *Studies in Linguistics Sciences* 27: 139-52.
- Liberman, Mark and Janet Pierrehumbert. 1984. Intonational invariance under changes in pitch range and length. *Language and Sound Structure*, ed. by Mark Aronoff & Richard Oehrle: 157-233. Cambridge, MA: MIT Press.
- Lindblom, Björn. 1963. Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America* 35: 1773-81.
- Misra, Dipti. 1989 Palatalization in Konkani. *International Journal of Dravidian Linguistics* 18: 105-10.
- Ohala, John. 1993. Coarticulation and Phonology. *Language and Speech* 36: 155-70.
- Ponelis, Fritz. 1974. On the dynamics of velarization and labialization: some Bantu evidence. *Studies in African Linguistics* 5: 27-58.
- Recasens, Daniel. 1991. An electropalatographic and acoustic study of consonant-to-vowel coarticulation. *Journal of Phonetics* 19: 177-92.
- Rubach, Jerzy. 1981 *Cyclic Phonology and Palatalization in Polish and English*. Warszawa: Wydawnictwa uniwersytetu warszawskiego.
- Rubach, Jerzy. 1993 *The Lexical Phonology of Slovak*. Oxford: Oxford University Press.
- Schachter, Paul. 1969. Natural assimilation rules in Akan. *International Journal of American Linguistics* 35: 342-55.
- Smith, Neilson Voyne. 1967. The Phonology of Nupe. *Journal of African Languages* 6: 153-69.
- Siptar, Peter. 1994. Palatalization rules in Hungarian. *Acta Linguistica Hungarica* 42. 5-32.
- Summers, W. Van. 1987. Effects of stress and final-consonant voicing on vowel production: Articulatory and acoustic analyses. *Journal of the Acoustical Society of America* 82. 847-63.

Why make life hard? Resolutions to problems of rare and difficult sound types

Ian Maddieson
University of California, Los Angeles

Introduction

Ideas inspired by the models of evolutionary biology have provided valuable insights into the shaping of the phonetic systems of human languages. One fruitful perspective is to view language as a field within which a variety of phonetic entities — the possible sounds that humans can make — are in competition for survival according to how well-adapted they are to linguistic function. The usual terms of this discussion are to view sounds, or systems of sounds, as being subject to twin pressures for articulatory economy and auditory distinctiveness. The 'fittest' are those which achieve a successful balance between these demands. Obviously the biological analogy is imperfect — there is nothing equivalent in sound systems to the genetic code — but nonetheless the notion that phonetic elements are in an ecological competition for survival is fruitful, and goes a long way toward providing an account of such things as why some sounds are more common and others are more rarely found in the sound inventories of the world's languages (see, for example, Lindblom 1980, Lindblom and Maddieson 1988).

One challenge which this approach presents, however, is to account for the stable occurrence — as part of the established system of a language, not as occasional variants, or transient 'experiments' during acquisition — of sounds that seem to fail to satisfy the ecological demands. In the biological world mal-adapted life forms die out quickly. How come phonetic elements which appear to be mal-adapted are able to have long life? This is the problem that will be the focus of this paper. It will start with an attempt to sketch a more satisfactory picture of the ecological demands that the phonetic patterns of language must satisfy. It will then discuss three classes of sounds found in traditional typologies that appear to be ecologically unfit for survival in light of these demands. An examination of three language-specific instances of sounds that putatively fall into these classes shows that in each of these cases the ecological damage is mitigated by arrangements of details of articulatory positions or timing.

Contrastivity and Connectedness

I have suggested elsewhere (Maddieson 1997) that the two fundamental 'ecological' conditions the phonetic patterns of language must satisfy are ones that may be conveniently subsumed under the labels Contrastivity and Connectedness.

Contrastivity is above all the requirement that a language must show differentiation in sound, rather than being an undifferentiated noise. Without differentiation, essentially only the message "I'm over here" can be transmitted and received (roughly the meaning of certain monotonous insect calls, or frog croaks). A language cannot be built unless messages can begin in different ways and continue in different ways, making possible the construction of a lexicon and the

other elements required. But mere difference is not enough. Both speakers and listeners need to be able to reliably identify the same message as being the same, and different messages as being different. This favors the selection of sound elements with stable characteristic 'signatures' in both their motor and auditory patterns so that they may be recognized and memorized. Note the stress placed here on the importance of distinctiveness among the motor patterns, not just on auditory distinctiveness. Production patterns must be as distinctive as auditory/perceptual patterns in order for a speaker to be able to encode the differences between one word (morpheme, etc) and another.

These requirements create the property that is usually labeled phonological contrast: utterances must contain parts that are differentiated from each other and recognizable when they recur. Contrastivity is also behind the alternation of louder and quieter sounds that is the basis of syllabification.

Equally, a language needs to be produced as a continuous stream, its parts connected to each other just as essentially as they must be differentiated from each other. This property is given the label Connectedness. Note that one of the implications of a connectedness requirement, since the position in which a given word or other element will occur is variable, is that the form of any item must be adapted for variable environments. The suggestion here is that it is the need to achieve connectedness — rather than an explicit minimization of articulatory effort — that favors limits on the range of articulatory gestures, especially in adjacent parts of an utterance. Making an extreme gesture, for example sticking the tongue out as far as possible between the teeth, makes it harder to connect to the next position (unless it happens to be the same, and variability of context ensures that it will more frequently not be the same). Connectedness therefore favors articulatory displacements close to the mean position rather than extreme ones, as well as moderate efforts to raise subglottal pressure, limited modulation of fundamental frequency, etc.

Variation in the articulatory instantiation, i.e. coarticulation, of a given item (word, etc.) facilitates its concatenation with other items in contexts that themselves vary. In the perceptual domain, coarticulatory variation spreads information over longer durations in the signal, and hence helps to reduce the likelihood of transmission error. Sound elements that resist coarticulation are therefore disfavored. However, beyond a certain degree, variation begins to conflict with the need to be able to identify repetitions of the same item, coming into conflict with the requirement of contrastivity. Variation will therefore tend to remain within limited bounds. (Of course, certain types of variation become conventionalized over a period of time and take on new roles in word identification, boundary marking, etc. In other words, variability is also a productive source of new contrasts.)

The two principles proposed here are characterized in very informal terms (though hardly less so than those usually given to the older formulations of articulatory ease and auditory distinctiveness). This is inevitable, given how little we know at this stage about such matters as how speech motor patterns are

represented in memory and how signals as complex as those in speech are perceptually processed. However, it is nonetheless possible to use these principles as the basis for making some predictions about expected and unexpected patterns. From this point on, the paper will focus on certain predictions concerning disfavored segments.

Disfavored segments

The requirement of contrastivity leads first to the expectation that languages will prefer sequences of sounds that contain sufficiently differentiated elements. However, because identifiability is also crucial, individual elements that are hard to identify as distinct in motor or auditory terms because they are inherently bland (non-salient), confusable with competing sounds, or produce unstable acoustic outputs from small variability in production details are among the kinds of segments that are predicted to be disfavored. For lack of connectedness it is elements which are resistant to contextual variation — for example, because they make use extreme gestures or require precise positioning of several articulators — which would be expected to be disfavored. Three classes of segments which, given their 'text-book' definitions, might be expected to be disfavored for one or both of these reasons are doubly-articulated plosives, doubly-articulated fricatives, and ejective fricatives. These classes are certainly comparatively rare, though all have been mentioned as occurring in at least some languages.

We might expect that doubly-articulated plosives are not well-connected since they seem poorly adaptable to context, given that two articulators are pre-empted. This is to extend to this case an argument that was first made in connection with clicks, another class of segments in which two articulators are employed. (Compare the 'back vowel constraint' limiting click-vowel co-occurrence in !Xóó (Traill 1985), and Sands' (1991) conclusion, based on acoustic measurements, that clicks in Xhosa do not coarticulate with surrounding vowel environments.) Perceptually, stops are primarily distinguished one from another by the transitions preceding and following them, and when a burst is present, by the spectrum of the burst. Given the text-book definition of doubly-articulated plosives as involving two simultaneous articulations, we might also expect that they are not well-contrasted since simultaneous articulations would make it hard to recover cues to the two separate motor gestures involved, impeding their identifiability. Moreover, given a target of simultaneity of articulation, small variations in inter-articulator timing would produce great variation in the acoustic signal, according to which gesture leads or lags the other and hence dominates the transitions and bursts.

In the case of doubly-articulated fricatives, there are reasons to think that the problems are more severe than for stops. Doubly-articulated fricatives might be thought to be not well-connected for the same reason as doubly-articulated stops: two articulators are involved and this impedes coarticulation. For fricatives, perceptual identification is usually taken to be primarily based on the spectral characteristics of the frication noise itself. The impedance of a fricative constriction creates an elevated air pressure level in the mouth behind the constriction as air flows from the lungs. The driving pressure of the air supply coming from the

lungs must be higher than this in order to maintain a sufficient flow of air past the constriction to generate the frication noise. Now, if there are two constrictions, as illustrated in Figure 1, then air flow past the further back of the two has to overcome the impedance created by the more forward of the two, demanding a higher driving pressure from the lungs. This higher subglottal pressure would only be required while this particular articulatory configuration was maintained and would be inappropriate for flanking sounds. Hence, there is an additional reason for thinking doubly-articulated fricatives are poorly-connected in speech.

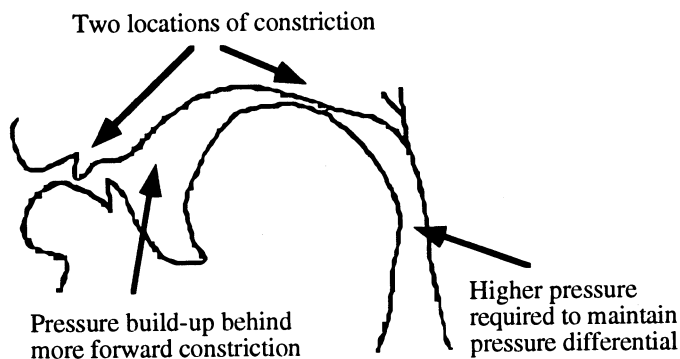


Figure 1. Schematic representation of production of a doubly-articulated fricative.

Moreover, doubly-articulated fricatives are not well-distinguished since the more forward of the constriction locations will act as an acoustic filter on the noise spectrum generated at the rear one, masking its place-specific characteristics, and making it hard to recover the articulatory gestures involved in production of the sound. We would expect doubly-articulated fricatives also to be acoustically unstable due to the very critical balance of various aerodynamic factors required to produce them; small variations around a mean would have greater consequences than would be the case for fricatives with only a single constriction.

Ejective fricatives, like other ejective segments, are judged to be poorly connected in speech because of the impossibility of producing a flow of speech using the ejective mechanism. By its nature, this can only be used for one segment-length interval before the laryngeal parameters of height and constriction must be re-set. However, ejective fricatives suffer from a further problem in that, in their textbook form, they would be of necessarily short duration and, for this reason, poorly distinctive. This is because raising the larynx with the glottis closed — the ejective mechanism — produces only a small change in oral cavity volume and this volume change is sufficient for brief frication only, as illustrated in Figure 2.

Based on data in the literature it is possible to make an approximate estimate of the anticipated duration of an ejective fricative. The average volume velocity estimates for the air flow in a pulmonic voiceless oral fricative are about $300 \text{ cm}^3/\text{s}$

(200-400 cm³/s in Shadle (1997); 330 cm³/s in Catford (1977)). An average fricative of 100 ms duration would therefore require a flow volume on the order of 30 cm³. We estimate the approximate magnitude of ejective larynx raising to be on the order of 1-2 cm and, using the volumes of the lower pharyngeal area shown in Baer et al's (1991) MRI study as a guide, the consequent reduction in the volume in the oral cavity to be 10 cm³ or less. This would produce enough outward air flow to sustain an ejective fricative with a duration on the order of 30 ms only, and quite likely less, as the estimates probably err on the generous side.

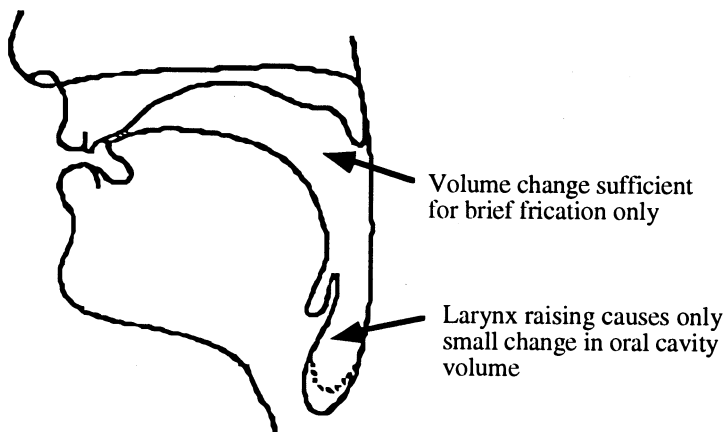


Figure 2. Schematic representation of production of an ejective fricative.

Thus, all three sound types appear disfavored for lack of contrastivity and connectedness. The severity of the problems mentioned also seem to correlate quite well with reported frequency. Doubly-articulated plosives are much rarer than singly-articulated ones; and the more seriously problematic doubly-articulated fricatives have been claimed to exist in only a handful of languages. Ejective fricatives are reported in far fewer languages than are ejective stops, and both are very much rarer than their pulmonic counterparts.

However, text-book descriptions and rough theoretical calculations do not always tell the whole story. As will be shown below, in individual cases the phonetic details concerning the way in which a 'text-book' category is realized show important deviations from standard descriptions. One might simply conclude that there are fewer cases of the textbook type than expected (and doubly-articulated fricatives may not exist at all, as suggested by Ladefoged and Maddieson (1996)), and leave it at that.

The perspective suggested here is to view the three types of 'ecologically unfit' segments discussed above in terms of the combinations of properties they represent, respectively:

— transitional cues to two places of articulation (doubly-articulated plosives)

- inherent cues to two places of articulation (doubly-articulated fricatives)
- laryngeal constriction + frication (ejective fricatives)

and to show how these combinations can get off the endangered list if they are ‘partitioned’, e.g. in the temporal or strictural domains, while retaining the organization typical of a single segment.

Doubly-articulated plosives

Segments written /kp/, /gb/ occur in many West African languages (and in some languages of New Guinea). These are typically described as doubly-articulated plosives with simultaneous bilabial and velar closures. A classic description of a /kp/ segment is that offered by Westermann and Ward (1933: 58):

“The two articulations must be simultaneous, i.e. when the sound occurs between two vowels there must be no onglide to the k heard before the lips come together for the p position.”

This description is closely echoed by Ladefoged (1968: 9) who talks of labial-velar stops produced by “the simultaneous articulation of k and p or g and b”.

We pointed out above the problems that strict simultaneity of articulations would engender. However, if the two articulatory gestures are slightly offset in time the presence of two articulations can be signaled much more strongly, since distinct onset and offset transitions can be heard. The risk of excess variability is also reduced by fixing the relative timing of the articulations. Maddieson (1993) showed that in Ewe the two articulatory gestures involved are indeed slightly offset in time in just this way, with the velar leading the labial one by some 20 ms.

Data on lip and tongue movement was collected using electromagnetic articulography from two Ewe speakers from Kpando, Ghana, reading short phrases including, among others, the following words:

ákpá	<i>too much</i>	ekpé	<i>stone</i>
àgbà	<i>load, trouble</i>	ekpo	<i>log</i>

Data on only one speaker will be shown here, but the patterns are the same for both. The onset and release of closure for the stops /kp/ and /gb/ were determined for each word from the acoustic records, and the movement tracks averaged for ten repetitions, aligned at the instant of acoustic release. Derivative measures, such as the distance between the upper and lower lip receivers (inter-lip distance) were also calculated. Figures 3 and 4 show results for ákpá and àgbà. Averaged vertical movement tracks over time of the tongue back and the lower lip, as well as the derived inter-lip distance measure (which effectively adds the much smaller movement of the upper lip to that of the lower lip) are plotted against time. The time interval shown is 400 ms, with the acoustically-determined release instant at 300 ms from the start of the window. In order to show the movements in a normalized space, the actual articulator heights were converted to standard scores. In both cases, the velar and labial gestures are highly overlapped, but are not simultaneous. Auditorily the onset sounds velar and the release labial. The production of these segments must explicitly be made contrary to the specification given by Westermann and Ward.

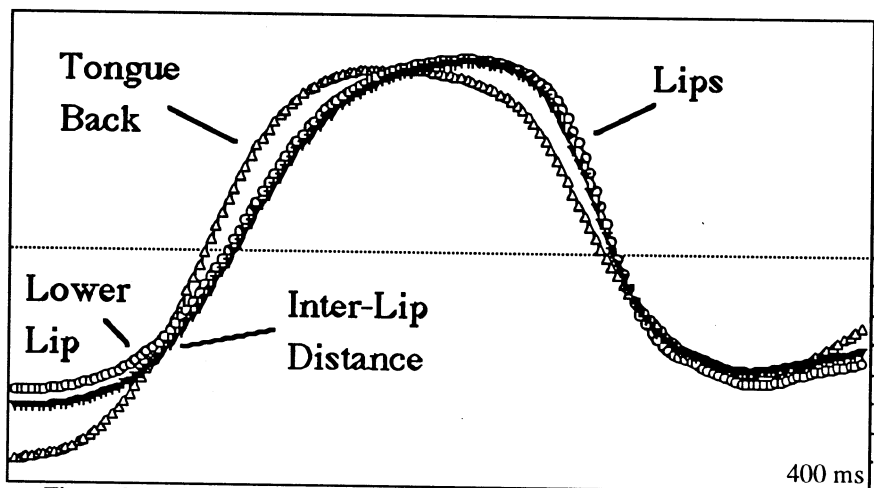


Figure 3. Ewe /akpa/ mean of ten repetitions. (Duration shown is 400 ms).

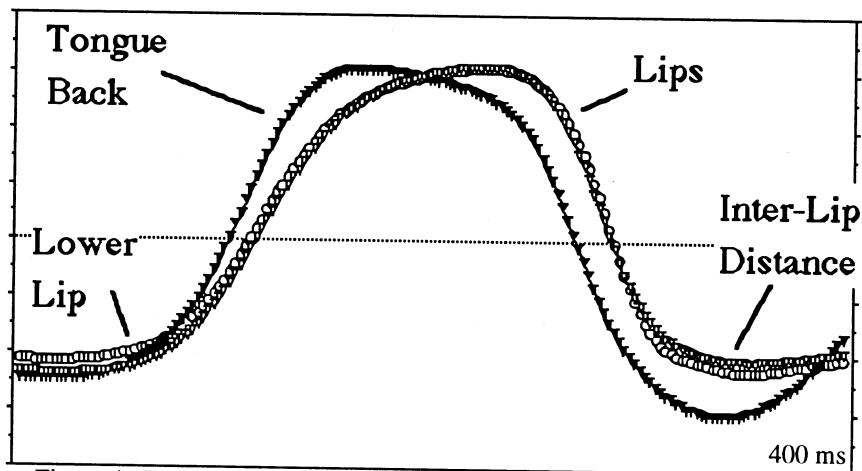


Figure 4. Ewe /agba/ mean of ten repetitions. (Duration shown is 400 ms).

The timing offset of the two articulations leads to a slightly longer total closure phase for labial-velar stops than for singly-articulated ones, as shown in Table 1.

Table 1. Mean consonant closure durations (in ms)

Velar:	/k/ 142	/g/ 133
Bilabial:	/p/ 158	/b/ 150
Labial-velar	/kp/ 174	/gb/ 179

Labial-velars are not, however, a sequence of separate segmental articulations, but a coordinated single entity, with very little timing variation. After averaging the

standard deviation of articulator positions for any time point is on the order of .1 to .2 mm, which is close to the accuracy limits of the measurement technique used. Contrast the English sequence /k/ + /p/, shown in Figure 5, which has generally less overlap of the two articulations and more variation in their timing.

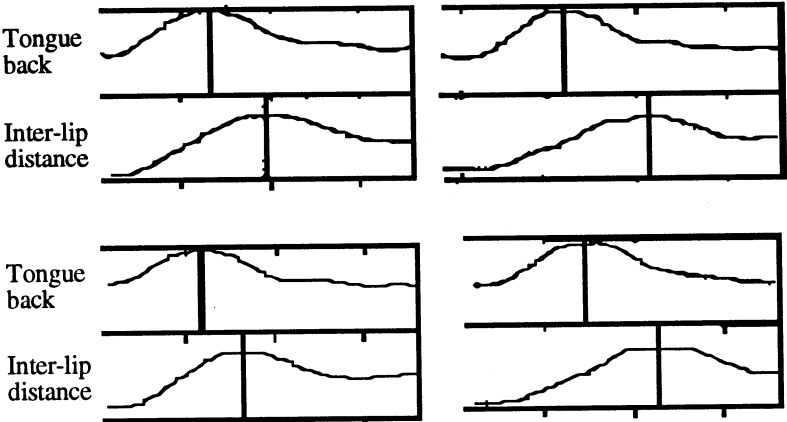


Figure 5. Movement traces (from x-ray microbeam) of 4 repetitions of the /k +p/ sequence in the English phrase ‘cock puddles’. Vertical lines show the consonant centers for the velar and labial gestures involved. (Data courtesy of Louis Goldstein, Haskins Laboratories.)

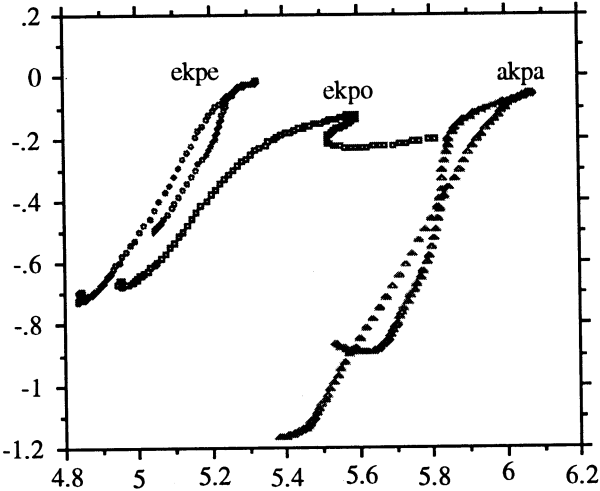


Figure 6. Coarticulation of Ewe labial-velars. Movement of the tongue back in two dimensions as a function of vowel environments. Distances on vertical and horizontal scales are in cm. The front is to the left of the figure.

Not only are the labial and velar gestures non-simultaneous, but contrary to a priori expectations, the velar component of labial-velars co-articulates well with the tongue body position of different vowel environments. This is shown in Figure 6. The same point on the tongue makes contact with the palate at three different locations over a span of almost a centimeter depending on the vowels in the three words shown.

Doubly-articulated fricatives

Avatime, another language of Ghana, according to Ford (1988), has a consonant inventory which seems to show a regular pattern of combining velar and bilabial articulations, so that a partial chart reads as follows:

p	k	kp
b	g	gb
ɸ	x	xɸ
β	ɣ	ɣβ

In view of the severe problems noted in producing doubly-articulated fricatives, it would be surprising if the combination of labial and velar articulation in fricatives was produced as indicated (Ladefoged (1964) wrote ɸ and ɣ for $x\text{ɸ}$ and $\text{ɣ}\beta$ but also describes these segments as fricatives).

As is shown in more detail in Maddieson (1998) the labial constriction in the fricatives which are [labial] + [dorsal] is not narrow enough to constrict air flow, and hence does not impede frication at the back closure. In fact, the nature of this articulation is quite different from that for a simple bilabial fricative, involving rounding and forward projection of the lips rather than a narrow approximation along their length. In other words, the putative $x\text{ɸ}$, $\text{ɣ}\beta$ are actually x^w , ɣ^w .

The method employed to study these productions was to video-tape the lip articulations of 4 speakers as they produced a variety of words containing labial articulations. Dots were placed on the outer lip surfaces to facilitate measurement. A frame from the videotape showing one speaker's most constricted lip position in the word /ax^wa/ "charcoal" is shown in Figure 7. (The video image has been dithered to improve printability.)

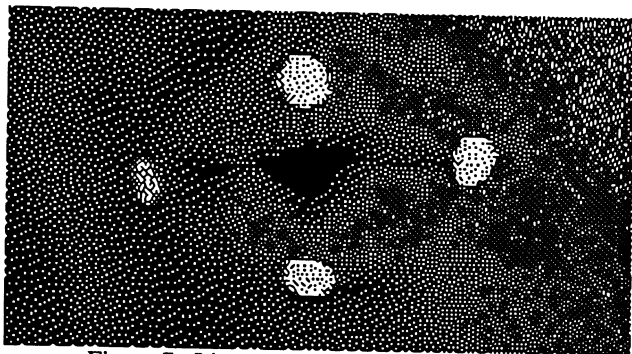


Figure 7. Lip-rounding in Avatime /x^w/.

We thus see that although frication at two locations is absent, a more general feature of labiality is easily combined with frication at another location by making it a vowel-like articulation with long-range transitional effects on adjoining vowels.

Ejective fricatives

For ejective fricatives, the problem is combining glottal constriction with sufficient duration of frication to cue place. In Yapese, an Austronesian language spoken in the Micronesian state of Yap, Hsu (1969) and Jensen (1977) specifically describe both the glottalized stops and fricatives as ejective. In his grammar, intended to be reasonably intelligible to a lay reader, Jensen puts it as follows:

"[...] for p', in contrast with p, while the lips are closed the glottis is also closed. With the lips closed, the closed glottis is raised very rapidly and suddenly. This motion compresses the air which is trapped in the mouth between the closed glottis and the closed lips, so that when the lips are suddenly opened, a popping sound is made. [...] Only then is the closed glottis opened. [...] The other glottalized stops are made in the same way, except [for place of articulation]." (Jensen 1977: 28)

"Glottalized fricatives are pronounced in a similar way to glottalized stops. [...] In pronouncing f', the lips are not closed completely. Rather, the lower lip is pressed against the upper teeth in such a way as to restrict the flow of air. So when the closed glottis is raised (as for p'), the air is forced out [...] making a hissing (friction) noise. Then the glottis is opened. th' is pronounced in the same way, except that the restriction [...] is produced at the teeth." (Jensen 1977: 31)

The production of the glottalized fricatives was examined using both audio recordings and aerodynamic data provided by a number of speakers of Yapese. The principal words used as examples of the sounds being examined are the following (maintaining the use of [''] to mark the glottalized category):

f	faaŋ	<i>platform</i>	f'	f'aaŋ	<i>eel</i>
f	faak	<i>his child</i>	f'	f'œθ	<i>apportion</i>
θ	θaam	<i>outrigger</i>	θ'	θ'aab	<i>cut</i>
θ	θiin	<i>language</i>	θ'	θ'iib	<i>pot</i>
θ	maaθ	<i>touch</i>	θ'	maaθ'	<i>severed</i>

Measurements of the frication duration showed that plain and glottalized fricatives had similar noise durations, contrary to the theoretical expectations for 'text-book' ejective fricatives. Mean frication durations for the four fricatives compared are shown in Figure 8. There is a salient difference in duration according to place, with the dentals longer than their labio-dental counterparts, but no significant difference between the pulmonic and glottalized categories.

Analysis of the production of the glottalized category shows that in fact these segments are not ejectives. Audio waveforms, and oral flow and intra-oral air

pressure traces are shown in Figures 9 and 10 for representative tokens of /f/ and /f'/ in word-initial position produced by a female speaker. Comparison of the oral pressure traces shows patterns corresponding to the fricative portions that are quite similar in amplitude and duration, giving them similar overall shapes. Differences lie in the fact that the pressure rise occurs considerably earlier with respect to vowel onset in Figure 10, and is then followed by a flat (silent) interval and a glottalized onset to the vowel. In other words, /f'/ is a sequence of an ordinary pulmonic fricative and a glottal stop.

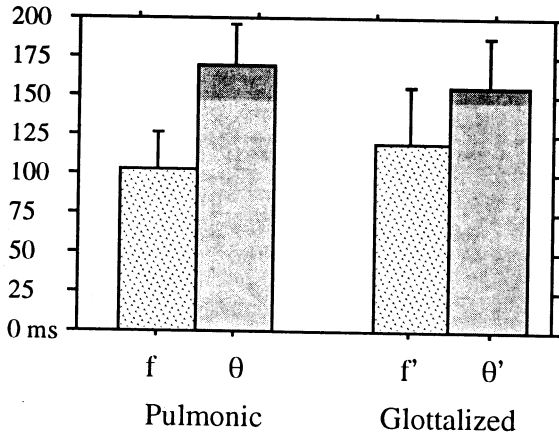


Figure 8. Frication noise durations of four Yapeese fricatives. Mean durations of 2-4 tokens per category from each of three speakers.

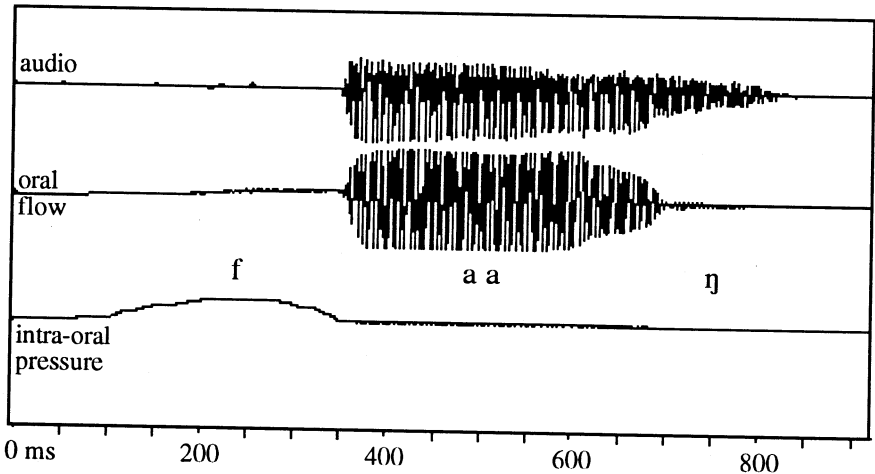


Figure 9. Aerodynamic records illustrating the Yapeese word /faay/ 'platform'.

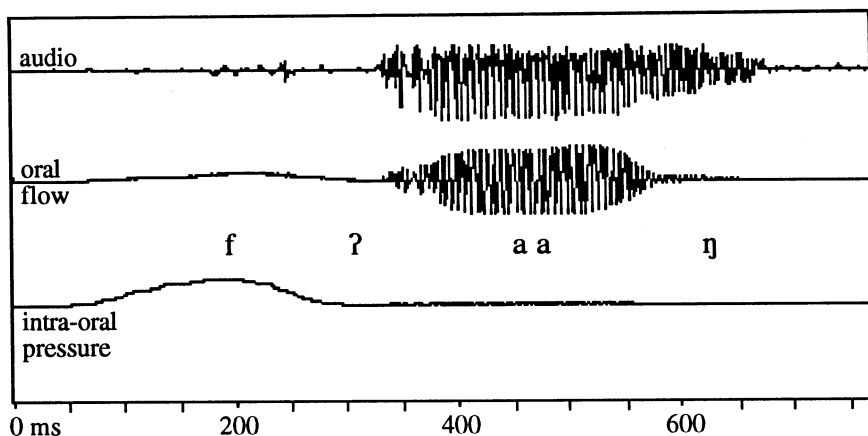


Figure 10. Aerodynamic records illustrating the Yapeese word /f'aar/ 'eel'

Yapeese does indeed use the ejective mechanism in its stops. Figure 11 shows records of a word beginning with /p'/. With the oral passage closed the ejective mechanism generates quite high intra-oral pressures (MacEachern 1996), higher than those typically generated by air pressure from the lungs during speech. Measurements of the peak intra-oral pressure during selected fricatives and stops in Yapeese are shown in Figure 12 separated by speaker. Note that the peak pressure is very similar for all the segments in which this peak value is attributed to pulmonic effort (all fricatives and the pulmonic stop /p/), but is much higher for /p'/. (These means are in fact an underestimate because some ejective stops showed clipping of the peak value, as in Figure 11.)

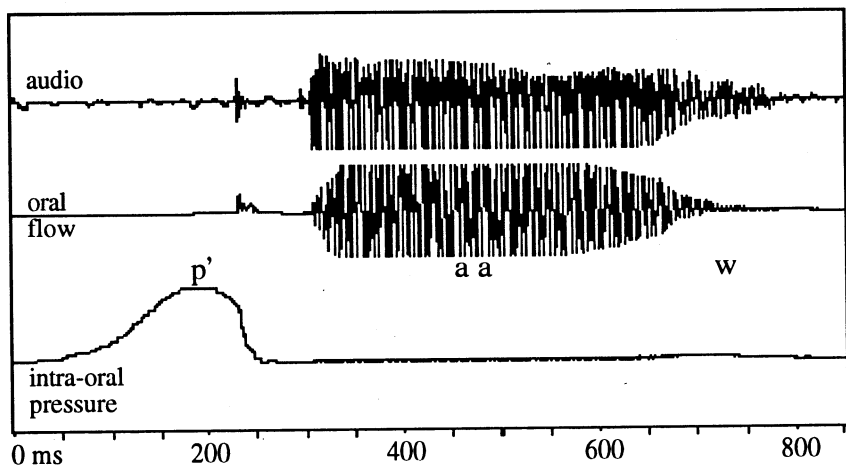


Figure 11. Aerodynamic records illustrating the Yapeese word /p'aaw/ 'banana'

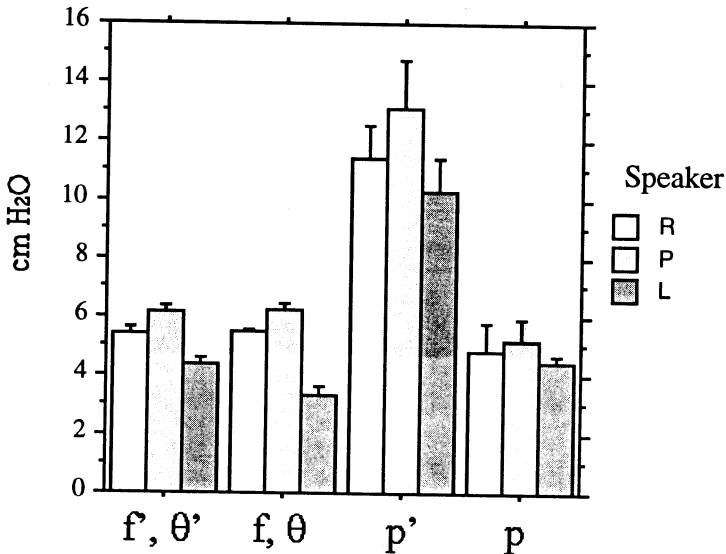


Figure 12. Peak intra-oral pressure: ejective stops vs fricatives and pulmonic stops (means of 3-12 tokens per category per speaker),

Although the Yapese solution is not the only way to combine glottal constriction and frication (see Maddieson, Bessell, and Smith 1996 for a different strategy employed in Tlingit), it is one way to be able to generate sufficient frication duration for place identification while maintaining distinctiveness from plain fricatives.

Conclusion

The three sound types which appeared disfavored for lack of contrastivity and connectedness if we took at face value the descriptions offered, are shown in specific cases to be organized so that the expected problems do not arise. In general the solution is to have a 'partition' of the components so the constituents are not so hard to identify or to produce:

- transitional cues to two places of articulation can be provided by having overlapping but temporally offset articulations
- the problem of providing inherent cues to two places of articulation in a fricative can be avoided by using different degrees of stricture
- laryngeal constriction + frication can be easily combined by having non-overlapping articulations.

In short, the difficulty of segments and their fitness for survival cannot always be appropriately determined from standard descriptions. We suggest that understanding of what conditions are truly imposed on the sound structure of language can be enhanced by detailed examinations of what happens in individual languages. It turns out life may not always be so hard as we make it appear.

References

- Baer, T., Gore, J. C., Gracco, L. C. and Nye, P. W. 1991. Analysis of vocal tract shape and dimensions using magnetic resonance imaging: vowels. *Journal of the Acoustical Society of America* 90: 799-828.
- Catford, J. C. 1977. *Fundamental Problems of Phonetics*. Indiana University Press, Bloomington, IN.
- Hsu, R. W. 1969. *Phonology and Morphophonemics of Yapese*. Ph. D. dissertation. University of California, Berkeley. Pp 183.
- Jensen, J. Thayer (with L. D. Pugram, J. B. Iou, and R. Defeg). 1977. *Yapese Reference Grammar* (PALI Language Texts: Micronesia). University of Hawaii Press, Honolulu.
- Ladefoged, P. and Maddieson, I. 1996. *The Sounds of the World's Languages*. Blackwell, Oxford.
- Lindblom, B. 1980. Models of phonetic variation and selection. *PERILUS* 11: 65-100.
- Lindblom, B. and Maddieson, I. 1988. Phonetic universals in consonant systems. In *Language, Speech and Mind: Studies in Honor of Victoria A. Fromkin*, ed. L. M. Hyman and C. N. Li. Routledge, London and New York: 62-80.
- MacEachern, M. R. 1996. Computer modeling of the production of ejectives using aerodynamic and acoustic data (Abstract). *JASA* 100: 2661.
- Maddieson, I. 1997. Phonetic universals. In *Handbook of Phonetic Sciences*, ed. W. J. Hardcastle and J. Laver. Blackwell, Oxford.
- Maddieson, I. 1998. Disappearing vowel harmony and doubly-articulated fricatives: two myths about Avatime phonology. In *Language History and Linguistic Description in Africa*, ed. I. Maddieson and T.. Hinnebusch. Africa World Press, Trenton, NJ.
- Maddieson, I., Bessell, N. and Smith, C. L. 1996. A Preliminary Report on the Phonetics of Tlingit. *UCLA Working Papers in Phonetics* 93: 125-148.
- Sands, B. 1991. Evidence for click features: Acoustic characteristics of Xhosa clicks. *UCLA Working Papers in Phonetics* 80: 6-37.
- Shadle, C. H. 1997. The aerodynamics of speech. In *Handbook of Phonetic Sciences*, ed. W. J. Hardcastle and J. Laver. Blackwell, Oxford: 33-64.
- Traill, A. A. 1985. *Phonetic and Phonological Studies of !Xóõ Bushman*. Helmut Buske, Köln.

Glottalized Sonorants: A Phonetic Universal?

Madelaine C. Plauché, Rosemary Beam de Azcona, Rungpat Roengpitya,
and William F. Weigel
University of California at Berkeley

1. Introduction

It has previously been argued that the glottalization of sonorants should be free to surface anywhere and that they tend to be preglottalized in the world's languages (Kingston, 1983). We have found in our preliminary phonetic investigation of Lai,¹ Coatlàn-Loxicha Zapotec,² and Yowlumne,³ as well as a survey of detailed phonetic descriptions of the segments in other languages, that for those languages that rely mainly on creaky voice, full glottal stop, and amplitude as phonetic cues for glottalization, these segments will surface as preglottalized in onset and postglottalized in coda. This pattern is thought to be the result of the obscuring nature of these particular cues, in which segments preserve the most information if the glottalization does not co-occur with the crucial sonorant to vowel or vowel to sonorant transition.

This paper begins by predicting a universal tendency for the phonetic structure of glottalized sonorants based on previous analyses and data on the perception of sonorant-vowel (NV) and vowel-sonorant (VN) sequences (Section 2). A preliminary acoustic analysis of both Yowlumne and Coatlàn-Loxicha Zapotec are shown to support these predictions (Section 3.1, 3.2), but Lai is found to work contrary to our claim (Section 3.3). The paper concludes with a discussion of how knowledge of universal physical and acoustic constraints can predict what we actually find in the phonology of the world's languages (Section 4).

2. A universal phonetic structure of glottalized sonorants

Glottalized sonorants are phonetically sonorants (usually nasals, rhotics, laterals, or glides) accompanied by a constriction of the glottis caused by tightening the cricoarytenoid, found, for example, in a variation of the word 'couldn't' in English: [kʌd̥n̥]. Although relatively rare in the world's languages as a distinct phoneme (Maddieson 1984), the glottalized sonorant surfaces with various salient acoustic cues to mark its distinction from plain sonorants. The main structure is a sonorant produced with creaky voice, an irregular voicing modality, 'in which the arytenoid cartilages are much closer together than in modal voice. Creaky voice also involves a great deal of tension in the intrinsic laryngeal musculature, so that the vocal folds no longer vibrate as a whole.' (Ladefoged et al. 1996:53). Sometimes this constriction is complete, yielding a phonetic glottal stop. Other secondary acoustic domains may cue the perception

of glottalized sonorants in a variety of languages as well. These include pitch, preceding vowel duration, sonorant duration, and bandwidth.

In his study of articulatory binding of glottalization, Kingston (1983) found that the complete oral closure of stops allowed a build-up of pressure in the oral cavity, causing a high-energy, salient acoustic event at the release of stops. Stops with a secondary laryngeal constriction tend to be postglottalized as a result of glottalization fixing itself to the salient stop release. In the case of sonorants, however, as a result of oral or nasal leakage, no equivalent air pressure build-up occurs. Thus, unlike stops, they do not have an equivalent asymmetrical design; their onset is much like a mirror image of their offset. Therefore, they should be free to have glottalization show up anywhere during the oral closure⁴. This paper will show, however, that many languages exhibit glottalized sonorants that are strictly preglottalized, but some show an even more specific structure: preglottalized in onset position, but postglottalized in coda.

In a study of voiceless nasals in Burmese and the Hmar dialect Mizo, Ladefoged and Maddieson found that voiceless nasals in initial position had substantial voicing in the last part of the oral closure.

'Ladefoged (1971) and Ohala (1975) suggest that an early onset of voicing helps to distinguish one voiceless nasal from another by making the place of articulation more apparent. This is because the voiced offglide from the nasal into the vowel displays formant transitions that are characteristic of each place of articulation.' (Ladefoged and Maddieson 1996:113).

Based on the behaviour of the voiceless nasals, Silverman (1995) then argued that because the CV transitions are primary in conveying information about the place and manner of nasals (Fujimura 1962), and

'a heavy glottal constriction may result in sufficient aperiodicity, or jitter, to disrupt the acoustic encoding of a salient nasal formant structure. ... if glottal pulse (quasi-periodicity) is markedly slow—a common result of creakiness—insufficient energy is present during the crucial transition period; transitions may take place during the relatively long periods of glottal closure. Consequently, formant transitions may be rendered unrecoverable.' (Silverman 1995:70)

By extending this argument to glides, he explains the tendency for languages to have preglottalized sonorants. This argument works well for sonorants in onset, but Silverman's concluding remarks of this argument reveal that it misses an entire set of sonorants: those in the coda: 'to optimize recoverability, sonorant consonants are realized with laryngeal gestures phased to the early portion of the

supralaryngeal configurations, ... in order to mitigate the potential non-salience of formant transitions into a following vowel.' (Silverman 1995:78)

The present paper extends this reasoning to those sonorants occurring in the coda and takes a closer look at the acoustic cues involved in glottalization of sonorants. Silverman mentions the irregular glottal pulses, the low amplitude, and lower F_0 , but we have found in Lai, for example, that secondary cues may include vowel length and sonorant length. If preglottalization is an effort to preserve as much information about the sonorant by keeping the inherently obscuring cues of glottalization (creaky voice and dip in amplitude) from compromising the crucial sonorant to vowel transition, then we would also predict that in coda position, where the crucial transition is not sonorant to vowel, but vowel to sonorant, we should find postglottalization.

Note that this analysis suggests that glottalized sonorants are likely candidates for sound change, as cues for either place or glottalization may be compromised, since they work antagonistically. If creaky voice obscures the cues for sonorants, we expect to find cases of de-glottalization or de-sonorization historically; below are two compelling examples.⁵ This is not to say that there is a universal tendency for glottalized sonorants to disappear; the creation of these segments from, say, a glottal stop and sonorant sequence is equally productive.

In Wakashan, 'glottalization floats off sonorants under certain circumstances, while remaining stably linked to stops under the same circumstances.' (Kingston 1983:320) Both loss of glottalization and loss of sonorancy are found in Kashaya: Glottalized sonorants de-sonorize (surface as [b] and [d]) word-initially and de-glottalize: 'if a glottalized sonorant at the end of a word resyllabifies as the onset, it loses its glottalization' (Buckley 1988:49):

/ma ² ne'mu/	→	mané'mu	'it's her'
/dolo ² me'mu/	→	do lo mé'mu	'it's a wildcat'

3. Preglottalized in onset, postglottalized in coda.

3.1. Zapotec: postglottalized in coda position.⁶

Coatlán-Loxicha Zapotec (CLZ) follows our first prediction: the glottalized sonorants [n^2 , w^2 , l^2 , j^2], which only surface in word-final position, consistently surface postglottalized. We find them only word-finally, the result of a historical vowel deletion which caused previously-medial sonorants to become word-final. Only when these sonorants became part of the rime could suprasegmental features such as tone and glottalization map onto them. Glottalized sonorants in CLZ are cued by either irregular F_0 and/or by a full phonetic glottal stop⁷. Even though

historically the glottal feature may have come from either a preceding or following syllable, these segments are always postglottalized, with the irregular voicing modality starting 50ms after the onset of the sonorant and persisting to the end of the sonorant, often with a glottal release marking the end of the word. In the table below, note that whether the glottalized sonorant arose from a plain sonorant combined with a preceding, a following, or a surrounding pair of glottal stops, they consistently surface as postglottalized segments in CLZ today.

<u>Original sequence</u>	<u>Proto-Zapotec</u> (Kaufman, 1995)	<u>Coatlán-Loxicha</u> <u>Zapotec</u>	<u>English gloss</u>
*ʔS	*p+e: ʔlla	mbǣl	‘snake’
*SVʔ	*ʃilaʔ	ʃil	‘cotton’
*ʔSVʔ	*pe: ʔlaʔ	bǣl	‘meat’

Table 1. Historical origin of glottalized sonorants in CLZ.⁸

Creaky voice is produced by tightening the cricoarytenoid which causes the vocal cords to increase in mass at the point where the air mass passes through. This causes not only an irregularity in the vocal pulse, but a general drop in the rate of vibration. Speakers of CLZ and other languages⁹ consistently produce glottalized sonorants in VN sequences with an associated high pitch on the preceding vowel, presumably as a re-interpretation of the contrasting drop in pitch on the sonorant itself when produced with glottalization. In CLZ, glottalized sonorants are found to cause pitch patterns to play out in higher frequencies on the preceding vowel, consistently enough to have developed into a fixed allotony in some cases. In the case of the rising tone, for example, this perturbation produces a regular allotony, conditioned by glottalization on either vowels or sonorants, such that the low-to-high rising tone surfaces as a mid-to-high rising tone in the Loxicha dialect, and a high-to-very high rising tone in the Coatlán dialect.

	<u>VN (Average frequency)</u>		<u>VN¹ (Average frequency)</u>	
Loxicha	120>180 Hz.	Low to High	140>180 Hz.	Mid to High
Coatlán	130>170 Hz.	Low to High	190>220 Hz.	High to Very High

Table 2. The frequencies of the beginning and end of 100 rimes for each dialect were measured from a pitch extraction and averaged. In both dialects, the phonemic rising tone, low-to-high rising in non-glottalized environments, surfaces with a higher allotone when preceding glottalized sonorants or glottal stop.

Despite the variety of acoustic cues available to these segments in CLZ, the glottalized sonorants consistently surface with either irregular glottal pulses or a

full glottal stop, compromising the communication of any brief formant structure, such as that found in the transition from vowel to the word-final sonorant. By consistently surfacing as postglottalized, these glottalized sonorants preserve the maximum acoustic cues about the place and manner of the sonorant itself.

3.2. Yowlumne: preglottalized in onset, postglottalized in coda.

A preliminary phonetic investigation of Yowlumne shows it complies to our prediction to the fullest extent: Glottalized sonorants in this language are consistently produced with preglottalization in onset position, but postglottalization in coda position. Unlike CLZ, Yowlumne has glottalized sonorants in onset (do not occur word-initially), where they are preglottalized, and in coda position (both preconsonantly and word-finally), where they are postglottalized. These segments are acoustically distinct from plain sonorants in their irregular F_0 and dip in amplitude, those cues previously discussed as being responsible for obscuring the place of the sonorant.

VNC		VN#		VNV	
<u>Yowlumne</u>	<u>Gloss</u>	<u>Yowlumne</u>	<u>Gloss</u>	<u>Yowlumne</u>	<u>Gloss</u>
lan [?] t'e	'left'	hajal [?]	'day'	ʔa [?] waŋ	'dislike'
xol [?] pojo	'lizard'	tsij [?]	'bone'	taa [?] muŋ	'whiskers'
ts'ol [?] lol	'white'	nukum [?]	'bend'	ti [?] miŋ	'eyebrow'
bim [?] t'ana	'stump'	laŋjaw [?]	'steep'	no [?] no	'man'

Table 3. Glottalized sonorants in Yowlumne: preglottalized in onset, postglottalized in coda (as determined by visual inspection of spectrograms (Figure 1)).

In Yowlumne, the most consistent and prevalent acoustic cue for glottalized sonorants versus plain sonorants is the presence of creaky voice. Creaky voice, or jitter, is not the only acoustic cue for glottalized sonorants: these segments often surface with no sign of creaky voice, but with a drop in amplitude across the onset of the sonorant in VN sequences and across the offset in NV sequences. As previously noted, creaky voice is not only characterized by irregular glottal pulses, but often by long and irregular intervals of time in between each pulse. This has the overall affect of dropping the amplitude during the constriction. In Yowlumne, the dip in amplitude may have become a primary cue for glottalization on sonorants: On average, glottalized sonorants exhibit a drop in amplitude of 10- 20 decibels for nasals and laterals (5-10 dB for glides) compared to the drop in amplitude from a vowel to a plain sonorant which is 1-5 dB for plain sonorants.

As we observed in CLZ, the glottalized sonorants in Yowlumne are found to devoice when in coda position. Devoicing happens after, not instead of, glottalization. This was noted by Newman for Yowlumne and other Yokuts

languages: ‘The glottalized consonants, w̥, j̥, l̥, m̥, n̥, and ŋ̥, when they occur finally in a word or in a closed syllable, are heard as -w̥^{hw} or even as -^{ch}w̥, etc.’ (Newman 1944:17)¹⁰ The tendency for voicelessness and glottalization to co-occur, is certainly related to the fact that voiceless glottalized sonorants do not exist distinctively in any language. (Maddieson 1984) This is to be expected, since the articulators needed to produce both of these voicing modalities are the vocal cords: it is physically impossible to constrict the glottis, causing the vocal cords to be closer together than in modal voice, while at the same time holding the vocal cords apart for true voicelessness. It is interesting to note that some of the overall acoustic effects for these two voicing modalities are similar: namely drop in amplitude and irregular output of energy.

3.3. Lai: an apparent exception.

Lai appears to contradict the principles discussed so far in this paper: Although this language only has glottalized sonorants [²m, ²n, ²ŋ, ²w, ²r, ²l, ²j] in coda position, they show variation of phonetic production, but are mainly preglottalized. This phenomenon cannot be fully explained by the obscuring nature of jitter on the NV transition, as explained by Silverman, as there is no ‘following vowel,’ not even historically: these segments were derived from a final -s suffix (Proto-Tibeto-Burman *zey-smans > zey-hman² ‘whatever’ (J. Matisoff p.c.)), presumably with an intermediate stage of glottal stop followed by a sonorant (cf. Roengpitya 1998). The sequence of a sonorant followed by a glottal stop was probably re-interpreted as a glottalized sonorant, which now surfaces as either preglottalized, simultaneously glottalized, or postglottalized.

<u>Verbal Forms:</u>	<u>Form I</u>	<u>Form II</u>	<u>Gloss</u>
	kaaŋ	ka ² ŋ	‘burn’
	ʔaj	ʔa ² j	‘eat’
	hjaɭ	hja ² ɭ	‘know’
	saaw	sa ² w	‘prolong’
<u>Nominal Form:</u>		hŋe ² r-tee	‘ants’

Table 4. Synchronic pattern of glottalized sonorants in Lai. The position of the glottalization relative to the sonorant varies, but it is most often preglottalized.

To understand the exception of Lai and other languages¹¹ to our prediction that languages should prefer postglottalized sonorants in the coda, we must look further at the phonetic and phonological structure of this language specifically. The phonetic structure of the glottalized sonorants differed from both CLZ and Yowlumne in that vowel and sonorant length varied consistently with glottalization. Also, a look at the phonology of Lai shows that stops are

unreleased in word final and word medial codas, suggesting a different structure for sonorants as well.

Although creaky voice, and often full glottal closure preceding, simultaneous to, and following the sonorant are found as acoustic cues for glottalization in Lai sonorants, other secondary cues may be involved. We found that both the length of the preceding vowel (The average vowel length across 10 minimal pairs for the three nasals were 135msec before plain and 102msec before glottalized nasals) and the length of the sonorant itself (The average sonorant length across 30 minimal pairs including all sonorant types were 268msec for plain and 77msec for glottalized sonorants) greatly differ depending on the absence or presence of glottalization on the sonorant. This was found both word-finally and in word-medial coda positions. If vowel and sonorant length are used by listeners to identify plain and glottalized sonorants, cues such as jitter and amplitude drop might be less essential to the production of glottalized sonorants. Jitter and amplitude are found in the acoustic signal of Lai glottalized sonorants, however, and so the problem of these cues masking those for place when they occur at the VN transition remains.

Another relevant fact about Lai is that all codas are unreleased for stops and sonorants. This suggests that although the production of preglottalized sonorants in the coda may obscure the place of articulation, the glottalization itself is more likely to be perceived if it occurs at the beginning of the sonorant than at the end, since there is no equivalent dynamic transition after the sonorant.

4. Conclusion.

Although the simultaneous oral and laryngeal constrictions for glottalized sonorant are not bound by production to surface with a certain temporal restriction relative to one another, and indeed in Lai and other languages they do exhibit large variations in production from speaker to speaker and utterance to utterance, Yowlumne, Coatlàn-Loxicha Zapotec, and other languages exhibit a particular phonetic structure for glottalized sonorants: preglottalization in the onset and postglottalization in the coda. The main acoustic cues for glottalization (creaky voice, amplitude, and bandwidth) may obscure those for the place of the sonorant itself. The pattern discussed here is perhaps an effort to preserve the most information about the sonorant by restricting the obscuring secondary cues to the non-essential part of the speech signal: the vowel-to-sonorant or sonorant-to-vowel transition.

In Lai, however, facts about possible secondary cues and the way stops and sonorants behave in the coda in general suggests that for this language, the tendency need not be followed. Although the phonetic structure discussed may

preserve the most information about the segments involved, it does not predict a phonological universal. It can only explain the cross-linguistic *tendency* for languages to preglottalize sonorants in the onset and postglottalize them in the onset. However, facts about the way specific languages function, including extra-phonetic facts (phonological, morphological, etc...), must be considered to determine whether this tendency is borne out on a case by case basis.

This paper is truly a preliminary look at the phonetic structure of glottalized sonorants in these languages. To determine what acoustic cues are used by listeners to identify these segments, it would be useful to run a series of perceptual studies, especially cross-linguistically. Also, as much of this and previous studies have mainly focused on glottalized nasals, more work should be done specific to laterals, glides, and rhotics.

Spectrogram.

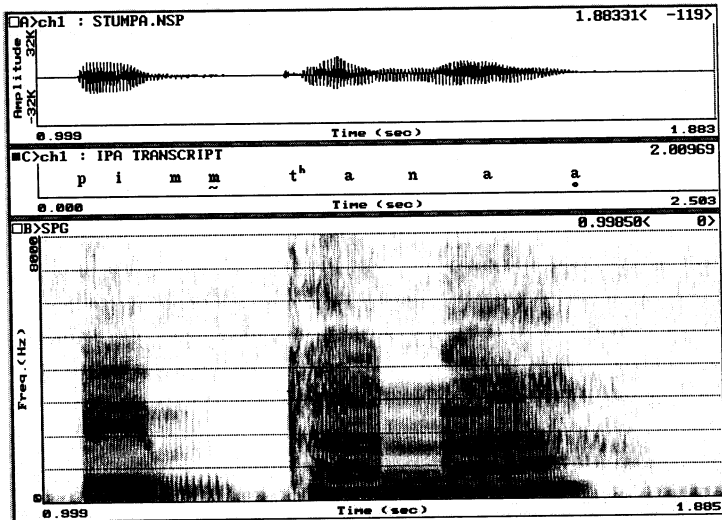


Figure 1. Yowlumne: pimʔana 'stump'. The irregular F₀ starts after the onset of the sonorant. This word is consistently postglottalized.

Notes

¹ A Tibeto-Burman language spoken in Northwestern Burma.

² A Southern Zapotec language spoken in Miahuatlán and Pochutla in the state of Oaxaca, Mexico.

³ A language of the Yokuts family, spoken in the foothills of the Sierra Nevada. Referred to in previous literature as Yawelmani.

⁴ In Montana Salish, we find true variation in the timing of the laryngeal constriction: 'Pronunciations of the word for 'soft' include [ʔə²m²m²móts], [ʔəm²m²:óts], and [ʔə²m²:móts], suggesting an underlying representation with three glottalized nasals, ʔə²m²m²móts, not all of which are realized.' (Flemming et al. 1994:16)

⁵ Also, glottalized sonorants de-glottalize or do not occur in postconsonantal position in Yowlumne and Shuswap. (Kingston 1983). In Klamath, sonorants de-glottalize word-finally (Barker 1964).

⁶ Other languages that follow this pattern (cf also Kraft 1973, Samarin 1966, Sapir et al. 1955, and Smith 1968): In Klamath, glottalized sonorants are mainly preglottalized. In coda position, however, glottalized sonorants 'are distinguished by a simultaneous glottal stricture instead of preglottalization.' (Barker 1964:27). 'Kashaya places the glottal constriction at the beginning of the nasal when the consonant is syllable-initial, and at the end when it is syllable-final.' (Ladefoged et al. 1996:110; Buckley 1990, 1993). In Logbara, the glottalized sonorants are preglottalized and only occur word initially (Crazzolara 1960).

⁷ Although historically the glottalization from both the first and second syllable of Zapotec roots was mapped onto the now-coda sonorant, synchronically suffixes such as [-m] and [-n] do not glottalize when added to roots whose vowels are checked.

⁸ Forms have been retranscribed to IPA.

⁹ In Acoma, 'the high accent on single vowels is usually realized as a high pitch with a slight fall if (1) the following syllable is unaccented and starts with a glottalized sonorant (...)' (Miller 1965:17). The stød in Standard Danish is a constriction of the vocal cords that is morphologically determined. When it occurs on a root that ends in a sonorant, a phonetic glottalized sonorant is the result, in which, 'the first part of the syllable is characterized acoustically by a higher pitch level and often a higher intensity level than syllables without stød (...). In the second part, the stød phase proper, there is a considerable decrease in intensity (...) and a noticeable decrease in fundamental frequency, and/or aperiodicity' (Fischer-Jørgensen 1989). In Kammu, glottalized sonorants only occur with high tone (Svantesson 1983).

¹⁰ This phenomenon has also been observed in Montana Salish: 'Glottalized nasals are preglottalized even in final position or before a voiceless consonant, as in sts'ó²m 'bone'. In these cases, the nasal portion is devoiced or creaky... In final position, the glide portion (of a glottalized sonorant) is realized as a voiceless, or very creaky, version of the equivalent vowel.' (Flemming et al. 1994:16)

¹¹ Gitksan has preglottalized sonorants in both word-initial and word-final position. (Wickstrom 1974). Montana Salish glottalized sonorants are typically preglottalized, even in final position (Flemming et al. 1994).

References

- Barker, M.A.R. 1964. Klamath Grammar. Berkeley: University of California Press.
- Buckley III, E. L. 1988. Theoretical Aspects of Kashaya Phonology and Morphology. Ph.D. dissertation. University of California, Berkeley.
- Crazzolaro, J. P. 1960. A Study of the Logbara (Ma'di) Language. London: Oxford University Press.
- Fisher-Jørgensen, E. 1989. Phonetic Analysis of the Stød in Standard Danish. *Phonetica* 46.1-59.
- Flemming, E., P. Ladefoged, and S. Thomason. 1994. Phonetic Structures of Montana Salish. *Fieldwork Studies of Targeted Languages II*. Los Angeles: UCLA Working Papers in Phonetics 87.
- Fujimura, O. 1962. Analysis of Nasal Consonants. *Journal of the Acoustical Society of America* 34: 1865-1875.
- Kaufman, T.S. 1995. Proto-Zapotec Reconstructions. Unpublished Ms.
- Kingston, J.C. 1983. The Phonetics and Phonology of the Timing of Oral and Glottal Events. Ph.D. dissertation. University of California, Berkeley.
- Kraft, C.H. and M.G. Kraft. 1973. *Introductory Hausa*. Berkeley: University of California Press.
- Ladefoged, P. and I. Maddieson. 1996. *The Sounds of the World's Languages*. Cambridge: Blackwell Publishers.
- Ohala, J. J. 1975. Phonetic Explanations for Nasal Sound Patterns. In *Nasálfest: Papers from a Symposium on Nasals and Nasalization (1975)*, ed. by C. A. Ferguson, L.M. Hyman and J. J. Ohala. Language Universals Project, Stanford: 289-316.
- Maddieson, Ian. 1984. *Patterns of Sounds*. Cambridge: Cambridge University Press.
- Miller, W. R. 1965. *Acoma Grammar and Texts*. Berkeley: University of California Press.
- Newman, S. 1944. *Yokuts Language of California*. New York: Viking Fund Publications in Anthropology No. 2.
- Roengpitya, R. 1998. Glottal Stops and Glottalization in Lai (Connected Speech)
- Samarin, W. D. 1966. *The Gbeya Language*. Berkeley: University of California Press.
- Sapir, E. and M. Swadesh. 1955. Native Accounts of Nootka Ethnography. *International Journal of American Linguistics* 21/4 Part II.
- Silverman, D. 1995. Phasing and Recoverability. Ph.D. Dissertation. University of California, Los Angeles. *Dissertations in Linguistics*. Number 1.
- Smith, K. D. 1968. Laryngealization and De-Laryngealization in Sedang Phonemics. *Linguistics* 38.
- Svantesson, J-O. 1983. *Kammu Phonology and Morphology*. Lund: CWK Gleerup. *Travaux de l'institut de linguistique de Lund* 18.
- Wickstrom, R. 1974. *A phonology of Gitksan*. M.A. Thesis. University of Victoria.

VOWEL HEIGHT: Reconsidering Distinctive Features

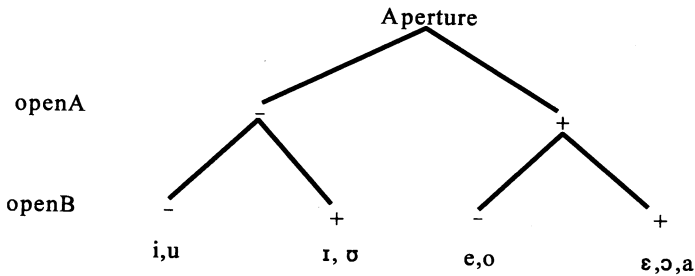
Don Salting

Communication Disorders Technology, Inc., Indiana University

1. Introduction

A model of vowel height consisting of a two-tiered, symmetrical hierarchy of autonomous nodes is presented as a descriptor of the segmental organization for languages which exhibit [ATR] type harmonies. This model is called the Nested Subregister model, and is illustrated in (1) below as it would describe a typical nine-vowel inventory.

(1) The Nested Subregister Model

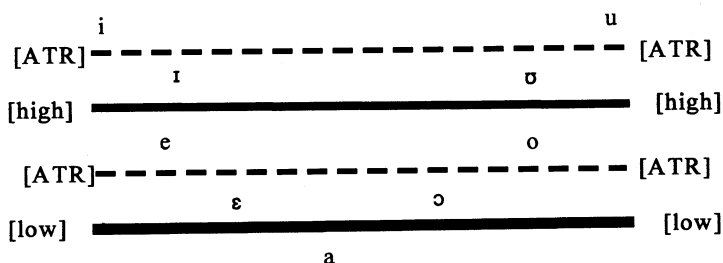


In the Nested Subregister model, the feature [openA] divides the height dimension in half. The [-openA] half is the less open, and thus the higher half, and the [+openA] half is the more open, and thus, the lower half. The feature [openB] represents a subregister, or subdivision of [openA]. As we will see in the two languages examined in this paper, the segmental makeup of the terminal nodes can vary, determined by the vowel inventory of the specific language. Following Clements (1991), a phonetic constant is that the leftmost (least open) vowels will always be the highest in the inventory, and the rightmost (most open) will be the lowest, with the remaining vowels arrayed by relative height.

The notion of an inventory-driven division of vowel features is contrary to the notions regarding distinctive features put forth in SPE. The traditional features as put forth in SPE are articulatorily based. They reference raising or lowering of the tongue body in relation to a "neutral" position defined as that in the English word 'bed' (Chomsky and Halle 1968:304). The assumption is that all languages divide the vowel space along the same parameters of musculature. It may be that the need for cooccurrence constraints and cleanup rules in so many harmony analyses stems from this assumption. In contrast, Archangeli & Pulleyblank (1994:135) cite evidence for cross-linguistic variability in the phonetic realization of F-elements. The Nested Subregister model allows for this sort of phonetic variability within the framework of a highly constrained hierarchy.

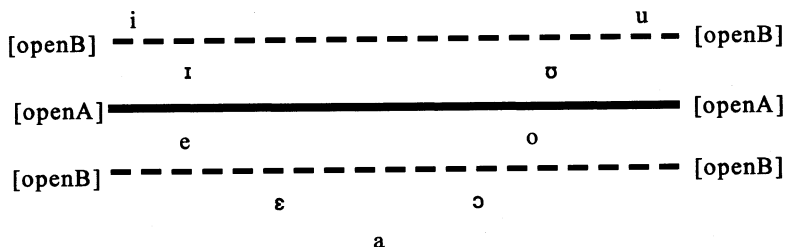
The departure from the divisions of vowel height referenced by traditional features is illustrated in Figs. (2) and (3) below.

(2) Vowel Space Divided by Traditional Features



Notice in (2) that the feature [low] separates the lowest segment from the rest of the inventory. The feature [high] then divides the remainder, with [ATR] subdividing the two divisions created by [high].

(3) Vowel Space Divided by Nested Subregister Model



In (3), as mentioned above, the feature [openA] divides the entire vowel-height space, while [openB] subdivides the two registers created by [openA]. Adopting this arrangement to account for a harmony process, one could just as easily argue for eliminating [low] from the language in question. One could then account for the divisions in (3) as spaces defined by [high] and [ATR] as they are currently understood. Likewise, one could consider /ɛ,ɔ/ to be [+low]. But, as will be seen, the segmental inventory of each subregister will vary cross-linguistically. Thus, phonetic correlates as would be referenced by traditional features cannot be viewed as fixed for the nodes in the Nested Subregister model.

Linguistic parallels for the subdivided hierarchy in the Nested Subregister model can be seen in models for tone put forth in Yip (1980), Clements (1981) and Pulleyblank (1986).

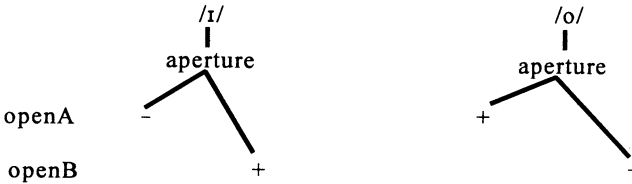
After a brief background, two languages with very different harmony systems

will be examined which are both accounted for equally well by the Nested Subregister model (hereafter NS model). Rather than a cursory look at a number of systems, this paper will examine two languages in relative detail and offer what is hopefully an explicit analysis. Further examples can be seen in Salting (1998).

2. Theoretical Background

The Aperture node and the feature [open] are adopted from Clements (1991). Also, following the notion of a feature hierarchy as put forth in Clements (1991), the nodes in the NS model are autonomous. Thus, in terms of feature spreading, segments from the inventory given in (1) above would be represented as in (4) below:

(4) Segmental Representation in Nested Subregister Model



The NS model departs from Clements (1991) on several points, however. First, the NS model employs letters (A,B) to denote successive tiers rather than numbers. Clements (1991) introduces successively numbered degrees of [open] ([open1, open2, etc.]) to account for Bantu languages which exhibit scalar harmony systems (Nzebi, Kimatumbi, i.a.). It is felt that Clements (1991) accounts very well for scalar harmony systems and these will not be addressed in this paper. The NS model is seen as a typological complement to Clements' scalar system.

For Bantu languages which exhibit non-scalar harmonies (Kinande, Sesotho), Clements (1991) introduces the notion of a branching hierarchy of the feature [open]. However, in a branching hierarchy, such as the NS model, successive tiers do not adhere to the definition of [open] as "...a uniform phonetic and phonological parameter..." (Clements 1991:38). Rather, the tiers denote subsets. Clements (1991) recognizes that the use of numbers for the hierarchic nodes is misleading, but argues for their use based on the fact that the languages in question (Bantu) are all closely related. In Salting (1998), data from two dialects of Italian (Southern Umbro and Northern Salentino) are offered to illustrate that this sort of inter-familial disparity occurs elsewhere. Therefore, in the NS model I adopt the use of a more arbitrary method of denoting hierarchic tiers, and use letters ([openA, openB]).

Second, in hierarchies of Clements (1991), the tier defined by [open1] is analogous to the feature [low], [open2] is analogous to [high], and [open3] is analogous to [ATR]. In the NS model, no reference is made to the traditional features of vowel height.

I assume several cardinal tenets of feature geometry:

- (i) "Spreading is the sole mechanism of assimilation." (Hayes 1986:467)

- (ii) The Constituent Spreading Hypothesis: "A single rule may spread no more than a single node of the feature hierarchy." (Pulleyblank 1988:314)

3. Lhasa Tibetan

Lhasa Tibetan has a 12 vowel inventory: /i,ü,u,ɪ,ɨ,ʊ,e,ø,o,ɛ,ɔ,a/. In addition to a full complement of vowels, Lhasa exhibits raising, lowering, and fronting harmony, which provide multiple triggers and targets in one language. These must, at some level, coexist. The data in this section are taken from Chang & Shefts (1964), Chang & Chang (1968), Dawson (1980) and Nornang (1978).

Raising harmony occurs across consonants and can affect any non-high vowel. This process is seen when the future suffix *-ki* is attached to verb roots:

(5) Raising Harmony

	<u>BASE</u>	<u>FUTURE</u>	<u>GLOSS</u>
(a)	še	ši-ki	'get'
(b)	kho	khu-ki	'hear'
(c)	tsö	tsü-ki	'sell'
(d)	phö	phü-ki	'flee'
(e)	nɛ	nɪ-ki	'sleep'
(f)	lo	lu-ki	'read'
(g)	ša	ši-ki	'leave'

This process also occurs across word boundaries in compound nouns. Raising harmony in Lhasa is not problematic for traditional features, and can easily be analyzed as the simple spreading of [+high] within the domain of the phonological word. But, as we will see, the other harmony processes in Lhasa prove problematic when analyzed with standard vowel-height features.

Both lowering and fronting harmonies occur in the morphophonology of Lhasa. When a root ends in a single, oral vowel, certain suffixes will create a harmonized, word-final long vowel. If the harmony environment is not met, the morpheme will surface as a default CV suffix. Some data are given below. For the sake of space, only harmonizing forms are given.

(6) Lowering Harmony

	<u>BASE</u>	<u>HARMONIZED</u>	<u>GLOSS</u>	<u>SUFFIX</u>
(a)	ci	cɪ	'heavy'	Comparative
(b)	tshüpu	tshüpu	'stitch(tailor)'	Agent/Resident
(c)	qhare	qhare	'kind(s)'	Dative/Locative
(d)	soso	soso	'oneself'	Dative/Locative

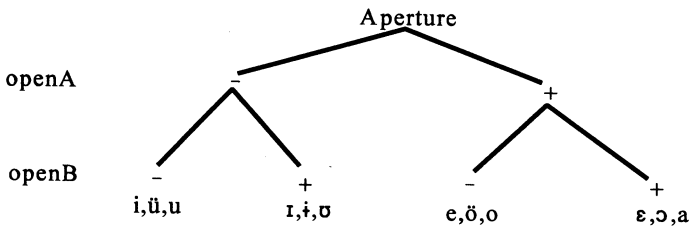
(7) **Fronting Harmony**

	<u>BASE</u>	<u>HARMONIZED</u>	<u>GLOSS</u>	<u>SUFFIX</u>
(a)	chu	chüü	'water'	Genitive
(b)	šɪŋpɪ	šɪŋpɪɪ	'farmer'	Genitive
(c)	cha	chɛɛ	'tea'	Instrumental

Salting (1995) and Salting (1998) argue that Lhasa has two allomorphs for each of the harmonizing suffixes, and that the harmonizing form consists of a mora with a single feature attached. In traditional terms, the active feature for the lowering harmony is [-ATR] and that for the fronting harmony is [-back]. For this paper, I assume the vowel place geometry of Odden (1991) which consists of a Back/Round node dominating the features [back] and [round].

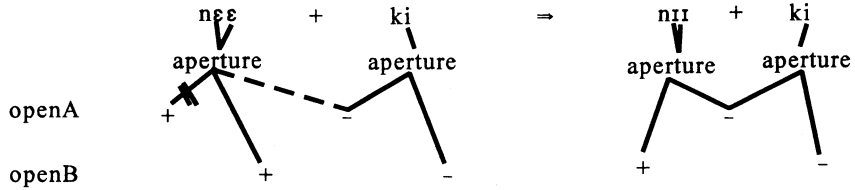
Notice in (6c) that when /e/ is lowered, the result is [ɛ], as one would expect with [-ATR] spread. Notice, however, in (7c), that when /a/ is fronted, it also surfaces as [ɛ]. Assuming the traditional features, the fronting of /a/ would create a disallowable combination of *[-back, +low], forcing a cleanup rule to create [-low]. An option would be to say that /ɛ, ɔ, a/ are all [+low] in Lhasa, but then lowering harmony to /e/ would create a surface [ɛ] specified for *[-low]. This would require an ad hoc rule to change [-low] to [+low]. In either case, traditional features cannot represent Lhasa vowels such that all harmony processes are representative of a single internal organization of vowel features.

Yet another alternative would be to eliminate the notion of the feature [low] altogether. This is accomplished with the NS model. As with the option above, this situation implies that /ɛ, ɔ, a/ are all the same height, but now they are not [+low]. Rather, they are all members of the lowest subregister. They are distinguished from each other solely by their Back/Round status. The vowels for Lhasa would be arrayed as in the hierarchy in (8):

(8) **Lhasa Vowel Hierarchy: the Nested Subregister Model**

Given the above configuration, the two height-harmonies in Lhasa can be expressed with one simple rule each, and no clean-up conditions are required. For the raising harmony, the rule is *spread [-openA]*. A derivation of the future tense for 'sleep' from (5e) is given in (9) below. For the sake of space, all irrelevant tiers are omitted from this and all subsequent derivations.

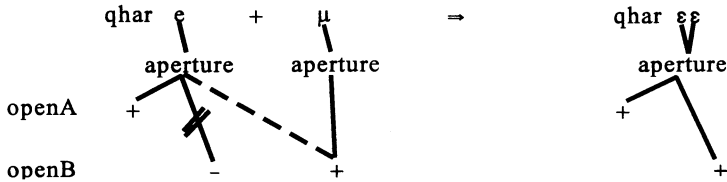
(9) Raising Harmony in Lhasa



More telling however, are the derivations in (10) and (11) in which /e/ and /a/ both surface as [ɛ] respectively due to occurrences of lowering and fronting.

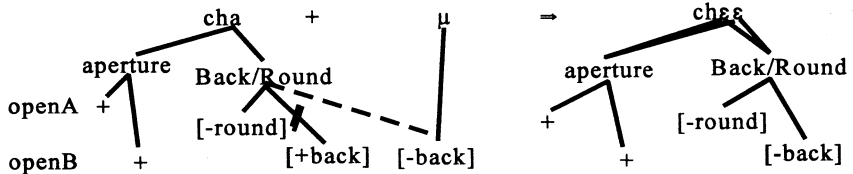
In (10) below, we see an example of the lowering harmony pattern. Following Salting (1995, 1998), the harmonizing suffix consists of a mora with a feature attached, in this case, [+openB]. The token is from (6c).

(10) Lowering Harmony in Lhasa



In (11), we see [ɛ] surfacing as a result of harmony with its other harmonic counterpart, /a/. This is due to fronting harmony. Again, we assume that the underlying trigger morpheme is a mora with a feature attached, this time, [-back]. The token in (11) is from (7c).

(11) Fronting Harmony in Lhasa



Most significant in (10) and (11) is the fact that the height features require no secondary rules or conditions to account for disallowable feature combinations. With the NS model, all harmony patterns in Lhasa fall out as natural reflections of an underlying organization within the phonological component. All segmental changes in Lhasa are shown in (12) below, with underlying and surface forms arrayed in respective order.

(12) Lhasa Harmonies

			Rule
(a)	<u>Raising</u>	e, o, ö, ε, ɔ, a → i, u, ü, I, U, ï	<i>spread [-openA]</i>
(b)	<u>Lowering</u>	i, u, e, o → I, U, ε, ɔ	<i>spread [+openB]</i>
(c)	<u>Fronting</u>	u, o, a → ü, ö, ε	<i>spread [-back]</i>

An argument could easily be made against the above analysis by claiming that either [low] doesn't exist in Lhasa, or that it remains unspecified until very late in the phonology. However, unrelated languages with differing inventories and processes provide evidence that the NS model represents a universal organizational parameter for vowel height. One such language is Ogori.

4. Ogori

Ogori is an Eastern Kwa language from Nigeria. It exhibits vowel-harmony patterns that are very difficult to analyze with traditional features, but it can be analyzed with the NS model without recourse to fill-in rules or constraints within the height component. All data are from Chumbow (1982). Ogori has a seven vowel inventory: /i, u, e, o, ε, ɔ, a/. There is a word-level constraint in Ogori barring the cooccurrence of /e, o/ with /ε, ɔ/. The vowels /i, u, a/ can occur with vowels from either of the 'mid' sets. Instances of vowel harmony in Ogori occur in the allomorphy of verb prefixes, in personal pronouns and in reduplication. For the first two cases, two distinct sets of vowels and a uniform pattern fall out. Reduplication exhibits a different pattern of vowel subsets than the other two types of harmony. I will illustrate these first two harmony processes then offer an analysis using hierarchical subsets. After that, reduplication will be examined as independent evidence for the analysis.

In the prefixal and pronominal harmonies in Ogori, two harmonic classes of vowels appear, which Chumbow (1982) labels as [+ATR] and [-ATR]. Each [+ATR] vowel is shown with its harmonic [-ATR] counterpart in (13) below:

(13) Ogori Harmonic Pairs	[+ATR]	[-ATR]
	i	ε
	u	ɔ
	e	a
	o	ɔ

Evidence for this pairing can be seen in the allomorphs for Ogori personal pronouns. These are given in (14) below:

(14) Ogori Personal Pronouns	SUBJECT		OBJECT	
	Sg.	Pl.	Sg.	Pl.
1p	ì/ě	tì/tě	mú/mó	tú/tó
2p	ù/ò	nì/ně	ú/ɔ	nú/nó
3p	è/à	bì/bě	é/á	bé/bá

When the verb contains any one of the [-ATR] vowels, the pronoun will agree in its specification for that feature as is shown in (15):

(15) Ogori Pronoun Harmony

<u>1st Person Singular</u>	<u>2nd Person Singular</u>	<u>3rd Person Singular</u>
i-jé 'I call'	ù-jé 'you call'	è-jé 'he calls'
ê-nê 'I fling'	ô-nê 'you fling'	à-nê 'he flings'
i-roró 'I think'	ù-roró 'you think'	è-roró 'he thinks'
ê-kpɔ 'I climb'	ô-kpɔ 'you climb'	à-kpɔ 'he climbs'

Evidence for the o ~ ɔ natural class is seen in (16) with the Agentive prefix:

(16) Ogori Agentives

	<u>Verb</u>	<u>Gloss</u>	<u>Agentive</u>	<u>Gloss</u>
a.	siútu	'do a job'	ò-siútu	'worker'
b.	bírépe	'plant crops'	ò-birépe	'planter'
c.	mòɔtélé	'make a pot'	ô-mòɔtélé	'potter'
d.	jájɔ	'dance'	ô-jájɔ	'dancer'

The active feature of the harmony can be seen in the Habitual prefix *dèkì-|dàkê-*. Though this prefix harmonizes, it blocks harmony to the word-initial pronoun, thus giving evidence that the default state is [+ATR].

(17) Harmony Blocking Affixes

bí dàkê bē bù úmú → (after truncation) [bí dàkê bē búmú]
they HAB beat their goat 'They always beat their goat'

With the data in (17) we can assume that, employing traditional features, the harmony rule for Ogori is *spread [-ATR]*. However, with one exception, all harmonic pairs require a variety of additional cleanup rules when analyzed with traditional features. This is illustrated in (18):

(18) Ogori Feature Changes

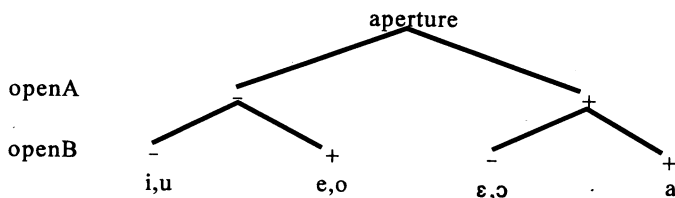
	[+ATR]	→	[-ATR]	ADDITIONAL FEATURE CHANGES
a.	i	→	ɛ	[+high] → [-high]
b.	e	→	a	[-low] → [+low]
c.	u	→	ɔ	[+high] → [-high]
d.	o	→	ɔ	no change

The most incongruous is (18b). Chumbow (1982:77) argues that it occurs "...in order to insure optimal utilization of all available vowels in the harmony process..." Were that the case, one would expect to find more cases of /e/ lowering to [a] 'over' [ɛ] cross-linguistically, but this seems to be very rare. On the other hand,

one might equally expect /o/ to lower to [a] as well, instead of lowering to [ɔ] as it does. Clearly, the traditional features of vowel height do not account for there being anything natural about the natural classes in Ogori.

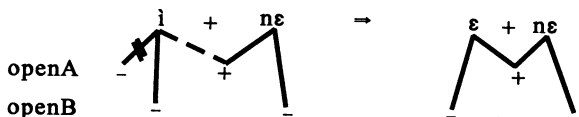
However, when analyzed with the NS model, harmony in Ogori can be explained with one rule and no conditions or constraints. As with Lhasa, the NS model divides the inventory along a center line and then subdivides each major division. Notice in (19) however, that the 'center' line as defined by [openA] separates different groups of segments for Ogori than it does for Lhasa (c.f. (8) above).

(19) Ogori Hierarchy



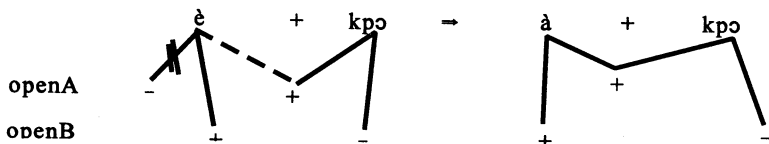
With this configuration, the harmony rule for Ogori is *spread [+openA]*. Some derivations from the data in (15) are given in (20,21) below.

(20) /i + nɛ/ → [ɛnɛ] 'I fling'



In the derivation in (21), we get a straightforward account of /e/ → [a] in Ogori harmony.

(21) /è + kpɔ/ → [àkpɔ] 'he climbs'



In both (20) and (21), the NS model is able to account for Ogori harmony in a way that does not require extra rules or constraints; all that is required is a simple rule spreading one feature. Furthermore, the NS model provides a paradigm that explains and predicts the natural classes. There is one exception, however, which does require a constraint. This occurs in /o/ → [ɔ] (Agentives) where the surface form also requires a change in the specification for [openB]. Though a change is required in the height component, it is required, not because the featural combination

is disallowed of itself, but because it cannot cooccur with the underlying Back/Round specification. Ogori has no segment which is *[+openA, +openB, +round]. Therefore, the Back/Round status of the vowel takes precedence over its height; the restriction is not against a specific height, but against a specific height-roundness combination. The issue of Place-Height interface will surface in the next section on reduplication as well.

A good corroboration for the NS model would be found in a separate harmony pattern in the same language -- a pattern which exhibits a different harmonic grouping. Such a pattern is found in Ogori reduplication. Reduplication in Ogori occurs in the nominalization of verbs and adjectives. The initial syllable of the root is reduplicated and the nominalizing prefix ò/ô which we saw in the Agentives (see (16) above) is added.

(22) Ogori Reduplication

A. VERBAL NOUNS

	<u>Root</u>	<u>Verbal Noun</u>	<u>Gloss</u>
1.	džé	ò-dží-dze	'eating'
2.	rúwá	ò-rù-rúwá	'division'
3.	wá	ò-wé-wa	'drinking'
4.	šá	ò-šé-ša	'coming'
5.	bálé	ò-bé-balé	'watching'

B. ADJECTIVAL NOUNS

	<u>Root</u>	<u>Noun</u>	<u>Gloss</u>
1.	gbódi	ò-gbí-gbodi	'bigness'
2.	kéré	ò-kí-kere	'smallness'
3.	lóré	ò-lé-lóre	'longness'
4.	dó	ò-dè-dó	'oldness'
5.	béré	ò-bé-bére	'softness'

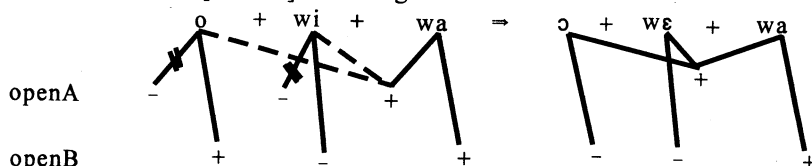
With the exception of *rúwá*, all reduplicative affixes contain either [i] or [ɛ]. Within the framework of the NS model, this means that the reduplicative affix is an open syllable containing a front vowel specified [-openB] which gets its [openA] status from the root vowel. The harmonic pairings of Ogori reduplication are presented in (23) below:

(23) Ogori Reduplication Schematized

	<u>RED. AFFIX</u>	<u>ROOT VOWEL</u>
a.	u	u
b.	i	e
c.	i	o
d.	ɛ	a
e.	ɛ	ɛ
f.	ɛ	ɔ

As with the previous harmony processes, one can assume a default status of [-openA] for the reduplicative affix making the underlying vowel /i/. With this, the rule for reduplication is the same as for the other processes, namely *spread* [+openA]. A sample derivation is given in (24) below. The subsequent change of [openB] for the nominalizing prefix is as per the discussion above.

(24) /o-wi-wa/ ⇒ [ɔwɛwa] 'drinking'



The incongruity of *rúwá* occurs because both Place and Height spread from the root. Space does not allow for a detailed explication of this issue, but this Place-Height interface parallels similar anomalies in Bantu languages (Luganda, Chichewa, Chithumbuka i.a.) and in Turkish. This issue deserves very involved research of its own, and is addressed in Salting (1998:143).

To reiterate, the NS model captures harmony processes in Ogori without the need for cleanup rules or constraints within the height component of the language. Rather, the NS model accurately describes the natural classes as they are delineated in Ogori harmony.

5. Conclusion

For the most part, distinctive features have been assumed to reference fixed phonetic correlates. The Nested Subregister model provides evidence for a linguistic universal in the organizational paradigm rather than in the phonetic referent of universal distinctive features. We have seen the power of the NS model to capture natural classes of segments as they manifest in harmony patterns in two languages, Lhasa and Ogori. Salting (1998) uses the Nested Subregister model to account for harmony in Kõnni and Somali as well. These are both languages which have proven very problematic to analysis with other models, but which were explained in a straightforward manner when the NS model was applied. Cursory looks at other [ATR] systems such as Yoruba, Akan and Okpe have shown great promise for the NS model as well.

The implication is that the symmetrical, hierarchic division of the vowel height space is a universal, adoptable by some languages. Other languages may employ a gradient division of the vowel height space such as that in Clements (1991), producing scalar harmony patterns. Salting (1998) shows that closely related languages can vary as to which type of organization they choose, as is seen in Bantu languages and dialects of Italian. A prediction is that one would never find scalar harmony in a language which also exhibits evidence for a hierarchical division, and *vice versa*.

OT works in vowel harmony (Pulleyblank 1994) still refer to the traditional features and thus, still require the same sorts of constraints for individual harmony processes. Within the framework of the NS model, there is no gradient violability within the height component. Rather, harmony is seen as a reflection of an internal organization of the inventory. OT may be of benefit in accounting for Height - Place

interactions as we saw in Ogori, but this will have to wait on further research.

Perhaps phonological distinctive features (for vowel height at least) are further removed from phonetic correlates than has been assumed, and instead reference a small variety of fixed parameters of organization. The fact that the NS model can so cleanly account for harmony systems which prove very problematic for other systems hints at the need for rethinking our approach to distinctive features. Further research will tell.

REFERENCES

- Archangeli, Diana and Douglas Pulleyblank. 1994. *Grounded Phonology*. MIT Press.
- Chang, Kun and Betty Shefts. 1964. *A Manual of Spoken Tibetan*. University of Washington Press. Seattle.
- Chang, Kun and Betty Shefts Chang. 1968. Vowel Harmony in Spoken Lhasa Tibetan. *Bulletin of the Institute of History and Philology Academia Sinica*. 40:53-124.
- Chomsky, Noam & Morris Halle. 1968. *The Sound Pattern of English*. Harper & Row. New York.
- Chumbow, Beban S. 1982. Ogori Vowel Harmony: An Autosegmental Perspective. *Linguistic Analysis* 10:1, 61-93.
- Clements, George N. 1981. *The Hierarchical Representation of Tone Features*. Harvard Studies in Phonology. Vol. II. IULC.
- Clements, George N. 1991. Vowel Height Assimilation in Bantu Languages. *Working Papers of Cornell Phonetics Laboratory*. 5:37-74.
- Dawson, Willa. 1980. *Tibetan Phonology*. Ph.D. Dissertation. University of Washington.
- Hayes, Bruce. 1986. Assimilation as Spreading in Toba Batak. *Linguistic Inquiry* v.17: 467-499.
- Nornang, Willa. 1978. Tibetan Vowel Process: A Case for the Feature, Advanced Tongue Root. D. Malsch, J. Heard and C. Sleat (eds.) *Proceedings of the 8th Annual Meeting of the Western Conference on Linguistics*.
- Odden, David. 1991. Vowel Geometry. *Phonology* 8:261-289.
- Pulleyblank, Douglas. 1986. *Tone in Lexical Phonology*. D. Reidel. Dordrecht.
- Pulleyblank, Douglas. 1988. Underspecification, the Feature Hierarchy and Tiv Vowels. *Phonology* 5:299-326.
- Pulleyblank, Douglas. 1994. Vowel Harmony and Optimality Theory. To appear in *Proceedings of Workshop on Phonology*. University of Coimbra, Portugal. September 1993.
- Salting, Don. 1995. Vowel Harmony in Lhasa Tibetan: Evidence for [hi] and [ATR]. *FLSM VI*, 1:39-49.
- Salting, Don. 1998. *The Nested Subregister Model of Vowel Height*. Ph.D. Dissertation. Indiana University.
- Yip, Moira. 1980. *The Tonal Phonology of Chinese*. Ph.D. Dissertation. MIT.

Phonological Universals: Trilling, Voicing, and Frication

Maria-Josep Solé
Universitat Autònoma de Barcelona, Spain

1. Introduction. This paper addresses the issue of why certain combinations of features, phonological patterns, and segment contrasts, are statistically preferred over others in the languages of the world. The principles that determine such universal preferences have been recognized as deriving from the physical and auditory properties of speech features (Ohala 1974, 1983; Lindblom 1983, 1990; Westbury and Keating 1986). In this view, the sound patterns which are more likely to be used in the world's languages are those reflecting the physical constraints of the speech production mechanism (and, consequently, not involving extra adjustments and increased articulatory cost), and having efficient acoustic consequences (*i.e.*, resulting in auditorily salient and distinct signals); or those representing an optimal balance between competing demands of perception and articulation.

The constraints of the speech production and perception mechanisms underlie the notions of 'gestural economy' and 'auditory distinctiveness' which define the combination of features into possible speech segments and their likelihood. For example, the feature values [+nasal] and [+fricative] do not co-occur due to *physical* constraints; a sufficient velo-pharyngeal port opening to create the percept of nasality would bleed the volume velocity of airflow required to create friction (Ohala, Solé and Ying 1997). Poor *auditory* result (as measured from confusion coefficients or dephonologization) can be evoked to account for the relatively low incidence of voiceless nasals, where the amplitude modulation for nasals is impaired by voicelessness. Conversely, some combination of features enhance the acoustic/auditory image, *e.g.*, the combined action of [+back] and [+rounded] heighten F2 lowering (Perkell *et al.* 1993), which is in line with the universal preference to round back vowels.

In this paper, we address the co-occurrence of tongue tip trilling with voicing, frication and nasalization. Exploring the physical and auditory properties of trills will allow us to account for their possibilities of combination with other features. The present study attempts (1) to characterize the aerodynamic conditions required for the production of lingual trills, specifically, the range of allowable variation in pharyngeal pressure to initiate and sustain trilling, and (2) to account for the combination of trilling with other speech events in terms of its aerodynamic and perceptual requirements. We address the questions of whether it is more 'natural' for trills to be voiced, why trilling tends to alternate with frication across languages, and why there is a lack of nasal trills.

This study is in line with the notion that in the search for universal patterns it is necessary to characterize a set of parameters -- physiological, aerodynamic, and acoustic/auditory -- their range of variation, and a set of categorial values along these parameters which exhibit stable relations in the articulatory and auditory domains. These categories or 'optimal settings' are the more likely seat of segments which are common cross-linguistically (Stevens 1989, Lindblom 1986, Ohala 1983). Some of the gradient physiological and aerodynamic variation will tend to result in discrete changes along these parameters due to the quantal nature of speech (Stevens 1972, 1989); thus, the articulatory robustness of speech features, specifically trilling, under different contextual and prosodic conditions can be characterized as well as their auditory distinctiveness.

Apical trills, such as [r], are sounds which are mastered late in the acquisition process (in fact, they are, along with sibilants, the last segment types to be mastered, Jiménez 1987, Vihman 1996), are not present in the babbling or 'vocal play' stage in infant vocal production when they are exploring the possibilities of the vocal tract (Stark 1980)¹, and present difficulties to second language learners and also to native speakers, *i.e.*, some speakers never succeed in rolling their [r]s, which suggests that they are sounds involving a complex production mechanism requiring positioning of the articulators, stiffness conditions, and aerodynamic requirements. Yet they are not uncommon sounds in phonological systems, half (47.5%) of the r-sounds in the languages of the world are trills, overwhelmingly dental/alveolar trills (Maddieson 1984). Ruhlen (1975) reports lingual trills in 79.5% of languages with an r-sound (78.3% of the sample).

The mechanics of tongue-tip vibration have been described by Catford (1977), Ladefoged and Maddieson (1996), Spajic, Ladefoged and Bhaskararao (1996), Barry (forthcoming), and modelled by McGowan (1992). These authors describe trills as the vibration of certain supralaryngeal articulators (tongue tip, uvula, lips) caused by *aerodynamic forces*, as opposed to taps and flaps, which involve active muscular movements of the tongue. The conditions for initiating lingual trilling involve (i) muscle contraction of the tongue to assume the position, shape and elasticity requirements, and (ii) a sufficient pressure difference across the lingual constriction. Once trilling is initiated, tongue-tip vibration is maintained as a self-sustaining vibratory system. Articulatorily, trills exhibit more predorsum lowering and postdorsum retraction than taps, thus leaving more room for the vertical movements of the tongue tip and blade, and a more retracted alveolar closure. In addition, the tongue body is more highly constrained for the trill than for the tap and the former coarticulates less with neighboring vowels (Recasens and Pallarès, *in press*). Unfortunately, the aerodynamic forces in trills have received little attention (but see McGowan 1992). Understanding the trade-offs between articulator movements and aerodynamic forces, and their acoustic result is essential for accounting for the phonological behavior of trills, for speech pathology and articulatory modelling and synthesis.

Every one of the requirements of positioning, shape, articulator mass, stiffness and pressure difference are necessary for trill production². Thus, lingual trills require fine neuromotoric adjustment of these different parameters, which accounts for their intrinsic difficulty in inexperienced (or immature, in the case of infants) speakers. Trills are very sensitive to variations in the articulatory and aerodynamic conditions, which may result in lack of tongue tip vibration. Thus, it is common that trills are realized as non-trilled variants (*e.g.*, in Spanish, Italian, Toda). In addition, trills alternate historically, dialectally and allophonically with taps, approximants and fricatives.

The main aim of this paper is to characterize the aerodynamic requirements for trills and their range of variation, and how they account for some common patterns in trills. In section 2 we present an experiment in which the pharyngeal pressure during trills was varied and the associated effects on the production and acoustic properties of trills were observed. In section 3 we provide an account for some phonological universals in the patterning of trills: the preference for voiced trills, the alternation between trills and fricatives, and the lack of nasal trills. Section 4 presents the conclusions.

2. Aerodynamic Characteristics of Trills. Aerodynamic conditions play a critical role in the production of trills, and their pathologically (*e.g.*, cleft palate),

contextually (coarticulation, speaking rate, etc), or artificially (e.g., experiments) induced modification may seriously affect the production of these sounds. The experiment reported here was designed to provide information on the allowable range of aerodynamic variation and compensatory articulatory maneuvers to sustain voiced and voiceless trilling, and their acoustic result. The results may throw light on the role played by aerodynamic factors in some universal tendencies in the patterning of trills.

2.1. Experimental Method. In order to identify the aerodynamic conditions required for trilling, intraoral pressure (P_o) and airflow were recorded simultaneously in two trained phoneticians producing steady state and intervocalic voiced and voiceless alveolar trills, as well as sustained trills with maximum exhalatory effort. Uvular trills and taps were also recorded for comparison. P_o was sampled via a catheter inserted into the pharynx through the nose and connected to a pressure transducer. Airflow was collected with a Rothenberg mask and a pneumotachograph. The oral pressure and airflow signals were low-pass filtered at 50 Hz. The aerodynamic and acoustic data were digitized and sampled at 16 kHz. Oral pressure during trills and taps was intermittently bled with catheters of varying cross-sectional areas (7.9, 17.8, 31.6 and 49.5 mm²), all 25 cm long, inserted into the speaker's mouth via the buccal sulcus and the gap behind the back molars. The impedance (i.e., resistance to exiting air) of the catheters for the range of flow used in voiced and voiceless trills was calculated, as well as the vocal tract impedance during the production of trills for each speaker. The catheters venting the P_o were intended to simulate variations in oral pressure present in speech due to contextual and prosodic factors, e.g., coarticulation with sounds of varying impedance, adjacent nasals, stress, speaking rate, phrasal position, etc. P_o and airflow were measured for the different conditions and the variation, impairment or extinction of trilling as a function of varying intraoral pressure was analyzed acoustically.

Masking noise was placed on the speaker through earphones to minimize auditory feedback. Kinaesthetic feedback could not be eliminated and measurements were made during the first 60ms after P_o was varied.

2.2. Results

2.2.1. Aerodynamic Features of Voiced and Voiceless Trills. Voiceless trills differ from voiced trills in the nature and rate of airflow which the tongue tip vibration modulates. While in voiceless trills the vibrating tongue-tip modulates a large and continuous airflow -- slightly turbulent due to impedance at the glottis -- voiced trills modulate periodic vibrations in airflow and pressure produced by the vibrating vocal folds. The laryngeal vibrations reduce the amount of air flowing into the oral cavity.

The higher rate of transglottal flow through the open glottis for voiceless trills, vis-à-vis voiced trills, resulted in the following differences: (1) Voiceless trills show a *higher* P_o than voiced trills as illustrated in Fig. 1, which shows peak intra-oral pressure for steady state and intervocalic voiced and voiceless trills for the two speakers (the difference in absolute values in the two speakers may stem from the net differences in the overall size of the vocal tract). Since there is no other narrow constriction downstream, the pressure difference across the lingual constriction for trills is equal to P_o ($\Delta P_{\text{supraglott}} = P_o - P_a$, where atmospheric pressure, $P_a = 0$). The pressure drop between the oral cavity and the atmosphere, which is larger for voiceless than for voiced trills, determines a *larger rate of flow* across the lingual

constriction (0.66 and 1.34 lit/sec for voiceless trills vs 0.22 and 0.56 lit/sec for voiced trills for speakers MJ and JJ, respectively). The larger flowrate results in a higher particle velocity and the generation of turbulence or friction noise across the lingual constriction for voiceless trills. Fig. 1 also shows that the P_o required to sustain tongue tip vibration is lower than that required to initiate it. Coarticulatory effects can also be observed; trills in the /i/ context exhibit a higher P_o (due to a smaller cavity volume for the vowel) than in the /a/ context.

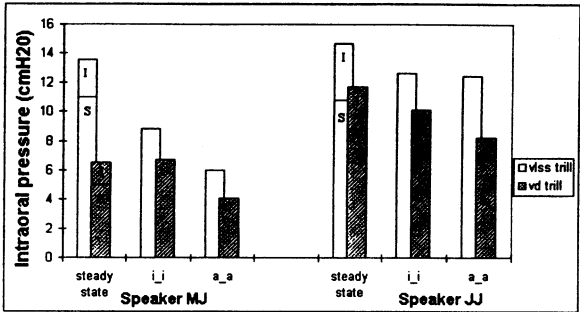


FIGURE 1. Mean peak oral pressure (ΔP across lingual constriction), in cmH2O, for voiced (grey bars) and voiceless (white bars) steady state and intervocalic trills for speakers MJ and JJ. In the steady state trills the difference in P_o needed to initiate (I) and sustain (S) a trill is indicated. Each bar represents an average of approximately 30 observations.

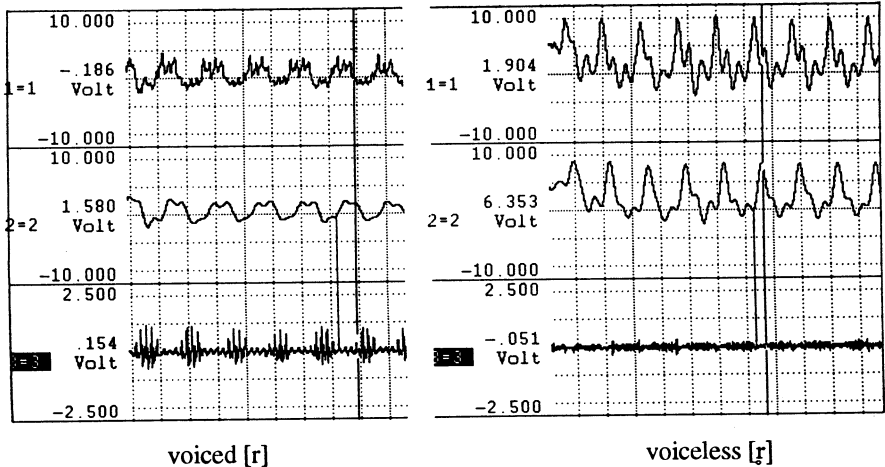


FIGURE 2. Unfiltered P_o (channel 1), low-pass filtered P_o (channel 2) and audio signal (channel 3), in volts, for sustained voiced and voiceless trills. The tongue tip contact period, showing a rise in P_o , is indicated between lines in the filtered P_o trace (a small phase shift is present between the filtered and unfiltered P_o traces). Voiceless trills exhibit a shorter closure period than voiced trills.

(2) A larger open to closed period ratio in tongue tip vibration was found for voiceless as opposed to voiced trills (1.96 vs 1.29). Fig. 2 illustrates the shorter closure period (indicated by a rising P_o) for voiceless than for voiced trills -- most

likely due to the higher oral pressure which overcomes the resistance of the tongue-tip in a shorter amount of time -- whereas the opening period is comparable in both. The proportionally longer open period in voiceless trills results in a lower impedance at the lingual constriction and longer periods of released (turbulent) energy, vis-à-vis voiced trills. As shown in the waveform in Fig. 2, voiceless trills also show frication during the closure, reflecting an imperfect palato-lingual closure.

(3) Voiceless trills exhibit a slightly *higher rate of vibration* than voiced trills (29.3 Hz, range 28-31.5 Hz vs 28.1 Hz, range 26-29 Hz) due to the larger pressure drop across the lingual constriction for the former. These results contrast with Lindau's (1985:161) finding that voiceless trills in Edo show a slower rate of vibration than voiced trills (22.5 Hz vs 25 Hz).

Voiced trills involve very precise aerodynamic conditions in order to maintain trilling and voicing. Po needs to be high enough to produce tongue tip vibration and low enough not to impair the transglottal flow required for voicing. Thus, the range of Po variation is bounded by the requirements for translungal flow needed for trilling and those for transglottal flow needed for voicing.

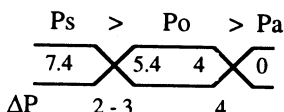


FIGURE 3. Estimated Po range for voiced trills for speaker MJ.

A pressure drop across the oral constriction of at least 4 cmH2O (depending upon tongue tip tension and mass) was found to be needed to sustain trilling, and the liminal volume velocity was about 0.175 lit/sec. Transglottal flow for voicing requires a minimum pressure drop of 2-3 cmH2O, and a minimum volume velocity of 0.050 lit/sec (Catford 1977: 98). To estimate the allowable range of Po variation, the peak intraoral pressure during voiceless trills was used as an estimate of subglottal pressure (Ps). Subject MJ had a peak, sustained Po of approximately 7.4 cmH2O for intervocalic voiceless trills (vowel contexts pooled), which can be assumed to be the value for Ps. If transglottal flow requires a pressure drop across the glottis (Ps-Po) of at least 2-3 cmH2O, that leaves a Po of at most 5.4 cmH2O, as schematically shown in Fig. 3. Trilling requires a minimum pressure drop of 4 cmH2O across the oral constriction, which means that Po may vary between a rather narrow range of 5.4 - 4 cmH2O in order to sustain voicing and trilling (the actual average Po value for speaker MJ was 5.39 cmH2O). The estimation of the range of Po variation for speaker JJ was between 8 - 11 cmH2O. Thus, the Po range for voiced trills is very narrow and unforgiving, and small pressure variations may lead to devoicing or cessation of trilling (a similar argument has been made by Ohala (1983) to account for the difficulty in maintaining voicing and frication). Ladefoged and Maddieson (1996:221) suggest that the rapid *increases* in Po during tongue tip contact in trills may impair the sufficient pressure drop for voicing, and be responsible for the tendency to devoice trills during the closure interval in a number of languages. Similarly, *decreased* Po may lead to devoicing: A lowered Po (due to decreasing subglottal pressure in utterance final position) endangers trilling, and active abduction of the vocal folds may be present to directly access the air reservoir and thus preserve trilling (see section 2.2.3), resulting in the commonly observed tendency to devoice final trills, *e.g.*, South American Spanish (Quilis 1981, Canfield 1981), Farsi (Ladefoged and Maddieson 1996). As for the cessation of trilling, non-

trilled allophones of trills have been reported in a number of languages (see section 3.2).

2.2.2. Variations in Oral Pressure. In order to determine the range of allowable variation in intraoral pressure in the production of trills, the backpressure during trills was intermittently bled with catheters of different cross-sectional area, and the articulatory and acoustic consequences were analyzed. Fig. 4 illustrates the effects of the reduction in P_o associated with a 17.8 mm^2 vent. Sustained voiced trilling is extinguished on vent, resulting in a fricative, whereas voiceless trilling is maintained, with peak oral P_o noticeably reduced and the amplitude of the acoustic energy diminished as can be observed on the waveform. Thus, for a given vent trilling was extinguished earlier in voiced than in voiceless trills. Tongue tip vibration was extinguished into a fricative/approximant as the reduction in P_o diminished the rate of flow through the oral constriction and the magnitude of the Bernoulli effect, which was not sufficient to suck back the tongue tip to the contact position. Fig. 4 shows that blocking the tube restores the original pressure conditions and associated acoustic effects in voiceless trills, but trilling fails to reinitiate in voiced segments. Fig. 5 displays an intervocalic voiced trill in normal pressure conditions (right), and with a vent area of 7.9 mm^2 (left), where the reduced P_o impairs tongue tip vibration, resulting in a fricative. Intervocalic trills proved to be more resistant to pressure changes between open vowels than between high front vowels. The lesser robustness of trills in the */i/* context is due to increased tension and mass of the tongue tip due to coarticulation with the high front vowel, rather than to intrinsic differences in the P_o of trills in different vowel contexts.

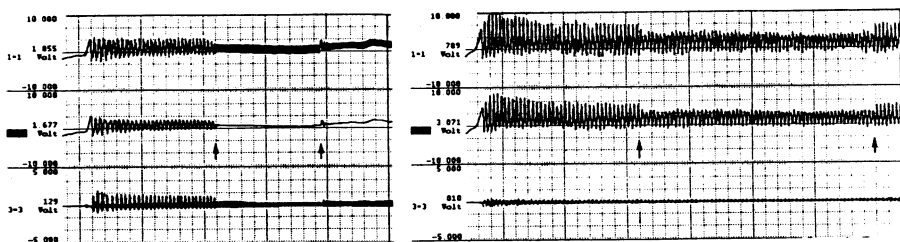


FIGURE 4. (1) Unfiltered P_o , (2) filtered P_o and (3) audio signal during the production of steady state voiced (left) and voiceless (right) trills. The unblocking and blocking of the catheter (area 17.8 mm^2) is indicated by arrows.

The reduction in oral pressure associated with venting the backpressure with catheters of different areas is presented in Fig. 6. When the P_o dropped below a certain threshold, trills were extinguished resulting in a fricative (a non-sibilant voiced alveolar fricative in the case of voiced trills, and a [h] sound for voiceless trills) or an approximant with catheters of larger areas ($\geq 31.6 \text{ mm}^2$). It is not possible to report absolute pressure values at which trilling was impaired, since depending on initial conditions (articulator tension, mass, cavity volume, articulatory position, compliance of the supraglottal walls, etc.) and speaker, the minimum oral pressure required for tongue tip vibration varied. A pressure drop of 2.5-3.5 cmH₂O impaired sustained trilling in voiced segments. A larger pressure drop, 5 cmH₂O,

was needed to impair voiceless trills. Thus, voiceless trills are more resistant to variations in P_o due to (1) a higher P_o which allows a larger reduction in pressure before reaching the minimum pressure drop across the lingual constriction required for trilling, (2) direct access to P_s to replace vented airflow, and (3) a smaller P_o reduction for the same vent aperture (because impedance is higher at higher flow rates).

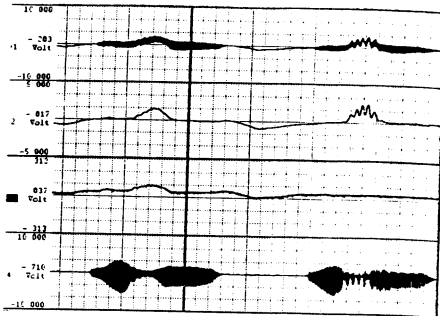


FIGURE 5. (1) Unfiltered and (2) filtered P_o , (3) airflow escaping through the catheter when unblocked, (4) audio signal for vented and unvented [i'ri].

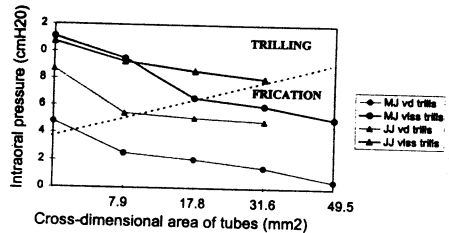


FIGURE 6. Reduction in intraoral pressure (y axis) per tube cross-sectional area (x axis) for steady state voiced (grey line) and voiceless (black line) trills produced by speakers MJ and JJ. The threshold of trilling is indicated by a hatched line. Each value represents an average of between 3 and 8 observations (mode 5). The number of vented tokens in each condition varied since intermittent venting had a random pattern.

2.2.3. Trill Extinction and Reinitiation. It was found that once trilling was extinguished, in the majority of cases it did not reinitiate when the initial pressure conditions were restored (by blocking the catheter, as shown in Fig. 4), suggesting that the pressure difference required to initiate tongue tip vibration is higher than that required to sustain it, in accordance with our measurements (Fig. 1) and with the reported differences in pressure drops necessary to initiate and maintain vocal fold vibration (Westbury et al. 1986). Alternatively, failure to reinitiate trilling may reflect differences in the initial tongue tip positioning for trills and fricatives.

The role played by initial pressure and articulatory conditions in trill production is illustrated in Fig. 7, which exhibits the initiation of trilling on vent (a), and the failure to sustain trilling in comparable pressure conditions (c) or to reinitiate trilling with increased P_o (d). The physiological adjustments (in cavity volume, mass and/or tension of the vibrating articulator, vocal tract compliance) to compensate for initial pressure conditions in (a), are most likely relaxed when P_o is reestablished in (b). When P_o is further reduced in (c), trilling cannot be sustained with the existing articulatory conditions. Similarly, trilling cannot be reinitiated when P_o is reestablished in (d). This illustrates once more the finely tuned aerodynamic and articulatory requirements to initiate and sustain trilling.

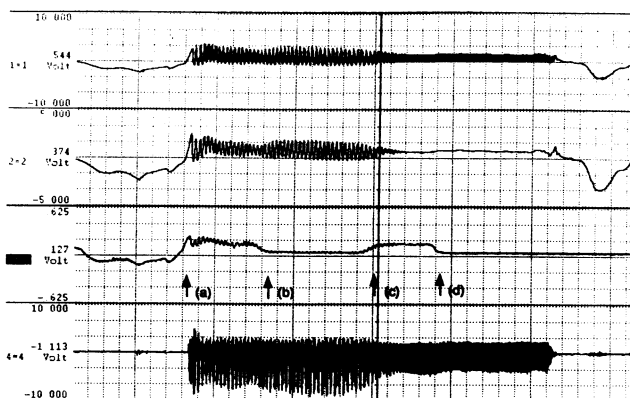


FIGURE 7. (1) Unfiltered P_o , (2) filtered P_o , (3) catheter vent, and (4) audio signal for a sustained voiced trill. (a) Trilling is initiated on vent (vent area 17.8 mm^2); (b) the catheter is blocked, P_o and amplitude of the signal increase; (c) extinction of trilling on vent; (d) failure to reinitiate trilling when the pressure conditions are reestablished by blocking the catheter.

In the cases when lingual vibration did reinitiate (usually when the venting period was very short) it generally did so through a transitional fricative, most likely reflecting the increase in P_o and in volume velocity before the Bernoulli force closed the alveolar channel. The extinction (and reinitiation) of trills into a fricative suggests that the aerodynamic range of variation for trills is narrower than for fricatives.

In prolonged trills, tongue tip vibration was sustained as long as sufficient airflow was available. When P_s (and consequently P_o) diminished thus endangering trilling, two possible acoustic outcomes were observed: frication and/or devoicing. In the great majority of cases the trill decayed into a fricative as airflow through the lingual constriction dropped due to diminished P_s . In a few cases, the trill or the resulting fricative were further devoiced. Devoicing can be seen as a maneuver to directly access P_s by removing the resistance at the glottis and thus to prolong trilling.

In opposition to trills, which extinguish when oral pressure is vented by approximately $2.5\text{--}3.5 \text{ cmH}_2\text{O}$, taps continue to exist on vent, which illustrates that the two sound types involve different primary energy forces: aerodynamic vs. muscular.

The behavior of trills in varying aerodynamic conditions parallels common processes and alternations found in languages (e.g., trill devoicing, detrilling, trill frication) and can account for observed phonological patterns.

3. Phonological Patterns.

3.1. Preference for Voiced Trills. We now address the issue of trilling co-occurring with voicing almost exclusively. The statistical preference for voiced over voiceless trills is evidenced in phonological inventories and diachronic variation. Trills are mostly voiced in the languages of the world; only 1.5% of the trills are voiceless (Maddieson 1994, Ruhlen 1975). This is a lower percentage than for other voiceless sonorants (3.17% for nasals, approximants and liquids).

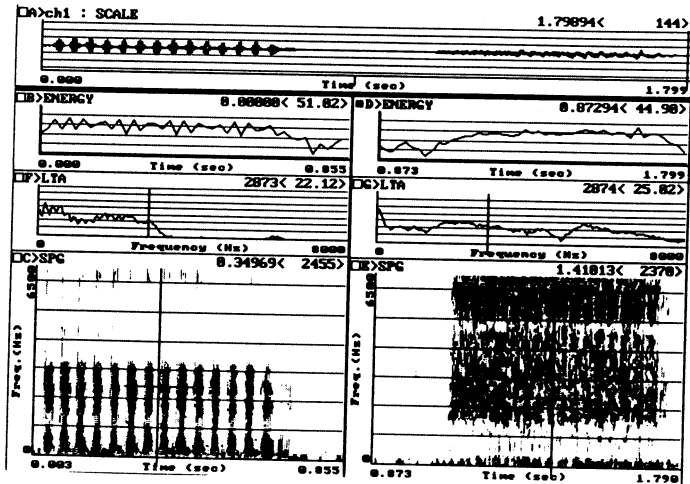


FIGURE 8. Waveform, energy plot, long term average spectrum and spectrogram of steady state voiced (left) and voiceless (right) trills produced by speaker MJ.

Fig. 9 shows contrastive voiced and voiceless trills for Lai-chin and Icelandic (the spectrograms show from 0 to 9 kHz). In voiceless trills the continuous high frequency component dominates the spectrum. Voiced trills show one clear contact followed by a burst of energy and subsequent decreases in intensity due to the vibration of the tongue tip approaching a closure.

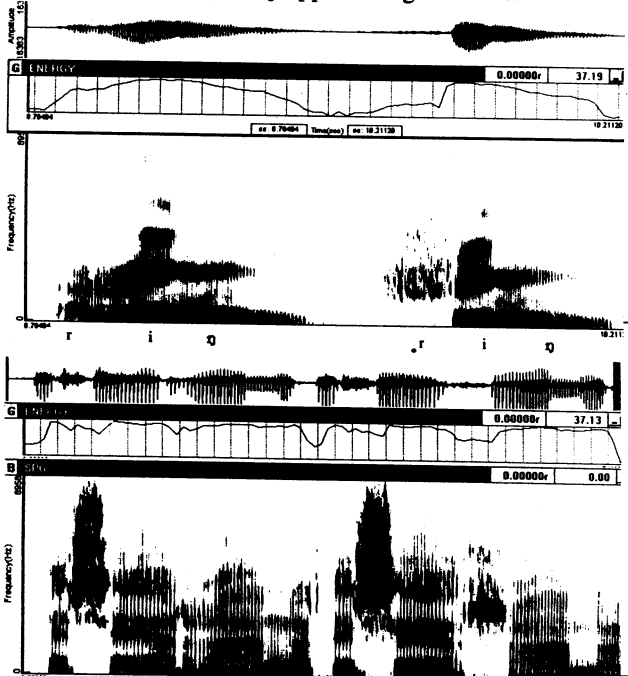


FIGURE 9. Voiced and voiceless trills in Lai-chin, [ri] 'loud' and [hri] 'green' (top), and Icelandic ([Teir 'seija 'ra:nar] 'they say trunks', [Teir 'seija 'hra:nar] 'they say boors' (bottom). Waveform, energy plot and spectrogram (0-9KHz).

The observed aerodynamic and articulatory characteristics of voiceless trills, vis-à-vis voiced trills (glottal friction; higher Po and larger flowrate across the oral constriction; larger ratio open to closed period of tongue tip vibration; and failure to achieve full palato-lingual closure) contribute to turbulent energy throughout the sound which makes them auditorily fricative-like. This is in line with Ohala's (1997) observation that voiceless sonorants tend to become fricatives. The reduction in airflow associated with voiced trills contributes to a regular alternation in the spectrum of bursts of energy; such temporal-spectral discontinuities result in an auditorily distinct signal. The co-occurrence of trilling and voicing can be seen as a natural byproduct of acoustic and auditory salience.

The preference for voiced trills in phonological systems seems to reflect a trade-off between articulatory stability (*i.e.*, preserving trilling in a narrow range of aerodynamic conditions, as opposed to voiceless trills) and acoustic/auditory saliency (*i.e.*, distinct signal modulation).

3.2. Trilling and Frication. A common cross-linguistic pattern is the alternation and co-occurrence of trilling and frication. Fricatives and trills tend to *alternate* synchronically and diachronically. Synchronically, apical trills exhibit non-trilled variants, taps, approximants, and fricatives. Fricative (and approximant) allophones of trills result when the vibrating tongue-tip fails to make contact with the palate, or apical vibration fails to occur. Non-trilled fricative variants have been reported in continental Spanish (Navarro Tomás 1950: 117) and American Spanish (Zlotchew 1971), Toda (Spajic, Ladefoged and Bhaskararao (1996:8), and Standard Swedish (Lindau 1985:164). This is most common in fast speech, and in the environment of high front vowels, where time constraints and muscular contraction for adjacent segments affect the tension of the tongue tip.

Historically, trills developed into fricatives in Tai dialects when voiceless (section 3.1) and devoiced, *e.g.*, *pr, *tr, *kr > ph, (t)h, kh, in Central dialects (Li 1977: 86, 118, 225). Other examples are the spirantization of palatalized trills in dialectal variants of Irish, *e.g.*, *ma:rla > [máza] 'Maire' or Proto-Slavic, where a palatalized trill developed into a trilled fricative [ř] in Czech, and a palatoveolar fricative [ʒ] in Polish (*e.g.*, Proto-Slavic *tsarja > Czech [tsařa] Polish [tsaʒa], 'Czar'). Palatalization involves active raising of the tongue dorsum which is antagonistic with the tongue shape (concave) and tension (relaxed tongue tip) required for trills (Recasens and Pallarès, in press). In addition, palatalization involves a more massive tongue tip and blade offering a higher resistance to trilling (Kavitskaya 1997).

Trilling tends to *co-occur* with frication. Trills with associated frication are most commonly uvular trills (in Southern Swedish, Standard French and Standard German (Ladefoged et al. 1996, Lindau 1985)), but apical trills involving frication have been reported in Toda and Spanish. Spirant trills [ř] (or assimilated /r/s) are the norm in approximately half of the American Spanish dialects and some dialects of Peninsular Spanish. Approximately half of these spirant trills are devoiced (Quilis 1981:301, Canfield 1981:7). Thus, *la ropa* 'the clothes' and *la sopa* 'the soup' are often confused by outsiders. Utterance and syllable final trills are assimilated in virtually all the American Spanish territory. Fig. 10 illustrates assimilation and devoicing of prepausal /r/ for a Mexican speaker. The intervocalic trill also shows some fricative component. Fricative trills result from imperfect linguopalatal closure or partial vibration of the tongue tip, which allows the high velocity air to flow continually through the aperture generating friction. The concomitance of trill devoicing and frication has been referred to in the previous section.

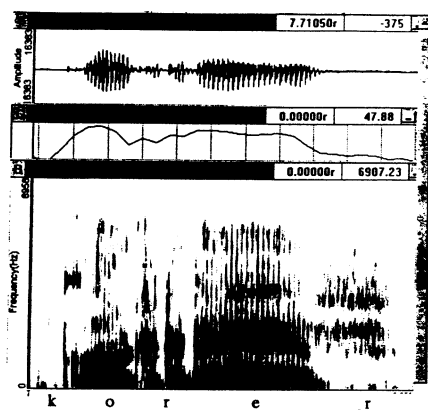


FIGURE 10. Waveform, energy plot and spectrogram of the word *correr* 'to run' uttered by a Mexican speaker. The intervocalic trill, showing three linguo-palatal contacts, is voiced with some frication, and the prepausal trill is aspirated and devoiced during most of its duration.

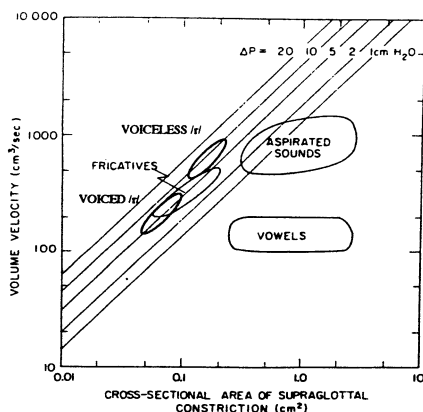


FIGURE 11. Pressure and airflow values for voiced and voiceless trills have been added to Steven's (1971) chart.

A related phenomenon is the *trilling of fricatives* (velar, uvular, and pharyngeal) in Northern Dutch (van den Berg 1958), dialectal German and French, (Ladefoged et al. 1996), !Xóó (Traill 1985), and peninsular Spanish (Quilis 1981). Occasionally presence of uvular/epiglottal trills reinforcing [-anterior] fricatives is found in these languages (mainly in contact with low back vowels).

Our experimental conditions replicated the phonological variation between voiced trills and fricatives. The results in sections 2.2.2 and 2.2.3 show that (i) when pharyngeal and subglottal pressure were reduced below a certain threshold trilling was extinguished into a fricative (or an approximant with larger pressure drops); (ii) trills may reinitiate as a fricative sound, and (iii) 'failing' trills -- due to changing aerodynamic conditions, imperfect articulatory positioning or increased tongue-tip tension and mass -- result in a fricative. Thus, trills seem to have very similar but more constrained aerodynamic and articulatory requirements than fricatives.

The similarities between fricatives and voiced trills are aerodynamic and muscular. Voiced trills and fricatives show similar pressure and airflow values. In Fig. 11, taken from Stevens (1971), a contour has been added to show the normal region of flows and pressure drops for voiced and voiceless trills. The region for voiced trills falls within (and is bounded by) the region for fricatives. The region for voiceless trills show flowrates similar to those of aspirated sounds. The degree of neuromotoric control and combined muscle contraction for trills is similar to that used in fricatives (Hardcastle 1976:132). In addition, fricatives and trills involve a highly constrained tongue dorsum and tongue tip to meet the critical positioning required for frication and trilling, resulting in their being very resistant to coarticulatory effects (Recasens et al., in press). The main differences between trills and fricatives seem to lie in the tension of the articulator (a stiffer articulator will not be free to vibrate); the configuration of the tongue (predorsum lowering and

postdorsum retraction for the trill to make possible the vertical vibration of the tongue tip as opposed to pre- and postdorsum raising and advancement for alveolar fricatives (Recasens et al., in press)), and initial tongue tip positioning (2.2.3).

The perceptual similarity between trills and fricatives is evidenced by the same sound being reported as a trill or a fricative by different investigators (Ladefoged and Maddieson 1996:241), and by the substitutions for the lingual trill by infants and unskilled adult native speakers -- intended lingual trills are rendered as a voiced uvular fricative [ʁ] (or uvular trill [R]), a postalveolar non-sibilant fricative, or a dental fricative [ʃ] (but not [w]).

In short, the results from the experiment (along with acquisitional and pathological sound substitutions) show that trills may become fricatives if the finely controlled articulatory or aerodynamic requirements for trills are not met, suggesting that fricatives involve a less complex articulation and allow a wider range of Po variation than trills. The reinterpretation of fricative variants of trills as the manifestation of phonological fricatives helps to explain the dialectal and diachronic change.

3.3. Absence of Nasal Trills. Another phonological universal involving trills is the absence of nasal trills. The non co-occurrence of trilling and nasality is dictated exclusively by aerodynamic reasons. An open velo-pharyngeal port for nasality would bleed the intraoral pressure required to make a relaxed oscillator vibrate for trills, as shown in the experiment where trilling ceased when the Po was vented with catheters that simulated nasal leakage. It has been shown that lingual trills require a high intraoral pressure ($>4\text{cmH}_2\text{O}$) to sustain tongue tip vibration (and even higher to initiate it), and that trilling was extinguished when the oral pressure was reduced by $2.5\text{ cmH}_2\text{O}$. Thus, trilling cannot combine with an open velopharyngeal port that reduces the Po by $2.5\text{ cmH}_2\text{O}$ or more (vent areas ≥ 8 or 17.8 mm^2 , depending on absolute Po values, from our estimates). The small velopharyngeal openings which do not impair trilling would most likely be insufficient to create a percept of nasalization. Thus, aerodynamic factors explain the lack of nasal trills. Nasal taps (as in American English 'winter') are, however, possible due to their ballistic muscular contraction being compatible with velopharyngeal opening.

4. Conclusions. This study provided empirical evidence on the aerodynamic conditions required for trilling and their compatibility with other features in making up speech segments. In addition, the universal principles in the phonological patterning of trills have been shown to reflect the constraints imposed by the physics and physiology of the speech production and perception systems, or to optimize competing requirements of perception and articulation.

Thus, voiced trills have been shown to have narrower aerodynamic requirements and to be less stable articulatorily than voiceless trills, but to exhibit a clearly modulated signal, clearly distinct from other speech segments, whereas voiceless trills are auditorily similar to fricatives. The statistical preference for voiced trills in phonological systems thus reflects a trade-off between competing demands of perception and production: preserving maximum auditory distinctiveness in a narrow range of aerodynamic conditions.

The common alternation and co-occurrence of trilling and frication result from similar production characteristics (aerodynamic, muscular and positional), with tongue tip trilling having more highly constrained requirements which may be easily thrown off, resulting in continuants. Devoicing of utterance final trills can be seen as a consequence of maintaining trilling with lowered subglottal pressure by removing the glottal resistance. Finally, the absence of nasal trills has been shown to be

dictated by the incompatibility of maintaining the relatively high oral pressure needed for trilling with an open velopharyngeal port required for nasality.

We have illustrated how aerodynamic factors, in combination with other constraints of production and perception, may determine feature cooccurrence restrictions and common phonological patterns. In Darwinian terms, we know that complex phonological structure can emerge from simpler (phonetic) mechanisms; and phonological universals seem to reflect the darwinian drift towards patterns that almost seem designed for their environment -- efficient speech production and perception --, eventually, the result of 'intelligent design'.

NOTES

* This research was supported by DGICYT grant PB 96-1158 to the UAB, Spain. The author gratefully acknowledges the help and support of John J. Ohala and Shawn Ying.

¹ Bilabial and (ingressive) uvular trills are reported in children's speech in the vocal play stage (Vihman 1996). The absence of tongue tip trills may be due to (i) the large volume of the tongue in relation to the oral cavity, leaving little room for distinct vertical movements of tongue tip or blade (Stark 1980), (ii) lack of maturity of the intrinsic muscles of the tongue to assume the position, shape and elasticity requirements of trills (Fletcher 1973), and lack of neuromuscular control of the tongue and kinaesthetic sensory information.

² The contribution of all these factors to trilling is evidenced by the conditions that have been anecdotally reported to aid trilling: (1) variations in air pressure (easier to trill after a [b], which involves increasing the airpressure downstream of the lingual constriction, or when singing, due to increased respiratory effort and subglottal Ps); (2) positioning (easier to trill when facing down, the gravitational force acting on the oscillator aids to closing the lingual channel); (3) tongue tension and mass (easier to trill between low vowels).

REFERENCES

- BARRY, W. Forthcoming. Another R-tickle. Ms.
- CANFIELD, D.L. 1981. Spanish pronunciation in the Americas. Chicago: The University of Chicago Press.
- CATFORD, J.D. 1977. Fundamental Problems in Phonetics. Bloomington: Indiana University Press.
- FLETCHER, S.G. 1973. Maturation of the speech mechanism. *Folia Phoniatrica* 25. 161-172.
- HARDCASTLE, W.J. 1976. Physiology of Speech Production. London: Academic Press.
- JIMÉNEZ, B.C. 1987. Acquisition of Spanish consonants in children aged 3-5,7 months. *Lang. Speech Hear. Serv. Schools* 18. 357-363.
- KAVITSKAYA, D. 1997. Aerodynamic constraints on the production of palatalized trills: The case of Slavic trilled [r]. *Proceedings of Eurospeech 1997, Rhodes, Greece, Vol. 4.* 751-754.
- LADEFOGED, P. and I. MADDIESON. 1996. The sounds of the world's languages. Oxford: Blackwell.
- LI, F. K. 1997. A handbook of comparative Thai. Hawaii: The University Press of Hawaii.
- LINDAU, M. 1985. The story of 'r'. *Phonetic Linguistics*, ed. by V.A. Fromkin, 157-68. Orlando, FL: Academic Press.
- LINDBLOM, B. 1983. Economy of speech gestures. *The Production of Speech*, ed. by P.F. MacNeilage, 217-246. New York: Springer-Verlag.
- , 1986. *Phonetic universals in vowel systems*. *Experimental Phonology*, ed. by J.J. Ohala and J.J. Jaeger, 13-44. Orlando, FL: Academic Press.
- , 1990. Explaining phonetic variation: a sketch of the H&H theory. *Speech Production and Speech Modelling*, ed. by W.J. Hardcastle and A. Marchal, 403-439. Dordrecht: Kluwer.

- MADDIESON, I. 1984. *Patterns of Sounds*. Cambridge: Cambridge University Press.
- McGOWAN, R. 1992. Tongue-tip trills and vocal tract wall compliance. *Journal of the Acoustical Society of America* 91 (5). 2903-2910.
- NAVARRO TOMÁS, T. 1950. *Manual de pronunciación española*. 20th ed., Madrid: CSIC.
- OHALA, J.J. 1974. Phonetic explanations in phonology. In Bruck, A., R. Fox & M. LaGaly (eds.), *Papers from the parasession on natural phonology*. Chicago: Chicago Linguistics Society. 251-274.
- , 1983. The origin of sound patterns in vocal tract constraints. *The Production of Speech*, ed. by P.F. MacNeilage, 189-216. New York: Springer-Verlag.
- , 1997. Emergent stops. *Proc. 4th Seoul International Conference on Linguistics [SICOL]* 11-15 Aug 1997. 84-91.
- OHALA, J.J., M.J. SOLÉ, and G. YING. 1997. The controversy of nasalized fricatives. *Proceedings of the 136th Acoustical Society of America*, Seattle.
- PERKELL, J.S., M.L. MATTHIES, M.A. SVIRSKY, and M.I. JORDAN. 1993. Trading relations between tongue-body raising and lip-rounding in production of the vowel /u/. *Journal of the Acoustical Society of America* 93. 2948-2961.
- QUILIS, A. 1981. *Fonética Acústica de la Lengua Española*. Madrid: Gredos.
- RECASENS, D. and M.D. PALLARÉS. In press. A study of /r/ and /r/ in the light of the 'DAC' coarticulation model. *Journal of Phonetics*.
- RUHLEN, M. 1975. *A guide to the languages of the world*. Language Universals Project. Stanford University.
- SPAJIC, S., P. LADEFOGED, and P. BHASKARARAO. 1996. The trills of Toda. *Journal of the International Phonetic Association* Vol. 26 (1). 1-21.
- STARK, R.A. 1980. Stages of speech development in the first year of life. In G. Yenikomshian, J.F. Kavanagh, and C.A. Ferguson (eds.), *Child Phonology 1: Production*. New York: Academic Press.
- STEVENS, K.N. 1971. Airflow and turbulence noise for fricative and stop consonants. *Journal of the Acoustical Society of America* 50. 1180-1192.
- , 1972. The quantal nature of speech: Evidence from articulatory-acoustic data. *Human Communication, A Unified View*, ed. by P.B. Denes and E.E. David Jr., 51-66. New York: McGraw-Hill.
- , 1989. On the quantal nature of speech. *Journal of Phonetics* 17. 3-46.
- TRAILL, A. 1985. *Phonetic and phonological studies of !Xóó Bushman (Quellen zur Khoisan-Forschung 1)*. Hamburg: Helmut Buske Verlag.
- VAN DEN BERG, B. 1958. *Foneiek van het Nederlands*. Dordrecht: Van Goor Zonen Den Haag.
- WESTBURY, J.R., and P.A. KEATING. 1986. On the naturalness of stop consonant voicing. *Journal of Linguistics* 22. 145-166.
- VIHMAN, M.M. 1996. *Phonological development*. Cambridge, Mass: Blackwell.
- ZLOTCHEW, C.M. 1971. The transformation of the multiple vibrant to the fricative velar in the Spanish of Puerto Rico. *Orbis* 23. 81-84.



BERKELEY LINGUISTICS SOCIETY

1203 Dwinelle Hall, University of California, Berkeley, CA 94720-2650 USA

Telephone and fax: (510) 642-5808; e-mail: bls@socrates.berkeley.edu

<http://faust.linguistics.berkeley.edu/BLS/>

Book Prices*

Parasession & Special Session (S) Topics		Individuals	Institutions
BLS-1	(1975)	\$10.00	\$15.00
BLS-2	(1976)	\$10.00	\$15.00
BLS-3	(1977)	\$10.00	\$15.00
BLS-4	(1978)	\$10.00	\$15.00
BLS-5	(1979)	\$10.00	\$15.00
BLS-6	(1980)	\$10.00	\$15.00
BLS-7	(1981)	\$10.00	\$15.00
BLS-8	(1982)	\$14.00	\$20.00
BLS-9	(1983)	\$10.00	\$15.00
BLS-10	(1984) Subordination.....	\$14.00	\$20.00
BLS-11	(1985) Poetics, Metrics, and Prosody.....	\$14.00	\$20.00
BLS-12	(1986) Semantic Typology.....	\$14.00	\$20.00
BLS-13	(1987) Grammar and Cognition.....	\$14.00	\$20.00
BLS-14	(1988) Grammaticalization.....	\$14.00	\$20.00
BLS-15	(1989) Theoretical Issues in Language Reconstruction.....	\$14.00	\$20.00
BLS-16	(1990) Legacy of Grice.....	\$16.00	\$20.00
BLS-16S	(1990) General Topics in American Indian Linguistics.....	\$8.00	\$12.00
BLS-17	(1991) The Grammar of Event Structure.....	\$16.00	\$20.00
BLS-17S	(1991) African Language Structures.....	\$8.00	\$12.00
BLS-18	(1992) Place of Morphology in a Grammar.....	\$16.00	\$20.00
BLS-18S	(1992) Typology of Tone Languages.....	\$8.00	\$12.00
BLS-19	(1993) Semantic Typology and Universals.....	\$16.00	\$20.00
BLS-19S	(1993) Syntactic Issues in Native American Languages.....	\$8.00	\$12.00
BLS-20	(1994) Anniversary Session: Dedicated to the Contributions of Charles J. Fillmore.....	\$16.00	\$20.00
BLS-20S	(1994) Historical Issues in African Linguistics.....	\$8.00	\$12.00
BLS-21	(1995) Historical Issues in Sociolinguistics/..... Social Issues in Historical Linguistics.....	\$16.00	\$20.00
BLS-21S	(1995) Discourse Issues in Southeast Asian Languages.....	\$8.00	\$12.00
BLS-22	(1996) The Role of Learnability in Grammatical Theory.....	\$16.00	\$20.00
BLS-22S	(1996) Historical Issues in Native American Languages.....	\$8.00	\$12.00
BLS-23	(1997) Pragmatics and Grammatical Structure.....	\$16.00	\$20.00
BLS-23S	(1997) Syntax and Semantics in Africa.....	\$8.00	\$12.00
BLS-24	(1998) Phonetics and Phonological Universals.....	\$20.00	\$25.00
BLS-24S	(1998) Indo-European Subgrouping and Internal Relations.....	\$8.00	\$12.00
Author Index to BLS 1-16S.....		\$2.00	(includes shipping)

California Sales Tax

All orders from residents of California must include sales tax according to county of residence.

	Within USA	Outside USA
Shipping and Handling.....	\$3.50	\$5.00

ORDERS MUST BE ACCOMPANIED BY PAYMENT

Make checks and money orders payable to the Berkeley Linguistics Society

*All prices subject to change without notice.