PROCEEDINGS OF THE TWENTY-EIGHTH ANNUAL MEETING OF THE

BERKELEY LINGUISTICS SOCIETY

February 15-18, 2002

GENERAL SESSION

and

PARASESSION on FIELD LINGUISTICS

Edited by

Julie Larson & Mary Paster

Berkeley Linguistics Society
Berkeley, CA, USA

PROCEEDINGS OF THE TWENTY-EIGHTH ANNUAL MEETING OF THE

BERKELEY LINGUISTICS SOCIETY

February 15-18, 2002

GENERAL SESSION

and

PARASESSION on FIELD LINGUISTICS

Edited by

Julie Larson & Mary Paster

Berkeley Linguistics Society
Berkeley, CA, USA

Berkeley Linguistics Society
University of California, Berkeley
Linguistics Department
1203 Dwinelle Hall
Berkeley, CA 94720-2650
USA

ISSN 0363-2946

LCCN 76-640143

Printed by Sheridan Books, Inc.
P.O. Box 1203
Ann Arbor, Michigan 48108
USA

TABLE OF CONTENTS

GENERAL SESSION

GENERAL SESSION

Base Selection in Analogical Change in Yiddish

ADAM ALBRIGHT
University of California, Los Angeles

0. Introduction

A notable difference between Yiddish and German verb paradigms is that Yiddish has no vowel alternations in the present tense.[1] Whereas Middle High German (MHG) and Modern German (NHG) often have alternations among the singular forms (1a), or between the singular and plural (1b), Yiddish never does (2).

(1) MHG present tense vowel alternations

a. 'dig'	sg.	pl.		b. 'know'	sg.	pl.
	1st	*grabe*	*graben*		1st	*weiȝ*	*wiȝȝen*
	2nd	*grebest*	*grabet*		2nd	*weist*	*wiȝȝet*
	3rd	*grebet*	*graben*		3rd	*weiȝ*	*wiȝȝen*

(2) Yiddish paradigms have no vowel alternations[2]

a. 'dig'	sg.	pl.		b. 'know'	sg.	pl.
	1st	*grɔb*	*grɔbən*		1st	*veys*	*veysən*
	2nd	*grɔbst*	*grɔbt*		2nd	*veyst*	*veyst*
	3rd	*grɔbt*	*grɔbən*		3rd	*veyst*	*veysən*

As I will show in Section 1, the form that has been extended in Yiddish is always the expected 1sg form. Interestingly, although this change is across the board in Yiddish, it is apparently unattested in any other German dialect.

Paradigmatic levelings of this sort, in which some members of the paradigm are rebuilt based on other forms, pose a well-known problem in historical linguistics. On the one hand, they occur frequently, and seem natural and unsurprising. In

[1]The Yiddish data in this paper concerns the eastern dialects of Central and Eastern Europe; I do not know if the same holds true of the western dialects.

[2]For Yiddish examples, I will use YIVO transliteration (http://www.yivoinstitute.org/yiddish/alefbeys.htm), with two minor modifications: I will use the IPA symbol ɔ instead of YIVO *o* for *komets-aleph*, and I will use *-ən* instead of YIVO *-en*/*-n* for the infinitive/1pl/3pl suffix. For MHG forms, I will use the standardized orthography of Paul, Wiehl, and Grosse (1989, §§18–20), in which ˆ marks long vowels, *ë* is a short open [e], and *ȝ* is a coronal sibilant fricative, possibly fortis, possibly postalveolar (Paul et al, §151). The change of MHG short [a] > Yiddish [ɔ] reflects a regular sound change; the correspondences between MHG ⟨w⟩ and Y ⟨v⟩, MHG ⟨ei⟩ and Y ⟨ey⟩, MHG ⟨s⟩ and Y ⟨z⟩, and MHG ⟨ȝ(ȝ)⟩ and Y ⟨s⟩ are also completely regular.

fact, it is often suggested that the desire for nonalternating paradigms is simply a primitive of language, sometimes refered to as "Humboldt's Universal", or, more recently, as Uniform Exponence (Kenstowicz 1995) or Paradigm Uniformity (Steriade 2000). On the other hand, a blanket preference for uniform paradigms can only go so far in explaining levelings: it can tell us an alternation is likely to be leveled, but not necessarily in which direction. Why was the 1sg extended, and not the 3sg, yielding paradigms like **greb*, **grebst*, **grebt*? And what distinguished Yiddish from other German dialects, otherwise very similar morphologically?

Many proposals over the years have attempted to explain the direction of leveling. The usual approach, pioneered by Kuryłowicz (1947) and Mańczak (1958) and continued by Bybee (1985) and others, has been to focus on tendencies, or groups of factors that compete to determine the direction of a change. Under such an approach, it is possible to derive typological predictions – leveling is often to the isolation form, the most frequent form, the 3sg, and so on – but it is impossible to make predictions about a given language at a given time, because we do not know which factors will win in that particular case. In this paper, I will pursue a different approach, in the tradition of Paul (1920) and Kiparsky (1965), which focuses on the role of the learner in language change. In particular, I will pursue the hypothesis that language learners impose a structure on paradigms that helps them to construct phonological and morphological grammars that generate unknown forms as accurately or as confidently as possible. The way that they do this, I will claim, is by seeking a base form within the paradigm that is "maximally informative" – that is, that suffers the least serious phonological and morphological neutralizations – and then deriving the remaining forms in the paradigm from the base form. Under this approach, we can use the direction of the grammar (base form → derived forms) to predict the direction of possible analogical changes. In section 2, I will show that the 1sg form preserves the most contrasts in Yiddish, and thus would be selected as the base form in the proposed model. Finally, in section 3, I will argue that the advantages of the 1sg are unique to Yiddish, due to small but crucial differences between Yiddish and other German dialects. Thus, the proposed analysis gives us insight not only into the question of why verbs were leveled to the 1sg in Yiddish, but also into why this did not occur elsewhere.

1. Leveling to the 1sg in Yiddish present tense paradigms

As shown in (2), Modern Yiddish has no vowel alternations in the present tense (Rockowitz 1979). In this section, I will show that in virtually all cases, the vowel of the 1sg has been extended to the rest of the paradigm. To show this, we will consider the following candidates as sources for the modern present tense stem, eliminating all except the 1sg[3]: 1sg, 2sg, 3sg, 1pl, 2pl, 3pl, infinitive, and UR. I

[3]This list includes almost all of the verb forms that occur in Yiddish; the only other forms are the present participle, stem (*shtam*), past participle, and imperative. The present participle and stem are always based on the infinitive, so any conclusions regarding the infinitive hold of them as well. The past participle is also demonstrably not the source of the modern present stem. The singular imperative form is always identical with the 1sg, and could equally well have served as the base for the leveling discussed here. For expository ease, I will refer to the 1sg, but I cannot preclude the possibility that it was the singular imperative instead.

will start with the assumption that the origin of Yiddish was some form of MHG, so it is useful to begin by considering the possible types of present tense paradigms that occurred in MHG (Paul, Wiehl, and Grosse 1989, chap. 7).

1.1. MHG present tense patterns

Most MHG verbs had the same vowel throughout the entire present tense, with no alternations, as in (3); verbs of this type included the "strong" classes I, IIIa, and some of VII, as well as all of the "weak" verbs.

(3) No alternations (Strong I, Stong IIIa, some Strong VII, all weak)

a.	'live'	*lëbe*	*lëben*	b.	'say'	*sage*	*sagen*
		lëbest	*lëbet*			*sagest*	*saget*
		lëbet	*lëben*			*saget*	*sagen*

In another set of verbs, an *a* in the root surfaced as an *e* in the 2sg and 3sg due to a process known as *umlaut*, as in (4). This occurred in strong class VI and the remainder of strong class VII.

(4) 2sg, 3sg different due to Umlaut (*a* ~ *e*): Strong VI, some Strong VII

'dig'	*grabe*	*graben*	(Also *varn* 'travel', *laden* 'invite', *slahen*
	grebest	*grabet*	'beat', *halten* 'hold', *lâȝen* 'let')
	grebet	*graben*	

Finally, a third set of verbs showed vowel alternations between the entire singular and the entire plural. This pattern occurred in two types of verbs. In some, the present tense derived from a Proto-Indo-European perfect, and the singular/plural alternation reflected a PIE alternation in the perfect tense (*ablaut*); these verbs are known as *preterite presents* (5a). In others, the alternation was due to a phonological process in Old High German that raised mid vowels before a following high vowel, causing the singular to diverge from the plural (5b); this pattern is sometimes referred to as *Wechselflexion* ("alternating inflection"), and occurred in strong class II (*ie* ~ *iu*), as well as IIIb, IV, and V (*ë* ~ *i*).

(5) Singular ~ plural alternations

a. Preterite presents

'know'	*weiȝ*	*wiȝȝen*
	weist	*wiȝȝet*
	weiȝ	*wiȝȝen*

(Also *kunnen* 'can', *durfen* 'need', *suln* 'should', etc.)

b. *Wechselflexion*

'give'	*gibe*	*gëben*
	gibest	*gëbet*
	gibet	*gëben*

(Also *nëmen* 'take', *ëȝȝen* 'eat', *gieȝen* 'pour', etc.)

1.2. Yiddish present tense patterns

Let us now consider the fate of these patterns in Yiddish. Unsurprisingly, verbs with no alternations in MHG continue to have no alternations in Yiddish, as seen in (6).

(6) Non-alternating verbs remain non-alternating in Yiddish

a.	'live'	*leb*	*lebən*	b.	'say'	*zɔg*	*zɔgən*
		lebst	*lebt*			*zɔgst*	*zɔgt*
		lebt	*lebən*			*zɔgt*	*zɔgən*

Turning to umlaut alternations (1sg *grabe*, 2sg *grebst*), these were leveled to the non-umlaut (*a*) alternant (7). Thus, it appears that the base, or pivot, of the leveling was not the 2sg or 3sg, or else the modern Yiddish paradigm would have *e*.

(7) Umlaut verbs leveled to non-umlaut (*a*) alternant: *grɔbən* 'dig'

1st	*grɔb*	*grɔbən*	infin.	*grɔbən*
2nd	**grebst* > *grɔbst*	*grɔbt*		
3rd	**grebt* > *grɔbt*	*grɔbən*		

Considering next the preterite present verbs, we find that for these, the Yiddish present tense forms come from MHG singular forms. This is shown in (8) for the verbs *darfən* 'need' and *visən* 'know', whose present tense forms are derived from the MHG singular forms *darf-* and *weiȥ-*, and not the plural *dürf-/durf-* and *wiȥȥ-*. Other examples include *muzən* 'must' (< MHG sg. *muoȥ*, not pl. *müeȥȥen*), *tɔrən* 'must' (< MHG sg. *tar*, not pl. *türren*), and *zolən* (< MHG sg. *sol*, not pl. *süln*)

(8) Preterite Present verbs leveled to singular

a.

'need'	sg.	pl.
1st	*darf*	**dürfən* > *darfən*
2nd	*darfst*	**dürft* > *darft*
3rd	*darf*	**dürfən* > *darfən*
infin.	**dürfən* > *darfən*	
UR	*/dürf-/*, */darf-/* > */darf-/*	

b.

'know'	sg.	pl.
1st	*veys*	**visən* > *veysən*
2nd	*veyst*	**vist* > *veyst*
3rd	*veys(t)*	**visən* > *veysən*
infin.	*visən*	
UR	*/vis/*, */veys/*	

We can conclude that the generalized form was not a plural form or the infinitive – in fact, most infinitives of preterite presents were also rebuilt based on singular forms. Furthermore, the two MHG stem alternants (*darf-*, *dürf-*) cannot easily be reduced to a single UR, since they involve an idiosyncratic alternation that occurs in only one other verb, and it is not clear how to derive *ü* from *a* or vice versa. The most promising analysis is to list two URs for these verbs (e.g., /darf/, /dürf/), in which case the generalized form does match one of the MHG UR's (/darf/). However, simply saying the UR has been generalized does not explain why one UR was chosen over the other. Putting this conclusion together with that from the umlaut verbs, we have now eliminated the 2sg, 3sg, all of the plural, the infinitive, and the UR as sources of Yiddish present tense forms. Thus, it appears that the 1sg is only possible source.

The data so far converge on the 1sg as the source for Yiddish present tenses. Unfortunately, when we turn to the *Wechselflexion* verbs, the situation is more complicated. From what we have seen thus far, we would expect these verbs to generalize the *i* of the singular, and indeed this is what we find with *gebən* 'give' (9a). For most MHG *Wechselflexion* verbs, however, Yiddish seems to have generalized the *e* of the plural/infinitive, as in *nemən* 'take' (9b).

(9) Fate of *Wechselflexion* verbs

a. Generalized *i* from sg: *gebən* 'give'

1st	*gib*	**gebən* > *gibən*
2nd	*gibst*	**gebt* > *gibt*
3rd	*gibt*	**gebən* > *gibən*
infin.	*gebən*	

b. Generalized *e* from pl: *nemən* 'take'

1st	**nim* > *nem*	*nemən*
2nd	**nimst* > *nemst*	*nemt*
3rd	**nimt* > *nemt*	*nemən*
infin.	*nemən*	

The pattern of generalized *e* is found not only in *nemən*, but also in *esən* 'eat' (*es*, **is*), *fargesən* 'forget' (*farges*, **fargis*), *zeyn* 'see' (*zey*, **zi*), *vern* 'become' (*ver*, **vir*), *helfən* 'help' (*helf*, **hilf*) etc. Why do these verbs show a different pattern from all other verbs? Is this an exception to generalization of the 1sg form?

I would like to argue that verbs like *nemən* are not exceptions, but rather that they already contained *e* in the 1sg at the time that Yiddish diverged from other German dialects.[4] I began this section with the assumption that Yiddish began as some form of MHG, exemplified by the literary MHG forms in (3)-(5). However, the history of the *Wechselflexion* in German is somewhat complicated, and it is not clear that the paradigm in (5) is the correct starting point for Yiddish. According to the standard account (Paul et al. 1989 §§31-35), *Wechselflexion* was due to a phonological process in OHG raising /e/ to [i] when a high vowel (*u*, *i*) was in the following syllable. Since singular suffixes had high vowels and plural suffixes had mid vowels, this led to an alternation between *i* in the singular and *e* in the plural (*issu* ~ *ëssêm* 'eat-1sg/1pl'). In MHG, all suffix vowels were reduced to schwa, making the *i*~*e* alternation a purely morphological difference between the singular and the plural (*isse* ~ *ëssen*). This pattern is found in all MHG texts until the mid-15th C (Dammers, Hoffmann, and Solms 1988, §148.4). Finally, during late MHG or early NHG times, the vowel of the 1sg lowered back to *e* (*isse* > *esse*), probably under the influence of the umlaut pattern (1sg vs. 2,3sg, as in (4) above). 1sg forms with *e* began to occur regularly in "middle German" (Fränkisch, Thüringisch, Böhmisch, Schlesisch) during the fifteenth century (Paul et al. 1989, §242, note 1; Philipp 1980, p. 66), appearing earlier in the west than in the east (Dammers, et al. 1988, §148.4). The change proceeded verb-by-verb, with considerable variation even between occurrences of the same verb in the same text (Kern 1903, pp. 47-60; Geyer 1912, §31-§32), but eventually all *Wechselflexion* verbs were affected.

What we see, then, is that 1sg forms with *e* in German predate the Middle Yiddish period (16th-17th C). Thus, I hypothesize that Yiddish already had *e* in the 1sg of *Wechselflexion* verbs prior to leveling. If this is the case, then the *e* of *nemen* is not an exception to the generalization that leveling was always to the 1sg.

I have found only three exceptions to generalization of the 1sg in Yiddish: (1) *zayn(ən)* 'to be' retains a suppletive paradigm (2) the future auxiliary *velən* derives from a conditional form, not the 1sg present indicative, and (3) *gefelən* 'be pleasing' is used predominantly in the 3rd person, and derives from a 3sg form (*gefelt* 'it is pleasing').[5] These exceptions are not all that surprising – two are extremely high frequency, and the third has semantic restrictions. In sum, for every type of

---

[4] I am not making any commitment as to when Yiddish ceased to be a sociolect of German, except to suppose that the two probably continued to co-evolve at least until the beginning of the Middle Yiddish period (c. 16th C), when Yiddish literature began to flourish in the east, eastward migrations trickled off, and significant east-west dialect differences emerged (Weinreich 1980, p.724-726).

[5] This effect, in which lexical semantics influences the direction of analogy, is discussed by Tiersma (1982) under the rubric of *local markedness*. However, this is the only example in Yiddish, so it seems extravagant to invoke local markedness to explain just one case. It may also derive from a MHG variant of *gefallen*; another example is Yiddish *fregən* 'ask', derived from MHG *vrëgen*, a variant of *vragen* (Paul et al. 1989, §30).

MHG verb, it appears that the 1sg form has been extended to the remainder of the paradigm in Yiddish. This leveling has been remarkably complete, affecting virtually all verbs. In the next section, I will consider the question of why Yiddish paradigms were rebuilt on the basis of this, and not some other form.

2. The 1sg as the optimal base in Yiddish
2.1. Identifying the optimal base

Why did the 1sg have a privileged status in Yiddish? In this section, I will argue that it was the "maximally informative"form, suffering from the fewest phonological neutralizations, and maintaining the most contrasts. In order to show this, I will examine a pre-leveling version of Yiddish, considering which parts of the paradigm suffered from neutralizations, and how many verbs were affected in each case.

Yiddish, like German and English, disallows coda clusters of obstruents with voicing disagreement ($*bs]_\sigma$, $*pd]_\sigma$, etc.). When a suffix containing voiceless obstruents (2sg *-st*, 3sg/2pl *-t*) is added to a root ending in a voiced obstruent, the root-final obstruent is devoiced. The result is that in the 2sg, 3sg, and 2pl, the contrast between root-final voiced and voiceless obstruents is neutralized.

(10) Neutralization in the 2sg/3sg/2pl: voicing assimilation to suffix

	libən 'to love'	*zipən* 'to sift'
1sg	*lib*	*zip*
2sg	***lipst***	***zipst***
3sg	***lipt***	***zipt***
1pl	*libən*	*zipən*
2pl	***lipt***	***zipt***
3pl	*libən*	*zipən*
infinitive	*libən*	*zipən*

This neutralization affects all obstruent pairs with a voicing contrast, of which Yiddish has seven (*p/b, t/d, k/g, f/v, s/z, ʃ/ʒ, tʃ/dʒ*). A hypothesis of the current approach is that the seriousness of a neutralization depends not only on the number of phonemes involved, but also on the number of lexical items whose underlying form cannot be recovered due to the neutralization. In order to estimate of the number of verbs whose final segment would be ambiguous due to voicing assimilation, I counted the number of verbs ending in these 14 obstruents in the German portion of CELEX.[6] For CELEX counts, I considered only verb lemmas that had a token frequency of 1 or greater and did not contain a separable initial element (separable prefix, incorporated object, adverb); this left a total of 4877 verbs. As it turns out, 1988 of these end in obstruents with voicing contrasts, meaning approximately 41% of all verbs are ambiguous in the 2sg, 3sg, and 2pl.

Another set of neutralizations in Yiddish comes from a ban on word-internal geminates. For example, adding the 2sg suffix *-st* to a verb ending in *s* or *z* should

[6]Ideally, we count a lexicon of Middle Yiddish, but this does not exist in searchable form, and counts from German form a reasonable approximation. There are certainly numerous lexical differences between Yiddish and German, and even some phonological ones – e.g., Yiddish has verb roots ending in [v] and [dʒ], which are rare or absent in German. However, most common Yiddish verbs are shared with German, and there is no reason to believe that the lexical differences would significantly alter the proportion of major classes like obstruent-final verbs, strident-final verbs, etc.

yield the sequence *-sst* (with devoicing of *z* to satisfy voicing agreement). This sequence actually surfaces as degeminated *-st*: /veys-st/ → [veyst], not *[veysst]. The result is that *s*- and *z*-final verbs are neutralized with vowel-final verbs in the 2sg, as seen in (11a). For the 3sg and 2pl, the suffix is *-t*, and an equivalent degemination of *tt* (fed by /d/ → [t] devoicing) applies (11b).

(11) Neutralizations caused by degemination

a. Neutralization in the 2sg: devoice *z*, degeminate of *ss*

	geyn 'go'	*visən* 'know'	*vayzn* 'show'
1sg	*gey*	*veys*	*vayz*
2sg	***geyst***	***veyst***	***vayst***
3sg	*geyt*	*veys(t)*	*vayst*
1pl	*geyən*	*veysən*	*vayzən*
2pl	*geyt*	*veyst*	*vayst*
3pl	*geyən*	*veysən*	*vayzən*
infin	*geyən*	*visən*	*vayzən*

b. Neutralization in the 3sg/2pl: devoice *d*, degeminate *tt*

	faltən 'fold'	*falən* 'fall'	*redən* 'talk'
1sg	*falt*	*fal*	*red*
2sg	*fal(t)st*	*falst*	*retst*
3sg	***falt***	***falt***	***ret***
1pl	*faltən*	*falən*	*redən*
2pl	***falt***	***falt***	***ret***
3pl	*faltən*	*falən*	*redən*
infin	*faltən*	*falən*	*redən*

How many lexical items would be affected by these neutralizations? The voicing neutralization of *s/z* and *t/d* was already included in the count for voicing assimilation above, but degemination means that vowel-final roots are also ambiguous in these forms – an additional 227 words in CELEX, or 5% of the verbal vocabulary.

So far, we have examined neutralizations in forms with obstruent suffixes – the 2sg, 3sg, and 2pl. Turning to the 1pl, 3pl, and infinitive forms, the suffix for all of these forms is *-ən*. Since this suffix is vowel-initial, and Yiddish allows vowels to occur in hiatus, it does not give rise to illegal sequences to trigger assimilation or deletion, with one exception: if the verb root ends in a schwa (e.g., *pɔrə-* 'fiddle with'),then the 1pl/3pl/infinitive form ends simply in *-ən*, not **-əən*. The reduction of /əə/ to [ə], motivated by a ban on long schwa (*[əː]), means that in these forms, schwa-final and non-schwa-final verbs are neutralized.

(12) Neutralizations in the 1pl/3pl/infinitive: stem-final /ə/

	pɔrən 'to match'	*pɔrən* 'to fiddle with'
1sg	*pɔr*	*pɔrə*
2sg	*pɔrst*	*pɔrəst*
3sg	*pɔrt*	*pɔrət*
1pl	***pɔrən***	***pɔrən***
2pl	*pɔrt*	*pɔrət*
3pl	***pɔrən***	***pɔrən***
infinitive	***pɔrən***	***pɔrən***

How serious is this neutralization? German does not have schwa-final verbs, so we cannot use CELEX to estimate the number of lexical items that would be affected by it. Instead, I took a sample from Weinreich (1990), counting all of the verbs beginning with [l]. (This segment was chosen to avoid skewing the sample by including uniquely Slavic onsets like *shtsh-* or *tl-*, or charactericaly Hebrew onsets like *mə-*; [l]-initial words [l] seem to come from Germanic, Slavic, and Hebrew in representative proportions.) Of the 90 [l]-initial verbs, 9 of them (10%) have stem-final ə. Thus, a contrast that is seen in a significant portion of the Yiddish verbal vocabulary is neutralized in the 1pl/3pl/infinitive forms.

As with other neutralizations, it is worth considering whether the presence of stem-final [ə] is truly neutralized in the 1pl/3pl/infinitive forms, or whether it could be predicted using secondary cues. In casual speech in many dialects, the [ə] of the *-ən* suffix may be lost, resulting in a syllabic nasal agreeing in place with a preceeding consonant: [libən] ~ [libm̩] 'love-1pl/3pl/inf.'. This process affects suffix [ə], but not stem-final [ə] – meaning that verbs with stem-final [ə] might be distinguished by lack of a [ə]-less variant (*pɔrən/pɔrn̩* 'match' vs. *pɔrən/*pɔrn̩* 'fiddle with'). This difference would be rather poor evidence about the status of final [ə], however. First, it requires distinguishing a syllabic nasal from a schwa-nasal sequence, which is not always easy to do, especially after continuants. Furthermore, this form is only informative if it is determined to end in a syllabic nasal; if it ends in *-ən*, no conclusion can be drawn. Finally, reduction of *-ən* to syllabic *-n* does not seem to occur in all environments. This is reflected in the YIVO orthography, which uses *-en* after *m*, *n*, *ng*, *nk*, and syllabic *l*, but *-n* elsewhere. In practice, reduction is probably not as categorically restricted as the orthography implies, but occurs most often after stops, least often after vowels, and so on. Therefore, we would be able to use the 1pl/3pl/infinitive form to infer a lack of final [ə] for at best only a subset of verbs in the language.

Another potentially relevant fact is that virtually all [ə]-final verbs come from Slavic or Hebrew. If a verb can be identified as non-Germanic, perhaps because it contains a sequence that is illegal in German (e.g., ***pyeshtshən*** 'caress', ***tliən*** 'smolder', ***strashən*** 'threaten'), it is much more likely to have a stem-final [ə]. In addition, there are two derivational suffixes with final [ə]: the verbal suffix *-eve* (e.g., *ratevə-n* 'rescue', *bushevə-n* 'rage'), and the mimetic suffix *-ke* (e.g., *shushkə-n* 'whisper', *hafkə-n* 'bark'). Therefore, verbs ending in *-ken* and *-even* are extremely likely to have final [ə]. These two facts make it somewhat easier to guess whether a new word should have final [ə], but it is still far from predictable. In fact, there are a number of other minimal or near-minimal pairs, including *bray-ən* 'brew' vs. *brayə-n* 'talk endlessly', *blank-ən* 'gleam' vs. *blɔnkə-n* 'stray', and *kvetsh-ən* 'squeeze' vs. *kvitshə-n* 'squeak'. The upshot is that although it may be possible to guess about the status of a final [ə] in some cases, it is still easier to choose a form that shows it unambiguously (a singular form or the 2pl).

The neutralizations discussed so far affected forms with overt suffixes – that is, all forms except the 1sg. The 1sg did not suffer such severe neutralizations, because no phonological processes affected segments in stem-final position.[7] It would not have been completely free from neutralizations, however; in fact, two properties of verbs could not have been predicted from the 1sg form alone. Umlaut verbs like *fɔrən* would have had the same vowel (ɔ) as non-umlaut verbs like *pɔrən* in the 1sg, and preterite present and *Wechselflexion* verbs would likewise have been indistinguishable from non-alternating verbs in this form. A crucial difference from

[7]It appears that an earlier stage of Yiddish did have final devoicing, but this was lost early on in most dialects; see King (1980) for discussion.

the neutralizations discussed above, though, is that umlaut, preterite present, and *Wechselflexion* verbs would have been ambiguous with non-alternating verbs not only in the 1sg, but in *all* parts of the paradigm. (Recall that we are considering here a version of Yiddish prior to paradigm leveling; in actual Modern Yiddish, all of these verbs have uniform paradigms (7-9).) The shading in (13) shows that in some parts of the paradigm, these verbs were neutralized with the vowel in the middle column, while in other parts of the paradigm, they were neutralized with the vowel in the final column.

(13) Neutralizations that include the 1sg (pre-leveling forms)

a. Umlaut and non-umlaut verbs

	fɔrən 'travel'	*pɔrən* 'match'	*hern* 'hear'
1sg	***fɔr***	***pɔr***	*her*
2sg	***ferst***	*pɔrst*	***herst***
3sg	***fert***	*pɔrt*	***hert***
1pl	***fɔrən***	***pɔrən***	*hern*
2pl	***fɔrt***	***pɔrt***	*hert*
3pl	***fɔrən***	***pɔrən***	*hern*
infin.	***fɔrən***	***pɔrən***	*hern*

b. Preterite presents and *Wechselflexion*

	visən 'know'	*heysən* 'order'	*vishən* 'wipe'
1sg	***veys***	***heys***	*vish*
2sg	***veyst***	***heyst***	*vishst*
3sg	***veys(t)***	***heyst***	*visht*
1pl	***visən***	*heysən*	***vishən***
2pl	***vist***	*heyst*	***visht***
3pl	***visən***	*heysən*	***vishən***
infin.	***visən***	*heysən*	***vishən***

Since these neutralizations affect all parts of the paradigm, they do not favor any particular choice of base, and it is perhaps unnecessary to count the number of lexical items involved. It may be noted, however, that compared with the neutralizations discussed above, these affected a very small number of words. In MHG, umlaut occurred in a handful of verbs, mostly in strong classes VI and VII – perhaps less than two dozen altogether (Paul et al. 1989, §§ 251-253). Added to these were about a dozen preterite present verbs (§§ 269-275) and around 70 verbs in the *Wechselflexion* classes (IIIa, IV, V; §§ 247-250), totaling about 2% of verbs.

The combined effect of these neutralizations is summarized in Figure 1, which shows the number of lexical items whose underlying form could not be unambiguously recovered from each part of the paradigm. In sum, the 1sg form preserves the greatest number of phonemic distinctions, including the voicing of stem-final obstruents, the presence of stem-final *t*, *d*, *s*, and *z*, and the presence of stem-final ə. Thus, given a 1sg form, it would be possible to predict every form of every word with absolute certainty, with the exception of the 2sg/3sg of umlaut verbs and the plurals of preterite present and *Wechselflexion* verbs.

2.2. Using the 1sg as the base to derive Yiddish verb paradigms

Suppose that you are acquiring a version of Yiddish prior to paradigm leveling. Your goal is to be able to produce and comprehend all forms of all verbs, and in order to do this, you need to learn their distinctive properties. I have shown that the 1sg provides almost all of them, and would thus be the optimal choice of base form to predict other forms. In the model proposed here, once the learner has identified the base, she goes on to develop a grammar to derive the rest of the paradigm from that form. For Yiddish, the grammar would include morphological rules like suffixing *-st* to form a 2sg, *-t* to form a 3sg or 2pl, *-ən* to form a 1pl/3pl/infinitive,

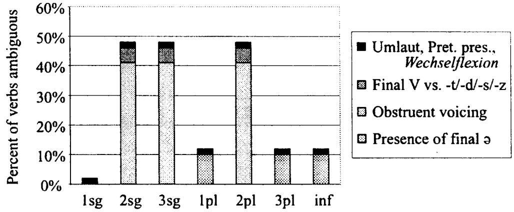


Figure 1: Summary of neutralizations affecting Yiddish verb forms

as well as phonological processes like obstruent voicing assimilation, degemination of /tt/ and /ss/, and elision of schwas in /əə/ sequences. If we use the 1sg as an input for these rules, they will yield the correct result for almost all forms of all words. The only exceptions are the 2,3sg of umlaut verbs, for which we predict incorrect forms like **fɔrst* and **fɔrt*, and the plurals of preterite present and *Wechselflexion* verbs, for which we predict incorrect forms like **veysən* and **gibən*. Under this approach, then, forms with umlaut (*ferst*) and with sg.∼pl. alternations (*visən*) must be learned as exceptions.[8] If a speaker forgets or is unable to access the correct exceptional form, she will use the grammar to produce an "overregularized" form (*fɔrst, veysən*). If these mistakes are accepted and adopted by the speech community, they will eventually replace the old, exceptional forms. There are clearly many factors that determine how willingly a community adopts new forms; the thoroughness of the change in Yiddish may have been facilitated by the lack of a standard language or widespread literacy, and perhaps even by a conscious desire to differentiate Yiddish from German. The model that I am presenting here simply attempts to predict what the *potential* overregularizations would have been.

3. Comparison with other German dialects

The Yiddish leveling seems like a very natural change, even if its completeness is a bit striking. If it is really so natural, however, we would expect that it might also have occurred in some other related dialects. An informal survey of dialect descriptions revealed several candidates for dialects that superficially resemble to Yiddish in their present tense forms, but all turned out to have different explanations. In Dutch and some northern German dialects, the present singular paradigm is always uniform (*graaf, graaft, graaft* 'dig'), as in Yiddish. However, these dialects never had umlaut to begin with, so this uniformity is not due to leveling. Some southern German dialects (Schwabian, Frankish, Bavarian, etc.) did historically have umlaut, and also have uniform present tense paradigms (e.g., Bavarian *grab, grabsd, grabd*) (Schirmunski 1962; Zehetner 1989). However, these dialects show leveling only of the singular forms, while maintaining singular ∼

[8]It does not matter for present purposes whether these are stored as whole-word exceptions, or whether we posit rules that apply only to words that are lexically specified for them. All that matters is that *ferst* and *visən* cannot be derived productively, and require an overriding word-specific mechanism that may fail.

plural alternations in preterite present verbs. It appears that these dialects have lost the umlaut rule for the 2,3sg, rather than undergoing true paradigm leveling. Finally, Early NHG texts occasionally have *e* in the 2,3sg of *Wechselflexion* verbs, such as Fischart's *schmelzt* 'melts-3sg' (Standard German *schmilzt*) (Geyer 1912, §23.6), reminiscent of the generalization of *e* in Yiddish (9b). This seems to happen mainly with verbs that are also sometimes given regular (weak) pasts, however. Weak verbs never have *Wechselflexion*, and its loss was probably part of a larger trend to create weak counterparts of strong verbs in Early NHG. In none of these cases do we find compelling evidence of paradigm leveling of the kind seen in Yiddish.

In contrast, there have been numerous changes in German that have introduced new alternations. The change from *i* to *e* in the 1sg of *Wechselflexion* verbs, for example, is usually seen as an extension of the umlaut pattern (raising/fronting in the 2,3sg), and umlaut has been extended to other verbs as well. In addition, many verbs have been rebuilt on the basis of 3sg forms, such as *ziemen* from Strong IV *zëmen*, and *wiegen* 'rock' from Strong V *wëgen* 'move'.

It would be difficult to prove that an equivalent leveling has never occurred in any other form of German, but my tentative conclusion is that German has generally gone in a different direction. So why would Yiddish have departed so radically in this respect? Considering the differences between Yiddish and German, we find that two of the neutralizations discussed above do not occur in German. First, the degemination of /dt/ and /tt/ to [t] ((11b) above) is found in only a few dialects (Schirmunski 1962), meaning the 3sg and 2pl forms preserve the voicing contrast between stem-final *t* and *d*, and keep both distinct from stem-final vowels. This is significant, because 562, or 12% of the verbs in the CELEX corpus end in coronal stops. Furthermore, German has no stem-final [ə], eliminating a major source of ambiguity in the 1pl/3pl/inf. forms. For these reasons, the 1sg form is not uniquely informative in Standard NHG; the plural and infinitive forms are just as good, and even the 2sg/3sg/2pl forms are not as ambiguous as in Yiddish.

4. Discussion and conclusion

In this paper, I have shown that the 1sg served as the base of a paradigm leveling in Yiddish, affecting almost every verb of the language. A comparison of neutralizations showed that before the change, the 1sg would have been uniquely revealing about the underlying form of the verb root. This is because the 1sg suffered from the fewest phonological neutralizations, involving the fewest lexical items. Comparing the neutralizations of Yiddish and German also provided some insight into why the 1sg may not be so privileged in related dialects.

The strategy of comparing neutralizations is rooted in a general model of paradigm acquisition, developed in Albright (in progress). The premise is that learners must be able to produce and understand forms they have never heard before, and they do this by focusing on the part of the paradigm that reveals properties of the word as unambiguously as possible. A hypothesis of this model is that learners must select a single surface form as the base or UR, even if it does not preserve every single contrast. This is similar in spirit to a proposal by Lahiri and Dresher (1984), who suggested that learners pay more attention to nominatives when learning the morphological class of nouns; the current approach is an attempt to generalize this, and explain how learners might discover which forms to pay more attention to.

The calculations that I have been using in this paper are rather crude, but may serve as a conceptual example for a more rigorously defined, computationally implemented algorithm described in more detail in Albright (in progress). This algorithm considers each member of the paradigm as a potential base, and constructs stochastic grammars of morphological and phonological rules to derive the remaining forms. It then compares how "effective" these grammars are by calculating their accuracy, the number of exceptions needed, the reliability of the stochastic rules, and so on. The algorithm is shown to select the right base in not only the Yiddish case, but in others as well, including in Latin noun paradigms (Albright, to appear), Lakhota paradigm innovations, and Spanish verbs.

It is useful to compare this model against one without the single surface form restriction. Under a traditional approach, learners could notice that some contrasts (like obstruent voicing and final schwas) are seen in some forms, while other contrasts (like umlaut) are seen in others, combining multiple surface forms to create a lexical entry that captures *all* unpredictable information. By comparing 1sg and 3sg forms, for example, learners could set up an underlying distinction between non-alternating forms (/pɔr-/ 'match') and alternating forms (/fɔr-/~/fer-/, or /fɔr-/$_{[+\text{umlaut}]}$ 'travel'). Under this model, there are various possible sources of error. A learner could have incomplete information about a word, failing to learn or recall that it has both [ɔ] and [e] allomorphs (or that it undergoes the *umlaut* rule), and incorrectly produce 2sg **fɔrst* without *umlaut*. A speaker could forget or not know that the singular of a particular *Wechselflexion* verb uses a different root allomorph, and incorrectly extend the plural vowel, producing 1sg **vis* instead of *veys*. There are few formal models of how learners learn URs and reason about only partially known words, and it is difficult to make exact predictions without one. However, the general point is this: if learners can construct URs from multiple parts of the paradigm, then we expect different verbs could potentially level to different parts of the paradigm. A traditional model does not explain why contrasts preserved in a particular surface form (the 1sg) were consistently maintained, while contrasts neutralized in that form were systematically lost.

The single surface base restriction, on the other hand, prevents learners from storing absolutely all unpredictable information in the UR, and requires more forms to be listed as exceptional. The learning procedure mitigates this problem by finding the URs and rules that minimize the number of stored exceptions, by selecting the base form that preserves the most contrasts. The prediction is that contrasts preserved in the base will be maintained, while contrasts neutralized in the base will be open for leveling – which, in the case of Yiddish, appears to be correct.

References

Albright, A. (in progress). *The identification of bases in paradigms: A computational approach.* Ph. D. thesis, UCLA.

Albright, A. (to appear). The morphological basis of paradigm leveling. In *Proceedings of the ZAS Conference on Paradigm Uniformity, March 2001.*

Bybee, J. (1985). *Morphology: A study of the relation between meaning and form.* Amsterdam: John Benjamins Publishing Company.

Dammers, U., W. Hoffmann, and H.-J. Solms (1988). *Grammatik des Frühneuhochdeutschen, Vol. 4: Flexion der starken und schwachen Verben.* Heidelberg: Carl Winter Universitätsverlag.

Geyer, A. (1912). *Die starke Konjugation bei Johann Fischart.* Hofbuchdruckerei C. A. Kaemmerer & Co.

Kenstowicz, M. (1995). Base-identity and uniform exponence: Alternatives to cyclicity. In J. Durand and B. Laks (Eds.), *Current Trends in Phonology: Models and Methods.* CNRS and University of Salford Publications.

Kern, P. O. (1903). *Das starke Verb bei Grimmelshausen: Ein Beitrag zur Grammatik des Frühneuhochdeutschen.* Chicago: University of Chicago.

King, R. D. (1980). The history of final devoicing in Yiddish. In M. I. Herzog, B. Kirshenblatt-Gimblett, D. Miron, and R. Wisse (Eds.), *The Field of Yiddish: Studies in Language, Folklore, and Literature, Fourth Collection*, pp. 371–430. Philadelphia: Institute for the Study of Human Issues.

Kiparsky, P. (1965). *Phonological Change.* Ph. D. thesis, MIT.

Kuryłowicz, J. (1947). The nature of the so-called analogical processes. *Diachronica 12*(1):113–145. (trans. Margaret Winters).

Lahiri, A. and B. E. Dresher (1984). Diachronic and synchronic implications of declension shifts. *The Linguistic Review 3*:141–163.

Mańczak, W. (1958). Tendences générales des changements analogiques. *Lingua 7*:298–325 and 387–420.

Paul, H. (1920). *Prinzipien der Sprachgeschichte* (5th ed.). Halle: Niemeyer.

Paul, H., P. Wiehl, and S. Grosse (1989). *Mittelhochdeutsche Grammatik* (23 ed.). Tübingen: Max Niemeyer Verlag.

Philipp, G. (1980). *Einführung ins Frühneuhochdeutsche: Sprachgeschichte, Grammatik, Texte.* Heidelberg: Quelle und Meyer.

Rockowitz, A. (1979). *201 Yiddish Verbs, fully conjugated in all the tenses.* Woodbury, NY: Barron's Education Series, Inc.

Schirmunski, V. (1962). *Deutsche Mundartkunde.* Berlin: Akademie-Verlag.

Steriade, D. (2000). Paradigm Uniformity and the Phonetics/Phonology boundary. In J. Pierrehumbert and M. Broe (Eds.), *Papers in Laboratory Phonology, vol. 6.* Cambridge University Press.

Tiersma, P. (1982). Local and general markedness. *Language 58*(4):832–849.

Weinreich, U. (1990). *Modern English-Yiddish, Yiddish-English Dictionary.* New York: Yivo Institute for Jewish Research.

Zehetner, L. (1989). *Bairisches Deutsch.* Hugendubel.

Dept. of Linguistics, UCLA
405 Hilgard Ave.
Los Angeles, CA 90095-1543

albright@ling.ucsc.edu

The Possessive NP Construction: Discourse Function and Discourse Profile

MIRA ARIEL
Tel Aviv University

1. Motivating Discourse Functions and Discourse Profiles

Pragmatists and discourse grammarians have benefited tremendously from examining naturally occurring data. An examination of all and only the **necessary** discoursal conditions obtaining when a certain linguistic expression occurs has paved the way for pragmatists to describe the appropriateness conditions placed on linguistic forms. These were identified as the **discourse functions** of those linguistic expressions. For example, based on an empirical study of English *it*- and *wh*-clefts in discourse, Prince (1978) defined distinct discourse functions for *it*-clefts and *wh*-clefts. She argued that the discourse function of the embedded clause in *it*-clefts is to mark "known (factual) information" and that of *wh*-clefts is to mark information "Given (currently available) to the addressee". These discourse functions should hold true in (virtually) every case the clefts are used.

Discourse grammarians broadened the definition of what discoursal factors count as relevant to the usage of a linguistic form. They often refer to **all,** or at least more than the minimally necessary discourse conditions obtaining when a certain linguistic expression occurs, provided these conditions repeat themselves frequently. The goal here is to characterize the **prototypical** (i.e., frequent) **discourse profile** of the linguistic expression at hand (Cf. Du Bois 1988).[1] This profile actually falls short of a full account for the form in all of its uses, since some of the conditions enumerated are not (currently) obligatory. The main rationale for this type of analysis has been that grammar emerges out of recurrent discoursal patterns (Hopper and Thompson 1980, Du Bois 1987, Hopper 1987, Comrie 1994, Bybee 1998, 2001, Bybee and Thompson 2000, Traugott and Dasher 2002). If two (or more) phenomena are concomitant in many instances (a specific form and specific contextual condition(s), in our case), they constitute a recurrent discourse pattern. And they may come to be obligatorily (grammatically) associated with each other, even if, initially, the specific

[1] I prefer 'prototypical' over 'frequent', because the prototypical discourse profile is a construct based on a number of frequent features, whose frequencies vary, however.

contextual factor was in fact incidental. This analysis explains how originally nonobligatory, even marginal contextual features, may come to form part and parcel of the discourse function of a particular linguistic expression (see Bates and MacWhinney 1989: 21-2). For such a linguistic expression, the **prototypical discourse profile** has become a grammatically specified **discourse function**.

For example, *while,* initially (and necessarily) used to code simultaneity, was most probably repeatedly used when the states described were also in contrast with each other (an optional contextual condition). This explains why we now have a separate concessive meaning (function) for *while*, where the original simultaneity requirement has actually been obliterated (see Traugott and König 1991, and see also Faltz 1989). The prototypical discourse occurrence in which the states of affairs were simultaneous **and** contrastive was reinterpreted to involve only contrast. Similarly, probably because most NPs in discourse are Given, many of them are modified by definite articles. This high frequency may bring about an increase in the contexts where the definite article is used. Once it extends to generics, proper names, etc., definite NPs become even more frequent, paving the way for a universal usage of the definiteness marker with all NPs. At this point the definite article no longer codes 'definiteness'. Rather, it codes 'nominality' (see Greenberg 1978). This is a case where the original discourse function (Givenness) was obliterated completely in favor of the prototypical discourse profile (noun modification), which replaced it. In addition, optional discoursal factors may also constitute crucial conditions for formal, rather than semantic changes in linguistic expressions, again, provided they cooccur with the forms frequently enough (see Ariel 2000a, Bybee and Thompson 2000, Bybee 2001).

There is thus no doubt that the prototypical discourse profile of linguistic forms is crucial in explaining how language works: Both functions and forms are modified by context, and specifically, by the prototypical context the form is used in. However, we must be careful to distinguish between the (nongrammaticized) prototypical discourse profile of form X (those conditions which prototypically but partly optionally accompany its use) and the discourse function(s) of X (those conditions which **must be** met for the form to occur properly). The discourse profile of an expression is a relatively concrete and detailed description of the conditions obtaining in its prototypical context. Its discourse function is more of an abstraction, since it must account for all of its (current) uses. It lists all but only the necessary conditions placed on its use (in any context).

Which of these characterizations is relevant for speakers? Focusing on the prenominal possessive NP construction (e.g., *his decision*), I will argue that both discourse functions and prototypical discourse profiles play a role in language use. The main body of data used here is a slightly edited first person narrative from Morris (1994: 129-139), which contains 133 possessive NPs. Additional data were drawn from the Santa Barbara Corpus of Spoken American English (Du Bois 2000).

Unfortunately, it is not always an easy matter to draw the distinction between conditions which are only part of the prototypical discourse profile and conditions obligatorily placed as discourse functions. Discourse functions too are often based on empirical findings which, although statistically very significant, rarely amount to 100% of the cases. If so, how can one tell if the identified variables constitute a necessary part of the discourse function of some linguistic form, or merely frequent but nonobligatory aspects of its occurrence, i.e., part of its prototypical discourse profile? To make things worse, in arguing for a particular discourse **function**, linguists have often quoted empirical results for particular discourse **profiles**. For instance, full proper names (e.g., *Charles Dickens*) tend to occur as first mentions (64%, data here based on Ariel 1990: 42-45). When discourse anaphoric, full proper names tend to be coreferent with antecedents which last appeared in a previous paragraph (16%), or at least two clauses away from the current mention within the same paragraph (8%). Do these findings attest to three distinct discourse profiles, or should we abstract away from the detailed profiles to derive one generalization about a discourse function of low accessibility? I have opted for the latter characterization.

While I have adopted the discourse function approach to referring expressions, Fox (1987) chose the discourse profile approach, concentrating on prototypical uses, emphasizing, moreover, the different discourse profiles the same referring expressions (definite descriptions and pronouns) have in different genres (see also Jucker 1996, Lord and Dahlgren 1997). Such analyses are in line with recent emphasis on low-level generalizations, rather than highly abstract, over-generating ones (see Bybee 1998, 2001, Thompson 2001). Do we use more concrete knowledge about recurrent patterns (i.e., discourse profiles) or do we rely on the most abstract discourse functions? I will suggest that both representations play a role in language use.

2. The Prenominal Possessive NP Construction: Prototypical Discourse Profile and Discourse Function

According to Langacker (1991: 170) and Taylor (1994, 1996), the anchor (the "possessor") in a possessive NP provides a reference point for the anchored (the "possessee"). This requirement should result in specific information statuses for the two entities. Section 2 analyzes the prototypical discourse profile (2.1) and the discourse functions (2.2) of the NPs in English prenominal possessive NPs. I show that characterizing possessive NPs according to their prototypical discourse profile accounts for only the majority of their occurrences, whereas characterizing them according to the discourse functions of the anchor and the anchored NPs enables us to generalize over (virtually) 100% of their occurrences. Nonetheless, in view of grammaticization patterns along prototypical discourse profiles, I argue that discourse profiles too must be represented by the native speaker (2.3).

2.1 The Discourse Profile of English Prenominal Possessive NPs

Here is an example with two prenominal possessive NPs from my data (I indicate in parentheses the page number and number I assigned the example):

(1) In 1982, graffiti referring to **my sexual orientation** appeared on **the men's room** wall (133: 116a, 117).

The referent of the anchor (*my*, *the men* in (1)) is invariably human (all 133 occurrences, 100% (see also Taylor 1996, Chappell and Thompson 1992). It is very often singular (110, 82.7%, and cf. with Taylor's (1996) 84%). For the most part, the anchor also represents Given information previously mentioned in the current paragraph (116 cases, 87.2% -- see also Brown 1983 and Taylor 1996). The anchor tends to be globally salient in the discourse (mentioned at least 10 times, in 87 cases, 65.4%). It is therefore prototypically coded by pronouns (121, 90.1%, 75% in Taylor's data), which refer to highly accessible entities (see Chafe 1976, Sanford and Garrod 1981, Givón 1983, Ariel 1990, inter alia).

The prototypical anchored entity is the mirror image of the anchor. It refers to nonhumans (96 cases, 72.2%, see also Brown 1983), and it presents new information (90 cases, 67.7%, see also Brown 1983, Givón 1983, Chappell and Thompson 1992). These new entities are, however, relatively easily inferable. Addressees rely on general stereotypic assumptions (e.g., people have offices, families, body parts, etc., hence, *my office* 133: 91; *my mother* 129: 8; *my head* 132: 73), or on assumptions appropriate for the specific discourse (academics have promotions, hence, *my promotion to full professor* 133:118). No wonder it is coded by a lexical NP (133 cases, 100%): Definite lexical NPs are used for entities which are accessible, but at a relatively low degree of accessibility.[2] The newness of the anchored, which is the head noun in possessive NPs, is compatible with the finding that possessive NPs tend to be coded as nonsubjects (86 cases, 64.7%, 87% in Brown's 1983 data). The anchored NP does not constitute a globally salient entity in the discourse (at most, it is mentioned once or twice again in the narrative in 116 cases, 87.2%). It is not even locally prominent. Hardly ever is it referred to again in definite form without the same anchor in the same paragraph (5, 3.8%, see also Brown 1983, but see Givón 1983 for different results).

Thus, the prototypical possessive NP is a nonsubject which introduces a new but inferable, transient, and relatively noncentral discourse entity by anchoring it to a highly accessible, prominent, human discourse entity (see Deane 1987).

2.2 The Discourse Functions of the NPs in Prenominal Possessive NPs

The discourse function approach to possessive NPs is that both anchors and anchoreds are selected according to Accessibility Theory (Ariel 1990, 2001). According to Accessibility Theory, speakers signal to the addressees how accessible the intended mental entity is by choosing a referring expression according to the specific degree of mental accessibility it codes. The form-function correlation between referring expressions and degree of accessibility is

[2] Note that it is not grammatically obligatory to have a lexical NP in the anchored position, in view of expressions such as *mine*, *Maya's (one)*.

highly motivated. Informative, rigid (roughly, uniquely referring) and phonologically prominent (long, or stressed) referring expressions, (e.g., definite descriptions) code lower degrees of accessibility, while uninformative, nonrigid, and phonologically attenuated forms (e.g., pronouns) code higher accessibility.

It is clear that the prototypical cases abide by Accessibility Theory. If the anchors have close-by antecedents, a high accessibility marker (a pronoun) is called for. And if the anchored NP presents a new inferable entity, then a low accessibility marker (a definite lexical NP) is called for. The crucial point is that the nonprototypical cases obey Accessibility Theory predictions just as prototypical cases do. In Ariel 2002 I have demonstrated in detail how each exception to the prototypical pattern can be accounted for if all we require is that each of the NPs involved correctly code the degree of accessibility with which the entity is entertained, without imposing any restrictions on the information status of the entities: Lexical anchors are used for anchors which are not highly accessible (mostly because they were last mentioned in a previous paragraph-- the distance criterion, or because of a competing antecedent). Indeed, the anchor entity may not be Given at all, provided an indefinite NP is chosen, as in:

(2) When **someone**'s personal and professional reputation has been besmirched... (138: 326a).

The same applies to nonprototypical anchored NPs. For example, even when Given, rather than New, the entities referred to are still entertained at a relatively low degree of accessibility (according to the distance and/or competition criteria), and hence the use of a lexical NP (understood as definite). Also, contra the prototype, over a quarter of the anchored NPs in the data refer to humans, and a few of them are even relatively prominent. Indeed, all persisting anchoreds in my data were human (for a more detailed analysis, see Ariel 2002). We cannot therefore maintain the prototype of anchored NPs (new, transient, nonhuman, nonsalient entities) as part of the generalization about their discourse function(s).

However, once we replace the detailed, lower order generalizations (the discourse profile) of possessive NPs with a more abstract definition of their discourse function, we can reach a characterization which is true in 100% of the cases. In fact, the findings indicate that in terms of their discourse functions, there is nothing special about the NPs in the possessive NP. Both the anchor and the anchored obey the form-function correlations established by Accessibility Theory, whereby relatively short and uninformative forms code a relatively high degree of accessibility and full/informative forms code a relatively low degree of accessibility.[3] In fact, while I have contrasted pronouns and lexical NPs here, Cohen 2002 analyzes similar more subtle phonological alternations reflecting degrees of accessibility (clitic Vs. pronominal possessives in Hebrew).

[3] But, of course, the requirement that some asymmetric association obtain between the two entities referred to remains.

2.3 Grammaticization Patterns of Possessive NPs

The arguments presented in section 2.2 suggest that the discourse function approach is sufficient, so that the prototypical discourse profile seems redundant. Formal and semantic grammaticization patterns, however, attest to the reality of at least the prototypical discourse profile. Taylor (1994, 1996) accounts for possible and impossible "objective" readings of possessive NPs as deriving from the topicality characteristic of anchors. Indeed, aspects of prototypical discourse profiles may have semantic and formal repercussions. The identity in form between ergatives and possessors in many languages (see Blake 1994: 151/2) is explained by Du Bois (1988) as deriving from the similarity between the **discourse profiles** of anchors/possessors and ergatives – both are prototypically human, Given and coded nonlexically. We could similarly propose that the similarity between the discourse profiles of grammatical patients (direct objects) and anchored entities (both predominantly nonhuman) motivates the assignment of the same case to both in some languages (see Blake 1994: 152).

Comrie (1981: 41) notes that high animacy, a prototypical discourse profile for anchors, tends to result in a preference for a certain structure of possessive NPs in Yidini, where two alternative possessive NPs are available. The high frequency of pronominal anchors we have seen is most probably responsible for the obligatory affixation of pronominal possessives in very many languages. Hebrew, for example, forces pronominal possessors (of one of its possessive constructions) to reduce to affixes, rendering **beit hi* 'house-construct state; 3rd fem pronoun' ungrammatical, as opposed to *beit+a*, 'house+3rd fem affix', 'her house' and *beit Maya*—'house Maya', 'Maya's house'). Italian forces a prenominal (rather than postnominal) possessive construction when the anchor is pronominal (Giorgi and Longobardi 1991). The high frequency of inalienable anchoreds (including kinship terms) in possessive NPs must have paved the way for different constructions to specialize for alienable and inalienable anchoreds, the latter being less marked than the former (see Haiman 1985, Croft 1990, Heine and Lébikaza 1997, and the articles in Chappell and McGregor 1996).

Similar cases of branching off of a general construction into a set of distinct, lower-level constructions with their own idiomatic characteristics and discourse profiles form the basis for Bybee and Thompson's (Bybee 1998, 2001, Bybee and Thompson 2000, Thompson 2001) hypothesis that constructions are much less general than has been thought. Indeed, contra Prince's 1978 claim for a general, discourse function for *wh*-clefts, Hopper (2001) and Hopper and Thompson (2001) argue for the formulaity of *wh*-clefts in conversations (88% of them restricted to 3 verbs). This is not only a much more concrete (discourse profile) characterization of *wh*-clefts. It also intentionally avoids the overgeneration of the abstract (discourse function) characterization (the claim is that the construction is rather unproductive in fact).

In English, and in many other languages, possessive NPs force a definiteness reading on both NPs (see Lyons 1999: 23-5, Haspelmath 1999), even though the definite article is obligatorily missing for the anchored NP. No doubt this derives from the prototypical discourse profile of possessive NPs, where for the most part, both entities are accessible (and therefore definite). It is this discourse profile of the entities involved in possessive NPs which may have given rise to a grammatical freezing of definiteness as the only option for possessive NPs. As Haspelmath (1999) describes it, however, the facts regarding the (non)marking of definiteness for the anchored NP are more complicated, and at the same time, highly instructive for the issue at hand. It is not just an anti-redundancy principle which is at work. The notion of an entrenched construction is heavily implicated. It seems that the relative chronology of the entrance to the language of the definite marker and the possessive construction plays a crucial role in predicting whether definiteness will redundantly be coded or not. Often, it is languages where the definite marker evolved later than the possessive construction which tend not to code the definiteness of the anchored NP. This phenomenon can be explained by analyzing the possessive construction (especially the prototypical pronoun + possessive marker + NP -- the order is language-specific, of course) as a construction stored in the lexicon, in the spirit of Bybee 2001 and Thompson 2001. Such constructions are said to have a storage and processing status for speakers, who tend to access them as inseparable chunks. Languages where the definite marker appeared later tended not to add the definite marker, for the possessive construction must have already been entrenched as a unit.

Similarly, languages where the definite marker had already been in use may have entrenched the definite-noun construction, so that when a possessive construction appeared, even though it would anyway be interpreted as definite in an overwhelming majority of the cases and the definite marking is therefore redundant, the entrenched pattern of definite marker + noun remained intact. Thus, the working of entrenched low-level constructions can explain why some languages manifest a "redundant" marking of definiteness in possessive constructions, while others do not. Such entrenchments demonstrate the important role of the prototypical discourse profile in language use and grammaticizations.

The prototypical discourse profile of possessive NPs can then become partially grammaticized. Since grammaticizations must happen in real discourse time, we must assume that speakers are operating with low-level form-function correlations, even though they are redundant and only optional in the early stages.

3. Implications for the Competence for Constructions

Both the form and the semantics of the prototypical possessive NP show that prototypical discourse profiles must be available to speakers (2.3). Still, the totality of language use can only be accounted for by reference to the discourse function (2.2). What then is the relevant level of representation for constructions? I suggest speakers simultaneously operate with discourse functions and discourse profiles.

First, note that in some cases, only low-level generalizations are in fact possible, for no single discourse function can account for the data. Thus, Chappell and Thompson (1992) are only in a position to argue for tendencies for either overt or lack of overt possessive marking in specific discourse profiles in Chinese (e.g., first person anchors + kin anchoreds are unmarked in 89% of the cases). Are significant but not absolute form-function correlations (say, 60-90%) to be ignored? Such discourse profiles seem to play a role in language use, and hence, merit representation, even if they only constitute patterned tendencies.

In fact, it is conceivable that the human mind does not actually operate with the highest generalizations that linguists can abstract away from the data. Bates et al (in press) emphasize the human ability to learn statistically significant patterns. Lakoff and Thompson (1975: 295) have claimed that "abstract grammars do not have any separate mental reality; they are just convenient fictions for representing processing strategies" (see also Langacker 1987, Fillmore et al 1988, Comrie 1994). Indeed, some recent research has emphasized the role of memorization in language competence. Once restricted to the lexicon, frequency-driven memorized formulas are now seen as replacing at least some phonological, morphological and even syntactic generalizations (Chafe 1992, Tannen 1989, Bybee 1998, Bybee and Thompson 2000, Bybee 2001, Hopper and Thompson 2001, Thompson 2001, Thompson and Hopper 2001, Thompson to appear).

Note that the low-level discourse profile approach entails a number of discourse profiles per linguistic construction (*a la* Fox 1987). Novel uses are then seen as analogical extensions of core uses (Lakoff 1987). However, if one wishes to do away with the discourse function generalization, one should demonstrate that the analogy works independently of the higher discourse function. I find it hard to conceive of an analogical pathway connecting the various discourse profiles of referential forms which does not involve relying on the general discourse function of coding specific degrees of accessibility. Discourse functions are then functional here too.

Polysemies involve a similar dilemma. Griceans have opted for one core meaning, deriving related meanings by inference, following Grice's 1978 Modified Occam's Razor: "senses are not to be multiplied beyond necessity". This is akin to the discourse function account. Traugott and Dasher (2002), however, propose that polysemous meanings are only partly derived from one abstract meaning. Partly, polysemous meanings are directly retrieved from "redundantly" specified lexical entries, akin to the redundant discourse profile(s). This "mixed" model of high and low-level generalizations seems right for constructions as well.

According to Israel 1996, the *one's way* construction (as in *Rasselas dug his way out of the Happy Valley*-- Israel's 1996 ex. 1) was initially restricted to a few high frequency motion verbs (*go*, *ride*, *run*), but it was not restricted to *way* (*street* and *path* are also attested). The restriction on verbs applied for a few centuries, justifying the assumption of low-level constructions privileging the specific verbs, but it has now generalized to accommodate any verb which can be construed as describing motion, justifying a general schema. Note, moreover, that whereas the verb category has been abstracted into a general discourse function, the direct object category has narrowed down to *way*, showing sensitivity to the prototypical discourse profile of the idiom. Speakers, then, shift in both directions: abstracting away from the discourse profile schema (to include many more verbs under a general discourse function) as well as narrowing a general schema (to the lexical noun *way* alone, due to a recurrent discourse profile). Such shifts in opposite directions for the same idiom demonstrate the simultaneous working of both representation levels.

Thompson (to appear), analyzing what are traditionally termed object complements in English (e.g., *I thought she might pull it out of the garbage* -- ex. (1)), also distinguishes between the majority cases accounted for by low-level discourse profiles (various formulaic constructions based on a few high frequency matrix verbs, all of which occur with first person singular subjects and no overt complementizer) and the nonformulaic minority cases where various verbal expressions occur (e.g., *we don't bother to ask the American people...* -- ex. (44). The latter are rare, but grammatical and useful nonetheless. Thus, entrenched formulaic subconstructions are probably stored alongside more general schemas. The fact that the majority cases can actually be accounted for by the higher-level schema does not render the representation of formulaic construction(s) redundant. Assuming such a double representation of both the abstract discourse function and a set of prototypical discourse profiles can account for the seemingly surprising fact that languages may share form-function correlations at the discourse function level, but not at the discourse profile level. For example, Delahunty and Gatzkiewicz (2000) identify one discourse function for the *'It's that...'* construction, but not all their Spanish examples are acceptable in other languages.

The most plausible option seems to be the least "elegant": speakers store both the abstract discourse function of the construction, and its variety of discourse profiles, or at least its prototypical discourse profiles. While the high level generalization about linguistic forms informs some of their uses, speakers often rely on local, conventionalized strategies when producing or understanding linguistic forms. As Langacker (1988) has argued, grammar contains both highly general schemas and redundant, idiosyncratic formulas (see Israel 1996 for an analysis in this spirit of the *way* construction). A reasonable variant of the universal "double storage" alternative is that whereas the abstract discourse function is relevant for infrequent linguistic uses, frequent uses are more automatically governed by reference to their prototypical discourse profiles. This would be in line with Bybee and Thompson's (2000: 381) morphological and syntactic claim that "high frequency sequences take on a life of their own." Thus, entrenchment makes such forms resistant to regularization, for example.

I doubt that linguistic usage alone can decide between these alternatives. The arguments in section 2.2 argue for a discourse function representation, while the findings in 2.3 support a discourse profile model of competence. I therefore propose we further examine the "double storage" option whereby language users represent both discourse functions and discourse profiles for constructions. I here see a clear role for psycholinguistic research.

References

Ariel, Mira. 1990. *Accessing Noun Phrase antecedents*. London: Routledge.

Ariel, Mira. 2000a. The development of person agreement markers: From pronouns to higher accessibility markers. In M. Barlow & S. Kemmer (eds.) *Usage-based models of language.* Stanford: CSLI.

Ariel, Mira. 2000b. Accessibility marking: One discourse function or genre-dependent referential styles? Ms., UC Santa Barbara.

Ariel, Mira. 2001. Accessibility Theory: An overview. In T. Sanders, J. Schilperoord & W. Spooren (eds.) *Text representation: Linguistic and psycholinguistic aspects.* Amsterdam: John Benjamins.

Ariel, Mira. 2002. Accessibility marking: Discourse functions, discourse profiles and processing cues. Ms., Tel Aviv University.

Bates, Elizabeth & Brian MacWhinney. 1989. Functionalism and the competition model. In B. MacWhinney & E. Bates (eds.) *The crosslinguistic study of sentence processing.* Cambridge: Cambridge University Press.

Blake, Barry J. 1994. *Case.* Cambridge: Cambridge University Press.

Brown, Cheryl. 1983. Topic continuity in written English narrative. In Givón, T. (ed.) *Topic continuity in discourse.* Amsterdam: John Benjamins.

Bybee, Joan. 1998. The emergent lexicon. *CLS* **34**: 421-35.

Bybee, Joan. 2001. Frequency effects on French Liaison. In Bybee & Hopper (eds.) *Frequency and the emergence of linguistic structure.* Amsterdam: John Benjamins.

Bybee, J. & P. Hopper (eds.) 2001. *Frequency and the emergence of linguistic structure.* Amsterdam: John Benjamins.

Bybee, Joan & Sandra A Thompson. 2000. Three frequency effects in syntax. *BLS* **23**: 378-88.

Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C.N. Li (ed.) *Subject and topic.* New York: Academic Press.

Chafe, Wallace L. 1992. The flow of ideas in a sample of written language. In W.C Mann & S.A. Thompson (eds.) *Discourse description: Diverse linguistic analyses of a fund-raising text.* Amsterdam: John Benjamins.

Chappell, Hilary. 1996. Inalienability and the personal domain in Mandarin Chinese discourse. In Chappell & McGregor (eds.) *The grammar of inalienability* Berlin: Mouton de Gruyter.

Chappell, H. & W. McGregor (eds.). 1996. *The grammar of inalienability.* Berlin: Mouton de Gruyter.

Chappell, Hilary & Sandra A Thompson. 1992. The semantics and pragmatics of associative *de* in Mandarin discourse. *C.L.A.O* **XXI**: 199-229.

Cohen, Evan G. 2002. The Pragmatics-Phonology Interface: Accessibility and Reduction - Hebrew Possessives. Unpublished Ms. Tel Aviv University.

Comrie, Bernard. 1994. Coreference: Between grammar and discourse. *Proceedings of the eighteenth annual meeting of the Kansai linguistic society.*

Croft, William. 1990. *Typology and universals*. Cambridge University Press.

Deane, Paul. 1987. English possessives, topicality, and the Silverstein hierarchy. *BLS* **13**: 65-76.

Delahunty, Gerald & Laura Gatzkiewicz. 2000. On the Spanish inferential construction *ser que*. *Pragmatics* **10**: 301-22.

Du Bois, John W.1987. The discourse basis of ergativity. *Language* **63**: 805-55.

Du Bois, John W. 1988. Discourse as pattern model for grammar: The possessor=ergator affiliation. Talk given at the University of Oregon.

Du Bois, John W. 2000. *Santa Barbara Corpus of Spoken American English,* Part 1. CD-ROM. Linguistic Data Consortium, University of Pennsylvania.

Faltz, Leonard M. 1989. A role for inference in meaning change. *Studies in Language* **13**: 317-31.

Fillmore, Charles, Paul Kay & Catherine O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language* **64**.

Fox, Barbara A. 1987. *Discourse structure and anaphora: Written and conversational English*. Cambridge: Cambridge University Press.

Giorgi, Alessandra & Giuseppe Longobardi. 1991. *The syntax of noun phrases*. Cambridge: Cambridge University Press.

Givón, T. 1983. Topic continuity in spoken English. In Givón, T. (ed.) *Topic continuity in discourse*. Amsterdam: John Benjamins.

Givón, T. (ed.) 1983. *Topic continuity in discourse*. Amsterdam: John Benjamins.

Greenberg, Joseph H. 1978. How does a language acquire gender markers? In J. H. Greenberg, C. A. Ferguson & E. A Moravcsik (eds.) *Universals of human language* **3**: *Word structure*. Stanford: Stanford University Press.

Grice, H.P. 1978. Further notes on logic and conversation. In P. Cole (ed.) *Syntax and Semantics* **9**: *Pragmatics*. New York: Academic Press. 41-57.

Haiman, John 1985. *Natural syntax*. Cambridge: Cambridge University Press.

Haspelmath, Martin. 1999. Explaining article-possessor complementarity: Economic motivation in noun phrase syntax. *Language* **75**: 227-43.

Heine, Bernd & Kézie Lébikaza. 1997. On attributive possession in Kabiye. In J. Bybee, J. Haiman & S. Thompson (eds.) *Essays on language function and language type*. Amsterdam: John Benjamins.

Hopper, Paul J. 1987. Emergent grammar. *BLS* **13**: 139-55.

Hopper, Paul J. 2001. Grammatical constructions and their discourse origins: prototype or family resemblance? In M. Putz, S. Niemeier & R. Dirven (eds.) *Applied cognitive linguistics I: Theory and language acquisition*. Berlin: Mouton de Gruyter.

Hopper, Paul J. & Sandra A Thompson. 1980. Transitivity in grammar and discourse. *Language* **56**: 251-99.

Hopper, Paul J. & Sandra A Thompson. 2001. Grammatical fragments and social action in conversation. Talk given at the International conference on Cognitive Linguistics, UC Santa Barbara.

Israel, Michael. 1996. The *way* constructions grow. In A. Goldberg (ed.) *Conceptual structure, discourse and language*. Stanford: CSLI.

Lakoff, George. 1987. *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago: Chicago University Press.

Lakoff, George & Henry Thompson. 1975. Introducing cognitive grammar. *BLS* **1**: 295-313.

Langacker, Ronald W. 1987. *Foundations of cognitive grammar*, vol. 1. Stanford: Stanford University Press.

Langacker, Ronald W. 1988. A usage-based model. In B. Rudzka-Ostyn, (ed.) *Topics in Cognitive Linguistics*. Amsterdam: John Benjamins.

Langacker, Ronald W. 1991. *Foundations of cognitive grammar*, vol. 2. Stanford: Stanford University Press.

Lord, Carol & Kathleen Dahlgren. 1997. Participant and event anaphora in newspaper articles. In J. Bybee, J. Haiman & S. A Thompson (eds.). *Essays on language function and language type dedicated to T. Givón.* Amsterdam: John Benjamins.

Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.

Morris, Celia. 1994. *Bearing witness*. Boston: Little, Brown and Company.

Prince, Ellen F. 1978. A comparison of WH-clefts and IT-clefts in discourse. *Language* **54**: 883-906.

Tannen, Deborah. 1989. Repetition in conversation: Toward a poetics of talk. In *Talking voices: Repetition, dialogue, and imagery in conversational discourse*. Cambridge: Cambridge University Press.

Taylor, John R. 1994. "Subjective" and "objective" readings of possessor nominals. *Cognitive Linguistics* **5**: 201-42.

Taylor, John R. 1996. *Possessives in English.* Oxford: The Clarendon Press.

Thompson, Sandra A. 2001. Constructions and conversation. Talk given at the International Cognitive Linguistics Conference in University of California at Santa Barbara.

Thompson, Sandra A. To appear. "Object complements" and conversation: towards a realistic account. *Studies in Language.*

Thompson, Sandra A. & Paul J Hopper. 2001. Transitivity, clause structure, and argument structure: Evidence from conversation. In Bybee & Hopper (eds.) *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins.

Traugott, Elizabeth Closs & Richard B. Dasher. 2002. *Regularity in semantic change*. Cambridge: Cambridge University Press.

Traugott, Elizabeth Closs & Ekkehard König. 1991. The semantics-pragmatics of grammaticalization revisited. In E. C. Traugott & B. Heine (eds.) *Approaches to grammaticalization*, vol. 1. Amsterdam: John Benjamins.

Tel Aviv University
Tel Aviv, 69978, Israel

mariel@post.tau.ac.il

FrameNet's Frames vs. Levin's Verb Classes[1]

COLLIN F. BAKER and JOSEF RUPPENHOFER
International Computer Science Institute & University of California, Berkeley

0. Overview

The classification of verbs in Levin's (1993) *English Verb Classes and Alternations: A preliminary Investigation*, on the basis of both intuitive semantic grouping and their participation in valence alternations, is often used by the NLP community as evidence of the semantic similarity of verbs (Jing & McKeown 1998; Lapata & Brew 1999; Kohl et al. 1998). In this paper, we compare the Levin classification with the work of the FrameNet project (Fillmore & Baker 2001), where words (not just verbs) are grouped according to the conceptual structures (frames) that underlie them and their combinatorial patterns are inductively derived from corpus evidence. This means that verbs grouped together in FrameNet (FN) might be semantically similar but have different (or no) alternations, and that verbs which share the same alternation might be represented in two different semantic frames.

1. Basic Comparison

Table (1) summarizes the two approaches. Note that the numerical comparison of coverage is misleading, because for Levin we are counting distinct lemmas, lumping together senses of polysemous and even homophonous words, but for FrameNet we are counting Lexical Units, which are defined as pairings of lemmas with semantic frames and thus represent separate word senses. (As usual, the problem of dividing and enumerating senses is difficult. Levin says (p. 22) that different senses of a verb will occur in different classes, but this does not always seem to be accurate (cf. Section 7.3).)

[1] We are grateful to the National Science Foundation for funding the work of the FrameNet project through two grants, IRI #9618838 "Tools for Lexicon Building" March 1997-February 2000, and ITR/HCI #0086132 "FrameNet++: An On-Line Lexical Semantic Resource and its Application to Speech and Language Technology" September 2000-August 2003. We are also grateful to the Principal Investigator of the FrameNet Project, Charles J. Fillmore, for extensive comments on this paper; any errors that remain are our own.

(1)

	FrameNet	Levin 1993
Groupings	230 semantic frames	193 verb classes
Basis	lexical semantics	argument syntax
Data Source	Corpora	linguistic literature
Coverage	2100 nouns, 1700 verbs (including multi-word expressions), 460 adjectives	3100 verbs
Results	Frame descriptions and annotated examples	Verb classes and alternations (most with descriptions)

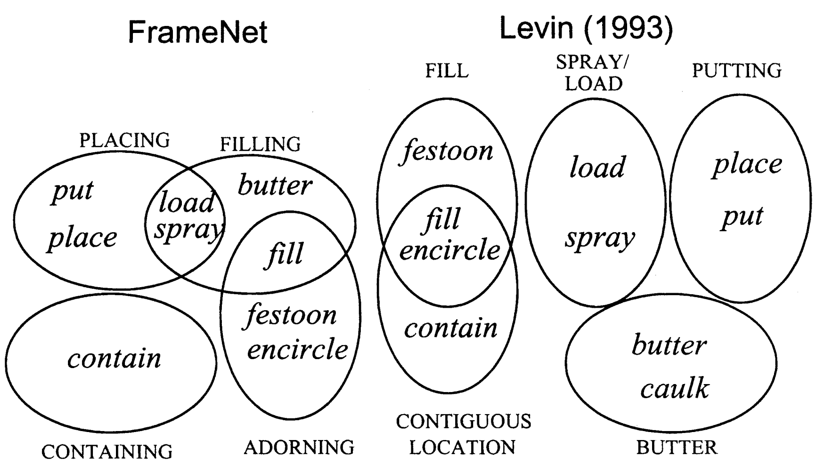
As the number of categories in the two studies is comparable (230 frames vs. 193 verb classes), one might expect that many of the FN frames would correspond to Levin's verb classes, and vice-versa. There are many partial correspondences, but there are many significant differences as well.

2. Basic Difference in Perspective

In FrameNet, predicates belong to frames based on a shared semantics. They need not exhibit all the same syntactic behaviors in order to be able to be grouped together. Thus, our frames can include alternators and non-alternators. Consider the verbs *load* and *fill*, which have long been central to the discussion of alternations.

(2)

Load, fill, and related verbs in FrameNet and Levin (1993)



According to FrameNet, *fill* and *load* are both in the Filling frame. *Load* is, additionally, in the Placing frame, while *fill* is also listed in the Adorning frame.

This reflects the facts (both syntactic and semantic) that Filling is causative (Theme-Object) and Adorning (Theme-Subject) is not.

Since alternation patterns are criterial in Levin's system, alternators and non-alternators cannot be in the same verb class, and the interchangeability between *fill* and *load* in the syntactic pattern exemplified in examples (3) and (4) is therefore not captured. (All examples are from the British National Corpus (BNC) unless otherwise noted.)

(3) While the shop assistant helped another customer and the children played, the 2 adults filled their pockets with rings and other valuables.
(4) He loaded the barrow with paving stones before running straight through a plate glass window at the B superstore in South Shields.

Likewise, Levin has a separate class for Butter Verbs, noting that they "all have zero-related nominals; their meaning can be paraphrased as 'put X on/in something'" but are otherwise similar to the Spray/Load Verbs. FrameNet records that, for certain verbs in this frame, information about the Theme is incorporated in the verb's meaning but does not regard the basic semantics as different enough from Filling to justify a separate frame. Figure (2) shows the relevant frames, including a few representative words treated in both sources.

3. Using Corpus Data

FrameNet's classifications and lexical entries are based on attested corpus examples. In many of Levin's classes there are certain members for which our corpus data do not support their use by speakers in the constructions that Levin predicts they should occur in. For instance, her Verbs of Instrument of Communication (*telex, wire, semaphore, phone, telephone, cable, telegraph, radio, fax*) are said to be able to occur as parentheticals in indirect quotations, e.g., *The winner, Heather cabled (Sara), would be announced tonight*. The verb *cable*, however, is the only one on the list attested in this construction in the BNC.

Let us look in detail at the verb *telephone*. Corpus study shows that some uses in the BNC match Levin's predictions:

(5) In October 1944, Mr. Argles telephoned the Birmingham office and said that his wife was severely indisposed.
(6) She might have backed off, gone into the pub and telephoned for a cab.
(7) [She] telephoned the young woman's mother to come …
(8) The following day, Moira telephoned the Daily Telegraph with profuse apologies for the misunderstanding.
(9) … you should telephone your flight-plan to Lisbon …
(10) My Chief-of-Staff telephoned to me that the attack had failed and that everywhere our men could be seen falling back.

Other uses are not discussed at all by Levin:

(11) Anyone who can help financially should telephone Clacton (0255) 426801

And some uses predicted by Levin are not attested at all:[2]

(12) ?Mom telephoned me the good news.
(13) ?Mom telephoned me that she was ill.
(14) ??My brother, mom had telephoned me, was now in the hospital.

4. Comparison of Groupings

Although the whole thesis of Levin's work is that grouping words according to alternations tends to produce semantically coherent classes, it can also split words that are close in meaning, or lump semantically disparate words. In this section, we will discuss examples of four types of mismatch with FrameNet frames.

4.1. Levin class roughly equivalent to FN frame

(15)

Levin Cooking Verbs	FrameNet Apply_heat Frame
bake, barbecue, blanch, boil, braise, broil, brown, charbroil, charcoal-broil, coddle, cook, French fry, fry, grill, hard-boil, heat, microwave, oven-fry, oven-poach, over-cook, pan-broil, pan-fry, parboil, parch, perk, plank, poach, pot-roast, rissole, roast, sauté, scald, scallop, shirr, simmer, steam, stew, stir-fry, toast, ...	*baste, blanch, boil, braise, broil, brown, char, coddle, fry, grill, microwave, parboil, poach, roast, sauté, scald, simmer, steam, steep, stew, toast, ...*

Levin defines this class partially on the basis of the three alternations:

(16)

Causative/ Inchoative	*Jan is baking the potatoes*	*The potatoes are baking*
Middle	*Jan baked Idaho potatoes*	*Idaho potatoes bake beautifully*
Instrument Subject	*Jan bakes the potatoes in that oven*	*That oven bakes potatoes well*

However, some of these alternates are rare; for example, of the 142 annotated examples of verbs in the Apply_heat frame in which the frame element Food appears, none permit a Middle interpretation. (But we should be cautious in ascribing significance to the number of annotated examples, see below).

[2] Some of these may be judged grammatical by some speakers, but they evidently are not found among the 1,200 examples of the verb *telephone* in the BNC.

4.2. Levin class narrower than FN frame

(17)

Levin (1993) Classes	FrameNet Placing frame
Pocket Verbs: *archive, bag, bank, …* Put Verbs: *mount, place, put, …*	*archive, bag, bank, mount, place, put, …*

The verbs of Putting and Placing are divided up by Levin into those which are morphologically related to a noun denoting the goal location and others for which this is not the case. FrameNet does not make such a distinction, given that these verbs do not incorporate a referential argument. Rather we assume that the incorporated arguments are interpreted like indefinite null instantiations in the sense of Fillmore 1986. Indeed, cognate location phrases are not categorically ruled out but are possible when more specific information is to be given about the goal location:

(18) The vinegar is then bottled in the traditional flask and sealed with a cork to preserve its high quality.

4.3. Levin class broader than FN frame

Reliance on syntactic alternations also leads Levin to posit some very broad, semantically very abstract classes. Examples of this are Levin's verb classes of social interaction (36): *correspond*, *marry*, and *meet* verbs. These three classes are defined syntactically by alternations reflective of the notion of reciprocality:

(19)

Collective subject NP	*The committee bantered/met*
Simple Reciprocal Alternation	*Pat bantered/met/*married with Kim* *Pat and Kim bantered/met/married*
Understood Reciprocal Object	*Pat married/met/*bantered Kim* *Pat and Kim bantered/married/met*

However, it appears that the alternations that Levin describes as characteristic of this verb class are not in fact diagnostics of reciprocality. For instance, even when the actions of the participants are not directed at each other but just jointly or simultaneously undertaken, plural (*John and Sue jogged*), coordinate (*John jogged with Sue*), or collective subjects (*The group jogged*) are acceptable. Furthermore, the encoding of one argument slot by a reciprocal is available with events that are not inherently reciprocal (*Larry and Moe looked at each other*).

Thus, verbs of social interaction, in so far as they are understood as involving reciprocal action of the participants, cannot be identified with the help of the above constructions. Proposing a more general verb class that includes all verbs denoting necessarily or optionally reciprocal events would result in a verb class of enormous size and semantic diversity. Clearly, the investigation in this case has to start from the semantics rather than the syntax.

Looking now at the semantics of Levin's three proposed subclasses, we find them internally rather heterogeneous.

(20)

Correspond verbs	Marry verbs	Meet verbs
agree, argue, banter, bargain, bicker, brawl, clash, coexist, collaborate, collide, combat, commiserate, communicate, compete, concur, confabulate	*court, cuddle, date, divorce, embrace, hug, kiss, marry, muzzle, pass, pet*	*battle, box, consult, debate, fight, meet, play, visit*

In our treatment of Levin's Correspond verbs, *argue*, for instance, is assigned to a Communication_conversation frame along with *bicker*, *chat*, *gossip* etc., whereas *struggle* is placed in the Hostile_encounter frame. Similarly, FN puts the verbs in Levin's Marry class (36.2) into different frames. *Marry*, *date*, *court*, *divorce* belong to the Personal_relationship frame along with other words such as *bachelor*, *boyfriend*, *break-up*, *wife*, *woo*. The remaining words in Levin's class are currently not in the FN database but arguably belong to something like a Display_of_affection frame, which would also contain *caress*, *pet* (an animal).

Levin's Meet class (36.3) differs syntactically from the other two by allowing for one party to appear as a direct object (*Pat met Kim*) as well as for it to appear in a *with*-phrase (*Pat met with Kim*). However, it seems to us that the verbs Levin puts in this class do not have the same meaning in the transitive and the *with*-PP use. For instance, *box* in *I ended up boxing with him* does not necessarily involve a competition in the same way as *Tyson will box Lewis*. Meaning differences exist also with *play* and *meet* (*My son played/met with your son ≠ My son played/met your son)*. Thus, the Meet class is not a genuine separate class from the other two and the pairs of senses in the Meet class can be distributed to other frames.

Note finally that FrameNet would not put the verbs that are only optionally reciprocal into separate frames in those uses. For these cases, we would rely on mechanisms of semantic composition to yield the right kinds of interpretation.

4.4. Overlapping Groupings: Communication verbs

The verbs of communication, like the *spray/load/fill* verbs discussed above, show a more complex sort of overlapping of FrameNet and Levin classes.

Communication verbs are one area in which the strategy of grouping by a verb's unique set of alternations leads to overly narrow classes. In the case of the verb *tell*, for instance, Levin, unlike FrameNet, does not distinguish the 'order' sense in (21), where tell takes a VP_{to} complement, from the 'inform/say to' sense in (22a-b), where it takes NPs and finite clauses as a complement:

(21) Maybe the French told her to act pregnant and so lengthen her stay in Scotland!

(22) a. I tell you I'm not satisfied with that pesky voyage.
b. "Yes, I told her my theory," he said under his breath.

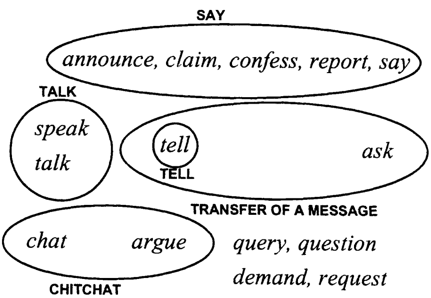
Similarly, Levin groups *speak* and *talk* into a unique class, whereas FrameNet groups the conversational uses in (23) with those of verbs like *chat* and *argue*, and the statement-like use of *speak* in (24) with those of *announce*, *claim* etc.

(23) a. He had spoken with Amelie who, though still incapacitated with her broken hip, was desperate to see Peach.
b. Would you like some coffee and then we could talk?

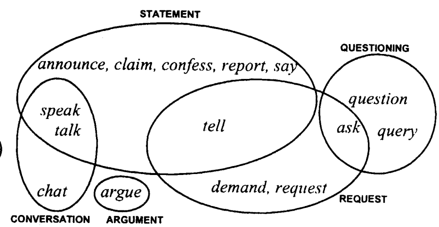
(24) 'He seems very nice,' Emily spoke guardedly, 'but what of his prospects?'

The two groupings are summarized in Figures (25) and (26). For Levin, the alternations dictate that *tell* be in a class by itself and that *speak* and *talk* be in a separate class. From our point of view, these are lexical idiosyncrasies within the semantic groupings.

(25) Communication verbs in Levin



(26) Communication verbs in FrameNet



5. Semantically meaningful Levin classes?

Levin says that "verbs in English and other languages fall into classes on the basis of shared components of meaning. The class members have in common a range of properties, including the possible expression and interpretation of their arguments, as well as the existence of certain morphologically related forms." But the meaning which is to be associated with a Levin class is often hard to define. As Dang *et al.* (1998) observe, "Of course, some Levin classes, such as *braid* (*bob, braid, brush, clip, coldcream, comb, condition, crimp, crop, curl*, etc.) are clearly not intended to be synonymous." In addition, many verbs are cross-listed in classes which pick out one aspect of their meaning but do not capture separate senses. An example of this are the hundreds of verbs found in Other Alternating Verbs of Change of State. The FN frame hierarchy allows us a more appropriate level of generalization for the facts relating to change-of-state verbs (see below).

The commitment to define separate classes of words according to their morphological make-up also causes Levin to make decisions differently from FrameNet. Thus, FrameNet could include *call* (i.e. on the telephone) and *write* in the Communication_means frame, but Levin would not, since *writing* is not zero-related, and a call is not an instrument.

We saw earlier that verbs of Instrument of Communication varied amongst themselves in respect to their valence, especially as found in the corpus, and we can see that *call* and *write* exhibit some of those same patterns, but not others. And in the same way that some of Levin's verbs have related result nouns (*wire* and *cable* but not *phone* or *telephone*), *call* has a result noun meaning but *write* does not. (*We received several important cables/wires/calls/*writes/*writings.*) Likewise, the commitment to syntactically defined alternations requires Levin to list, e.g. *track* twice, both in the Stalk verbs (35.3) and the Chase verbs (51.6); most people would probably consider this one sense, rather than two.

6. Comparison of Hierarchies

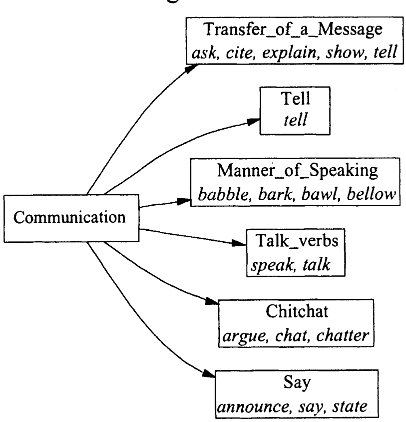
Levin's (1993) list of verb classes is divided into 51 sections, with two further levels of subdivision. The sections

> reflect a limited attempt to group verb classes related by meaning together. However, there is little hierarchical organization compared to the number of classes identified. This lack of structure reflects not only the preliminary nature of the investigation, but also the fact that it is an open research question whether a complete hierarchical organization of English verbs is possible or even desirable (see Fellbaum 1990 and Miller & Fellbaum 1991 for some discussion)" (Levin 1993:23).

As an example, Figure (27) shows the semantic relations indicated by Levin's subclasses of the Communication verbs, while Figure (28) shows the FrameNet inheritance structure in this area.

(27)

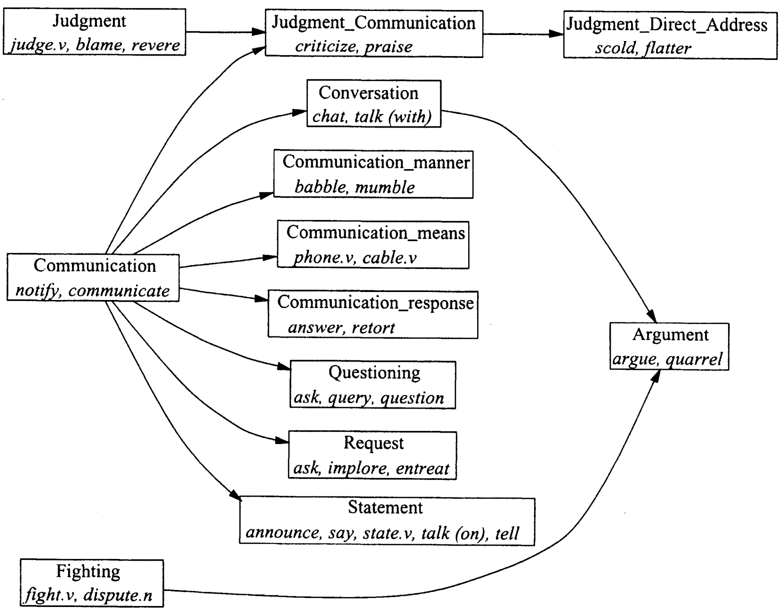
Relations among Levin's verb classes



Not only does FN define inheritance lattices relating frames, but we have a parallel, detailed set of inheritance relationships among their Frame Elements. (These are not illustrated here because the graph is quite complex.) Although the FN inheritance relations do not yet connect all frames into larger structures and although it is not clear what form such larger structure(s) might take, multiple inheritance should allow us a fuller representation of the relations among frames.

(28)

Relations among FrameNet frames



7. Challenges for both treatments

7.1. Alternations not related to lexical semantics of the verb

It is important to notice that neither Levin nor FrameNet assume that all verb syntax reflects only the inherent semantics of the verb. Of course, some alternations do reflect the underlying lexical semantics. An example is the alternation between the conative and the simple transitive construction, which differentiates, e.g., the verbs *bite* and *break*. Both verbs can occur in the transitive construction. *Bite* can also occur in the conative construction (*The dog bit at Sue*) since it does not lexicalize success in biting, only the attempt. By contrast, *break* cannot be used this way (**Bill broke at the vase*) since it lexicalizes a completed change of state. Levin depends on this alternation in many decisions about grouping.

Two constructions that do not seem to relate to inherent lexical meaning in the same way are Locative Inversion and There-insertion, shown in (29) - (31):

(29) Into the room came Harry. (Bresnan)
(30) From the speakers drones the voice of Max Von Sydow. (BNC)
(31) Out of this blur there stares a single set of eyes. (BNC)

As pointed out in much of the syntactic and discourse-pragmatic literature, these constructions are used when new referents are to be introduced into the discourse. However, we do not want to posit an 'appearance/existence' sense for all the verbs that can occur in them. Rather, we assume that the semantics of verbs belonging to frames such as Perception_noise (*drone*), Perception_active (*stare*), and Self_motion (*come*) is compatible with the discourse-pragmatic function of introducing new referents. For instance, by having only one participant, these verbs allow speakers to introduce a new referent in one clause before predicating about it in the next clause, which is the preferred discourse strategy. In short, not all argument syntax reflects lexical semantics in a narrow sense.

7.2. Frequency Data

Currently, neither approach incorporates information about the frequency of a verb's syntactic patterns in text genres. While the relevance of probabilistic knowledge is contested in the field of theoretical linguistics (despite some influence from experimental studies such as Schuetze (1996) and from the emerging field of corpus linguistics), it is clearly a useful notion for NLP (e.g., Manning (1993)).

It is important to recognize that, although one can count the annotated examples in FrameNet, they cannot be taken as representative of the frequencies of FEs or valence patterns in running text. FrameNet annotations exist for the purpose of documenting the range of syntactic possibilities for lexicographic purposes. Research on the real frequencies of FEs and their valence patterns in running text is being carried out by our associates at the University of Colorado (e.g., Roland *et al.* (2000)).

7.3. Relationships between senses

Another theoretical limitation of both the FrameNet approach and the Intersective Classes approach of Dang *et al.* (1998) is that relationships between senses of words can be characterized only by positing a more general sense that relates the two more specific ones. (Levin (1993) does not discuss this issue.) In FrameNet this is done by inheritance; in the intersective class approach, by intersection. Neither treatment is adequate for all types of sense relations. For instance, the uses shown in (32) – (35) of a word like *argue* belong to two different frames, Evidence and Statement, whose relationship is not discussed explicitly by FrameNet.

(32) ... all of which argues that only a small minority of policemen and women define their work in terms of social service.
(33) Seurat and his colleagues argued that Pointillism was a scientifically modern style,
(34) *The facts clearly announce that we should do it this way.
(35) ?The spokesperson substantiates that the company will build the plant in Pittsburgh.

One solution to this problem would be to remove the words from Statement and move them into a frame that inherits from both Statement and Evidence. This is an appealing solution, since both the notions of evidence and of communicating are simultaneously required. However, the inheritance solution would run counter to the intuition that the speaking sense of words like *argue* is more basic than the evidential one; the same intuition also holds for other word doublets: *My brother persuaded me to drive more slowly* versus *My accident persuaded me to drive more slowly.* But there are differences among words in these frames in regard to this polysemy, as shown by examples (34) and (35).

8. Conclusions

Levin (1993) demonstrated that syntactic alternations can be the basis for groupings of verbs that make some semantic sense, and that accord reasonably well with linguistic intuitions. Detailed examination of Levin's classes and alternations, as in Dang *et al.* (1998), reveals that (a) the classes are not simply the product of automatic application of a set of rules about participation in alternations, but are partially semantically motivated, and (b) a classification rigorously and solely based on alternations would give much finer distinctions, including splitting of many semantically coherent classes. The FrameNet project is producing a lexicon with roughly comparable coverage of verbs, but with much more detail concerning the semantics and syntax of their arguments, more semantically consistent categories, and a richer set of relations among them. The FrameNet lexicon also covers nouns and adjectives, using the same semantic frames.

References

Bresnan, Joan. 1994. Locative inversion and the architecture of universal grammar. *Language*. **70**:72-131.

Chafe, Wallace. 1987. Cognitive constraints on information flow. In Russell Tomlin (ed.) *Coherence and Grounding in Discourse*, 21-51. Amsterdam: John Benjamins.

Dang, Trang Hoa, Karen Kipper, Martha Palmer, & Joseph Rosenzweig. 1998. Investigating regular sense extensions based on intersective Levin classes. *Proceedings of COLING-ACL*. Montréal. 293-99.

Fellbaum, Christiane. 1990. English verbs as a semantic net. *International Journal of Lexicography.* **3**:278-301.

Fillmore, Charles, & Collin F. Baker. 2001. Frame semantics for text understanding. *Proceedings of WordNet and Other Lexical Resources Workshop.* NAACL. Pittsburgh. 59-64.

Jing, Hogyan, & Kathleen McKeown. 1998. Combining multiple, large-scale resources in a reusable lexicon for natural language generation. *Proceedings of COLING-ACL.* Montréal. 607-13.

Kohl, Karen T., Douglas A. Jones, Robert C. Berwick, & Naoyuki Nomura. 1998. Representing verb alternations in WordNet. In Christiane Fellbaum (ed.) *WordNet: An Electronic Lexical Database,* 153-178. Language, Speech, and Communication. Cambridge: MIT Press.

Lakoff, George. 1987. *Women, Fire, and Dangerous Things.* Chicago: University of Chicago Press.

Lambrecht, Knud. 1994. *Information Structure and Sentence Form.* New York: Cambridge University Press.

Lapata, Maria, & Chris Brew. 1999. Using subcategorization to resolve verb class ambiguity. *Proceedings of the 1999 Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora.* College Park, MD. 266-74.

Levin, Beth. 1993. *English Verb Class and Alternations: A Preliminary Investigation.* Chicago: University of Chicago Press.

Manning, Christopher. 1993. Automatic acquisition of a large subcategorization dictionary from corpora. *Proceedings of the 31st Annual Meeting of the Association for Computational Linguistics.* 235-42.

Miller, George A., & Christiane Fellbaum. 1991. Semantic networks of English. *Cognition.* **41**:197-229.

Roland, Douglas, Daniel Jurafsky, Lise Menn, Susanne Gahl, Elizabeth Elder, & Chris Riddoch. 2000. Verb subcategorization frequency differences between business-news and balanced corpora: the role of verb sense. *Proceedings of the Association for Computational Linguistics Workshop on Comparing Corpora.*

Schütze, Carson T. 1996. *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology.* Chicago: University of Chicago Press.

Collin F. Baker & Josef Ruppenhofer
International Computer Science Institute
1947 Center St., Suite 600
Berkeley, California 94704

collinb@icsi.berkeley.edu
josef@icsi.berkeley.edu

Phonetic Analogy and Schwa Deletion in French*

JONATHAN BARNES
University of California, Berkeley

DARYA KAVITSKAYA
Yale University

1. Phonetic Analogy

Phonetic analogy as a diachronic phenomenon has been the subject of no small interest at least since the early part of the last century. Bloomfield (1933: 366), for example, notes the following:

> Indeed, when we observe sub-phonemic variants, we sometimes find them distributed among speakers or systematized among forms, quite in the manner of linguistic borrowing and of analogical change.

What is interesting in this for Bloomfield is the fact that analogy in its traditional preserve is held to apply only to elements which are phonemically contrastive in a language. The intuition here seems to derive from the notion of contrastiveness in sound systems as the measure of "linguistic relevance" or "conscious opposition", with all that is noncontrastive or subphonemic pertaining only to a lower, more mechanical level of linguistic realization. As such, changes at the level of subphonemic detail are better expected in, for example, the phonetic drift of the neogrammarians than in so rational a process as analogy.

In more recent years, the issue of phonetic or subphonemic analogy has entered the debate surrounding the nature of the phonetics-phonology interface in synchrony as well. Steriade (2000: 332) defines it in this context thus:

> ... Some processes located in the 'phonetic implementation component' are qualitatively the same as the ones classified as phonological. Phonetic analogy is qualitatively the same process as cyclicity, the paradigmatic extension of contrastive properties.

The problem is essentially the same as before: in the classic Lexical Phonology model of cyclicity, only lexical processes are held to be cyclic. Postlexical processes, operating on elements of subphonemic detail, should not display this

* We owe much gratitude to the audience of the Fall 2001 Phonology Forum at UCB, and particularly to Andrew Garrett for discussion and inspiration. Many thanks are due to Louis Goldstein, Khalil Iskarous, Doug Whalen and Vladimir Rybin for encouragement, advice and technical assistance.

characteristic. If it can be shown that they in fact do, this would be a challenge not only to the division of processes into lexical and postlexical, but also to the notion of any meaningful split between phonology and phonetics.

An illustrative example of the putative workings of subphonemic analogy is the familiar English vowel duration before flap problem shown in (1).

(1) Vowel duration before flap

a.	rate	raid	b.	rated	raided
	[rĕjt]	[rējd]		[rĕjɾəd]	[rējɾəd]

Earlier held to be a product of rule-ordering, the difference in vowel duration in the two suffixed forms could be ascribed to the operation of phonetic analogy. In the unsuffixed 'rate' and 'raid', the duration of the vowels is determined in part by the identity of the following consonant; before the voiceless /t/, the vowel is phonetically shorter than the vowel before the voiced /d/.[2] Upon the addition of the suffix, the voicing distinction in the obstruents is neutralized, both now being realized as [ɾ]. The difference in duration of the preceding vowels, however, remains held over, as it were, from the unsuffixed forms. This durational difference can no longer be attributed to the phonetic influence of the following stops; it must therefore be the result of analogical extension through the paradigm of certain phonetic characteristics of the base form.

In the remaining sections of this paper, we discuss one potential instance of subphonemic analogy proposed recently by Steriade (2000), schwa deletion in French. Section 2 describes the facts of the case, and Steriade's experimental results and analysis. In section 3 we present the results of our own experiment, designed to test the veracity of the claim for subphonemic analogy in French schwa deletion. General discussion of the issue follows in section 4, and we conclude in 5.

2. French Schwa Deletion as Phonetic Analogy

Schwa deletion in French has received voluminous attention in the phonological literature (e.g., Anderson (1982), Dell (1973), Rialland (1986), Tranel (1981, 1995, 1998), among others). Many researchers have been concerned with the correct division of instances of deleting schwa into separate types or rules. We follow Steriade here in discussing only one of these types, the so-called Optional Schwa Deletion exemplified in (2) and characterized in (3).

(2)	a.	pas de role	[pa də ʁol]
	b.	pas d'role	[pa d ʁol]

[2] A common (if not universal) pattern cross-linguistically. The effect is exaggerated in English, making plausible the claim that the difference here is just as phonological as the voicing distinction itself.

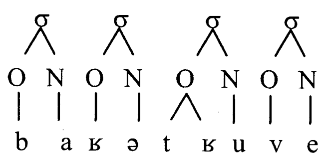
(3) Optional schwa deletion: (V)#C__.C

In question is the optional deletion of schwa following a single postboundary consonant after a vowel or pause.[3] This yields the competing forms [pa də ʁol] and [pa d ʁol], the second being favored in casual speech.

Rialland (1986) observed a curious fact concerning certain instances of deletion of French schwa. Specifically, she noted that the preceding consonant, in non-postpausal contexts ostensibly resyllabified as a coda, nonetheless appears in spectrograms to retain much of the phonetic character of its corresponding onset variant, and not to lengthen the preceding vowel, as it would be expected to do were it in fact in the coda.

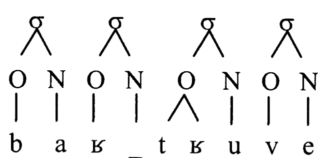
These facts lead Rialland to posit that while the schwa itself is deleted, the syllabic nucleus it projected is nonetheless retained in the representation. Because of this, the preceding consonant is never in fact resyllabified, continuing instead to occupy the onset of its original syllable, the nucleus of which is now "empty". This is shown in (4a) and (4b) below.

(4) a.



bas retrouvé

b.



bas r'trouvé

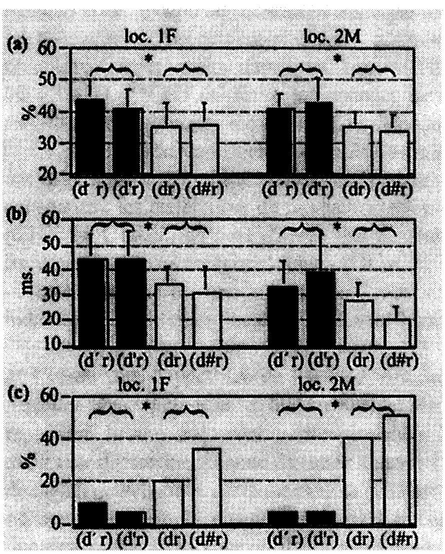
Building on Rialland's work, Fougeron and Steriade set out to confirm experimentally the phonetic peculiarities found in the neighborhood of a deleted schwa. To this end they take minimally-differing phrases from Rialland such as those shown in (5) as tokens for an electropalatographic study. The phrase in token (5a) 'pas de role' contains a schwa optionally subject to deletion by the rules discussed above. (5b) is the same phrase with the schwa now deleted. (5c) contains the same segmental sequence as (5b), but with no underlying schwa. Here the two consonants in question are part of a single complex onset. (5d) again contains no underlying schwa, but the d-ʁ sequence now occurs across a word boundary.

[3] Essentially the conflation of Dell's rules INI Optional and VCE_1 Optional (1973). Whether or not certain internal schwas meet the same description is not at issue here.

(5) a. pas de role [pa də ʁol]
 b. pas d'role [pa d ʁol]
 c. pas drole [pa dʁol]
 d. jade rose [ʒad ʁoz]

The results of this study, shown in (6) (tables reproduced from Fougeron and Steriade 1997), support Rialland's observations. For all parameters measured, forms undergoing schwa deletion patterned consistently with the forms still containing schwa, both of these differing significantly from the forms with underlying consonant sequences. (6a) shows that for both speakers, the amount of linguopalatal contact recorded in the articulation of the /d/ was roughly the same for both "pas de role" and "pas d'role", while "pas drole" and "jade rose" formed another class. (6b) shows the mean duration in ms. of the same occlusion, again with the same grouping of results. (6c) records the frequency of lenition of /d/ in each context. Again, the phrases with underlying schwas pattern together, giving the appearance at least of a /d/ in "pas d'role" which continues to be realized with the phonetic characteristics of its onset allophone, despite the fact that, on the surface at least, the nucleus vowel making this possible is absent. Fougeron and Steriade interpret the lesser linguopalatal contact and greater frequency of lenition of the /d/ in underlying /d-ʁ/ sequences as the result of the shorter duration of occlusion of the consonant in both these contexts.

(6) Fougeron and Steriade (1997)



So far then Fougeron and Steriade's results support at least the intuition of Rialland's analysis. But there is a problem. Steriade argues that there is evidence to suggest that the sequence in question does not in fact constitute a syllable on its own in French. In support of this view she brings the following line from a poem by Georges Brassens:

(7) Il en est de pires il en est d'meilleurs.
[il ɑ̃n e də piʁ il ɑ̃n e d mɛjœʁ]
'There are worse and there are better ones.'

On the basis of (7), Steriade (2000: 328) concludes the following:

> If d'meilleurs of 'better ones' contains schwa, as [də mɛjœʁ], the verse is unacceptably long. Therefore we cannot adopt Rialland's suggestion that the invariant property in C(ə)C sequences is the number of syllables: when schwa deletes, the syllable count is correspondingly decreased.

As a solution to this problem, Steriade proposes that the source of the phonetic effects surrounding the deletion of schwa not surface syllable structure at all, but rather, an instance of subphonemic analogy. Under this interpretation, the schwa deletes in its entirety phonologically, while the durational characteristics of surrounding consonants are inherited from the related utterance without schwa deletion. Steriade formalizes this process as a synchronic Paradigm Uniformity effect. The specific constraint she devises is shown in (8).

(8) Paradigm Uniformity (Left: Duration)
If two consonants, C and C' stand in correspondence and C is morpheme-initial in the careful pronunciation of the relevant morpheme, C' is durationally equivalent to C.

In other words, morpheme-initial consonants in paradigmatically-related forms are required to have the same phonetic durations, even if their immediate phonetic environment might warrant otherwise. This constraint differs in one fairly obvious and significant way from more familiar Faithfulness constraints of OT Correspondence Theory (McCarthy and Prince 1995): Here, rather than the usual correspondence relationship between segments (as in MAX and DEP constraints) or abstract phonological features (as in IDENT constraints), Steriade's proposal requires correspondence between real-time physical characteristics of the realizations of related output forms. This proposal constitutes part of a larger research program questioning the reality of the theoretical distinction between phonetic and phonological representation.

3. Experiment: lip rounding and schwa in French

In order to test the validity of Steriade's Paradigm Uniformity hypothesis, we designed an experiment, taking advantage of one particular phonetic characteristic of schwa in French. Specifically, French schwa, when realized, is produced with significant rounding of the lips. The question our experiment sets out to answer is this: while deleted schwa certainly seems inaudible, is it in fact dispatched in its entirety, or is any of that rounding gesture retained?

3.1. Methodology

For this experiment, we videotaped a native speaker of French producing tokens of the three utterances shown in (9).

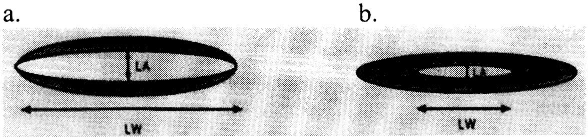
(9) full schwa: Il était bien plus p'tit **que La**nnes.
deleted schwa: Il était bien plus p'tit **qu'la** femelle.
no schwa: Ce n'était qu'un tout p'tit **cla**vecin.

These particular sentences are taken directly from Fougeron and Steriade (1997). Tokens were repeated 5 times in random order, for a total of 15 repetitions. The resulting video was subsequently digitized using MyTV digital video software from Eskape Labs. QuickTimePro was used for analyzing the video and selecting the individual frames exhibiting a maximum of lip rounding for each context. These frames were then excised from the video, and degree of lip rounding in each still was measured using NIH Image (by Wayne Rasband, National Institutes of Health).

The phonetic literature contains numerous potential strategies for the quantification of degree of lip rounding, such as measurement of lip protrusion, aperture, and width of opening. For this experiment, we use the measuring technique proposed by Goldstein (1991), which argues that the most reliable correlate of degree of roundedness for vowels is in fact the side contact of the lips. This is illustrated in Goldstein's figures, reproduced in (10).

(10) Side contact of the lips as the primary articulatory correlate of rounding

a. b.



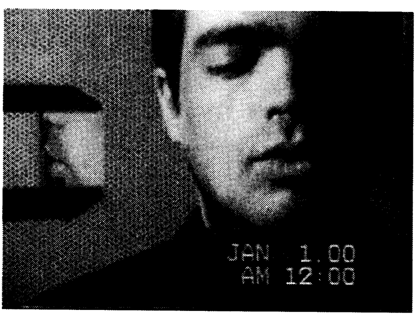
In these figures, (10a) representing an unrounded articulation and (10b) a rounded one, vertical and horizontal arrows indicate dimensions of lip aperture and lip width. Goldstein's measure of side contact refers to the portions of the lip width dimension in which there is contact between the upper and lower lips.

In order to obtain measurements of side contact from our video stills, it was necessary first to set a scale by which NIH Image would establish a correspondence between some unit of distance and the number of pixels between any two points in a given frame. To this end we shot several seconds of the initial video with a measuring tape positioned immediately adjacent to the subject's lips at the appropriate depth in the scene. NIH Image was then able to calculate distances marked off in our stills in millimeters.

3.2. Results

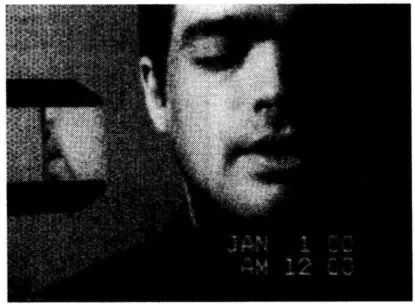
The stills shown in (11-13) contain lip rounding maxima for the relevant portions of the three utterance-types in question (full schwa, deleted schwa, and no schwa). The results of our study are evident here upon even a visual inspection: unsurprisingly, the phrases with a fully realized schwa show the most lip-rounding, and the phrases with no schwa show the least. The phrases with a deleted schwa, however, form a middle category, in which some rounding is present, but not as much as if there would be for a fully realized schwa. The graphs in (14) provide the numbers.

(11) Full schwa: [kə la]
Il était bien plus p'tit **que La**nnes.
Il était bien plus p'tit **que la** femelle.



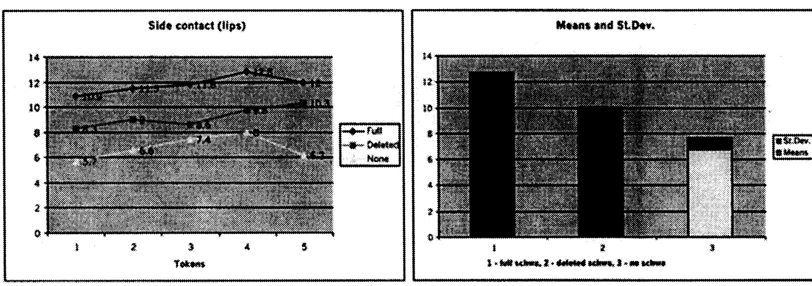
(12) Deleted Schwa: [k_la]
Il était bien plus p'tit **qu'la** femelle.
Il était bien plus p'tit **qu'la** table.

(13) NoSchwa: [kla]
Ce n'était qu'un tout p'tit **cla**vecin.
J'ai vu beaucoup de fli**cs là**-bas.



While the size of the sample here makes statistical analysis essentially beside the point, the patterning of the results into three completely non-overlapping classes, each with a small standard deviation is suggestive indeed. It is now necessary on the basis of these results to conduct an expanded version of this study to confirm the pattern emerging here.

(14) Side contact of lips in mm for three contexts



4. Discussion

Our results indicate that schwa is not in fact deleted, insofar as some of its component gestures remain articulated. Instead, what we have in French schwa deletion is another instance of articulatory overlap producing the appearance of categorical deletion, as in the famous "perfect memory" facts of Browman and Goldstein (1990). In their words, "much coarticulation and allophonic variation occurs as an automatic consequence of overlapping, invariant underlying gestures" (Browman and Goldstein 1990). As rate of speech increases, so does the overlap of neighboring gestures. In fast speech, then, the inherently short schwa in the French examples above is overlapped completely by surrounding consonants. The gestures comprising schwa may remain (at least in some reduced-magnitude variant), but the schwa itself becomes inaudible. Optional schwa deletion is thus not categorical deletion at all, but rather the result of a gradient process reorganizing the phasing of the consonant gestures of the string relative to those of the nuclear schwa. Under this account, paradigm uniformity is extraneous here.

Steriade's Paradigm Uniformity (PU) analysis of French relies on categorical deletion of schwa (with concomitant loss of syllabicity), together with analogical extension of one particular phonetic characteristic from a related form. To extend this account in light of our results, additional copresent phonetic analogies could be posited, one for each feature of schwa ultimately located in the deletion environment. For example, Smorodinsky (1986) shows that vertical displacement of the tongue dorsum is significantly different in deleted schwa contexts in French than in comparable consonant clusters. Minimally this adds to the PU account constraints establishing correspondences for not only rounding, but tongue dorsum position as well.

The PU account now says, in essence: apply categorical phonological deletion of schwa, and then restore multiple phonetic features associated with its presence through a list of paradigm uniformity constraints referring to a related utterance of the same string without schwa deletion. The PU account thus casts the facts of schwa deletion as the accidental product of a particular ranking of a list of PU constraints. The prediction is that rerankings could produce the analogical retention of any one of the phonetic characteristics associated with schwa, or indeed any arbitrary constellation thereof. Our account, by contrast, sees here only the effect of a single phonological process, a specific rephasing of gestures accompanying casual speech.

Finally, we are left with the problem of the syllabification of sequences with deleted schwa, as exemplified by Steriade's verse from Brassens above. Steriade asserts that for a syllable to scan in French verse, it must have an audible nucleus. Our schwa is present articulatorily, but inaudible. It is therefore passed over for purposes of metrification. The process of extracting firm generalizations about phonological syllabification norms from poetry, however, is far from simple. As Trubetzkoy himself notes, "any metrics is of course violence" (letter to Jakobson, 20 December, 1922 [2001]). Consider, for example, the following quatrains from John Donne. For the first line of the poem in (15) to scan as tetrameter, it is necessary that the words "heaven's influence" contain only three syllables. Likewise, for the first line of the poem in (16) to scan as pentameter, "ignorant" must be disyllabic and "the experienced" must have only three syllables. Leaving aside the question of whether such a parsing of "ignorant" implies the acceptability of /nr/ onset clusters, we are still left to wonder just how much the possibility of metrifications such as Donne's actually tells us about the syllabification norms of 17th century English.

(15) Donne's tetrameter

On man **heaven's influence** works not so,
But that it first imprints the air,
So soul into the soul may flow,
Though it to body first repair.
(*The Ecstasy,* lines 57-60)

[on . man][heaven'. s in][fluence . works][not . so]
[but . that][it . first][im . prints][the . air]
[so . soul][in . to][the . soul][may . flow]
[though . it][to . bo][dy . first][re . pair]

(16) Donne's pentameter

Those wars the **ignorant**, these **th' experienced** love,
There we are always under, here above.
There engines far off breed a just true fear,
Near thrusts, pikes, stabs, yea bullets hurt not here.
(*Love's War,* lines 35-38)

[those . wars][the . ig][norant . these] [th' ex . per][ienced . love],
[there . we] [are . al][ways . un][der . here][a . bove]
[there . en][gines . far] [off . breed][a . just][true . fear]
[near . thrusts][pikes . stabs][yea . bul][lets . hurt][not . here]

5. Conclusions

French schwa deletion is not an instance of phonetic analogy. Rather, gradient reorganization of gestures in a manner familiar from the literature on Articulatory Phonology is enough to account for the facts of this case. The gestural reduction account assumes only the existence of gradient phonological processes capable of targeting a single dimension of the realization of a string. In the case of French schwa, at issue is the phasing of consonant gestures relative to one another. Other phonetic effects associated with this so-called Schwa Deletion (such as the decreased magnitude of the rounding gesture) result directly from accommodations related to this rephasing. Whereas the Paradigm Uniformity approach posits first categorical deletion of schwa and then subsequent reintroduction of phonetic features such as the duration of the preceding consonant, in our account these phonetic features are never altered to begin with. Instead of the sum effect of an arbitrarily large set of analogical repairs to the site of an overzealous deletion, we see the result of a unitary phonological process.

References

Anderson, Stephen R. 1982. The analysis of French schwa, or, How to get something for nothing. *Language* **58**: 534–573.

Bloomfield, Leonard. 1933. *Language.* New York: Henry Holt & Co.

Browman, C.P. & L. Goldstein. 1990. Tiers in Articulatory Phonology with some implications for casual speech. In J. Kingston & M.E. Beckman (eds.) *Papers in Laboratory Phonology I: Between the grammar and physics of speech*, 341-376. Cambridge: Cambridge University Press.

Dell, François. 1973. *Les règles et les sons*. Paris: Hermann.

Donne, John. 1971. In A.J. Smith (ed.) *The Complete English Poems.* London: Penguin Books, Ltd.

Fougeron, Cecile & Donca Steriade. 1997. Does deletion of French schwa lead to neutralization of lexical distinctions? In *Proceedings of Eurospeech 97* **2**: 943-946.

Goldstein, Louis. 1991. Lip rounding as side contact. *Proceedings of the XIIth International Congress of Phonetic Sciences* **1**: 97-101. Université de Provence.

Kiparsky, Paul. 1985. Some consequences of Lexical Phonology. *Phonology Yearbook* **2**: 85-138.

McCarthy, John J. & Alan S. Prince. 1995. Faithfulness and reduplicative identity. Ms., University of Massachusetts, Amherst and Rutgers University.

Rialland, A. 1986. Schwa et syllabes en Français. In L. Wetzels & E. Sezer (eds.) *Studies in Compensatory Lengthening.* Dordrecht: Foris.

Smorodinsky, Iris. 1996. Schwas in French: an articulatory analysis. Ms., Yale University.

Steriade, Donca. 2000. Paradigm uniformity and the phonetics-phonology boundary. In Michael B. Broe & Janet B. Pierrehumbert (eds.) *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, 313-335. Cambridge: Cambridge University Press.

Tranel, Bernard. 1981. *Concreteness in Generative Phonology: Evidence from French.* Berkeley, California: University of California Press.

Tranel, Bernard. 1995. Current issues in French phonology: liaison and position theories. In John A. Goldsmith (ed.) *The Handbook of Phonological Theory*, 798-816. Oxford: Blackwell Publishers.

Tranel, Bernard. 1999. Optional schwa deletion: on syllable economy in French. In J.-Mark Authier, Barbara E. Bullock, & Lisa A. Reed (eds.) *Formal Perspectives on Romance Linguistics*, 271-288. Amsterdam: John Benjamins.

Trubetzkoy, N.S. 2001. In Anatoly Liberman (ed.) *N.S. Trubetzkoy: Studies in General Linguistics and Language Structure.* Durham: Duke University Press.

Zsiga, Elizabeth. 1993. *Features, Gestures, and the Temporal Aspects of Phonological Organization.* PhD Dissertation, Yale University.

jbarnes@socrates.berkeley.edu

Dept. of Linguistics
University of California, Berkeley
1203 Dwinelle Hall
Berkeley, CA 94720

darya.kavitskaya@yale.edu

Dept. of Linguistics
Yale University
P.O. Box 208366
New Haven, CT 06511

Unpacking the Okanagan Person-Marking Conundrum

PAUL BARTHMAIER
University of California, Santa Barbara

0. Introduction

Okanagan, a Southern Interior Salish language spoken in northern Washington state and southern British Columbia, exhibits a peculiar set of pronominal morphemes that surely are a testament to a diverse and varied history. From the outside, the pronominal markers associated with Okanagan clauses appear to be a disparate group of morphemes. A lack of formal similarity frustrates attempts to characterize them as either nominative-accusative or ergative-absolutive. Morphologically the pronominal forms appear to be the typologically rare tripartite system. Yet, speakers have little trouble using the different markers in their appropriate contexts. In what follows, I will propose an analysis of how the person marking in the language has come to have such an interesting shape. I will offer internal and external motivations that the system responded to as it evolved into its current form.

Okanagan has been investigated by several different researchers, the most notable of whom, Anthony Mattina, has continuously studied the language since the late 1960s. In his recent paper (A. Mattina 2001), he describes four pronominal sets, summarized in the tables below. If four sets of pronominal marking were not complicated enough, a language learner must also contend with homophonous forms overlapping paradigms and a host of allomorphy some of which has no reasonable explanation in the phonology. (Parenthesized elements in the above tables are subject to regular phonological processes in the language such as deletion in unstressed environments and coronal simplification.)

(1)

	Intransitive	
	sg	Pl
1	kn=	k^wu=
2	k^w=	p=
3	Ø	Ø...-lx

	Genitive Subject[1]	
	sg	pl
1	i(n)-	-tt
2	a(n)-	-mp
3	-s/-c	-s-lx/-c-lx

[1] A. Mattina (2001) labels this paradigm 'possessive,' but elsewhere he has used 'genitive'. For further discussion of this category label see A. Mattina (1993) and N. Mattina (1996).

	Transitive Subject	
	sg	Pl
1	-(í)n	-(í)m/-t
2	-(í)xʷ	-(í)p
3	-(í)s	-(í)s-lx

	Transitive Object	
	sg	pl
1	kʷu=	kʷu=...-m
2	-s/-m	-ɬ(úl)m
3	-Ø	-Ø...-lx

Finding the same consonants, particularly *-m, -s,* and *-t*, recurring across so many forms is somewhat surprising in a language like Okanagan, which has almost forty consonants in its inventory. But, if we focus on individual sections of the grammar one at a time, we can explain why we see the same consonants in so many forms. Borrowing may account for some of the changes, but the modern system is also the result of detransitivization, grammaticization, and reanalysis. This process has led to a radical restructuring of a crucial portion of core grammar which we still see today.

1. Possessive marking

The first step towards simplifying the person-marking puzzle is to point out that one set of markers, the 'Genitive Subject', is identical in form to the possessive markers in the language. Although genitive subjects and possessive morphemes are employed in distinct sections of the grammar, their formal similarity warrants a unifying diachronic explanation. The examples in (1) have morphemes circled which correspond to underlined pronouns in the glosses.

(1a) n.t'k'ʷìkiʔ - ɬt - xʷ i- stʔəwtílt (gw26)[2]
saddle.up - DITR -2ERG 1POSS-youngest.boy[3]
You will saddle it up for my youngest boy

(1b) way' x̌às - t a- sq'líps (gw352)
DM be.good-STA 2POSS-handkerchief
Your handkerchief is pretty

(1c) way' kən - xìt - əm iʔ t l.ʔìw -s (gw30)
DM help - BEN - INTR DET OBL father-3POSS
His father helped him

The above Okanagan examples show typical predicate-initial clauses with single arguments. The predicate describes an event or state and is marked with verbal

[2] The majority of examples come from Mattina (1985) *The Golden Woman*, an Okanagan narrative tale told by Peter J. Seymour to Anthony Mattina. Examples numbers correspond to the line number of the published text.

[3] Abbreviations used are: ASP=aspect, BEN=benefactive transitivizer, DET=determiner, DIR=directional, DITR=ditransitive, DM=discourse marker, DST=distributive, ERG=ergative, EVID=evidential, INTR=intransitive, IMP=imperative, IRR=irrealis, LOC=locative, HTR=high transitivity, LTR=low transitivity, MUT=mutative, NC=non-control, NEG=negative, NOM=nominalizer, OBL=oblique, POSS=possessive, RCP=reciprocal, RDP=reduplication, RES=resultative, STA=stative.

morphology. In each example the argument is a nominal which is affixed by a possessive marker. In the following, we will see that the construction types found in (1) serve as a model for many other constructions.

Cross-linguistically, nominalizations have been recognized as a common strategy for identifying complements. In English, nominalizations are used as one type of complementation strategy. Consider the following examples.

(2a) The Taliban destroyed the Buddha statue.

(2b) [*The Taliban's destruction of the Buddha statue*] angered many people.

The clause in (2a) is nominalized in order to be used as the subject argument of (2b). Note that word order is preserved, the subject and verb of (2a) are in a genitive relationship and the object is an oblique.

Nominalization also appears to be an important grammatical resource of Okanagan as a way of identifying arguments.

(3a) c - my - st - ìs i - s - kʷúl - st - əm (gw137)
ASP-know - LTR -3ERG 1POSS-***NOM***-send-LTR-INTR
He'll know that I sent you (lit. He'll know it was my sending)

(3b) way' x̌às - t iˀ s - c - k'əɬpáx̌- tət (gw352)
DM be.good-STA ART ***NOM***-ASP-think-1pPOSS
Yes, we've done good thinking (lit. our thinking is good)

The intransitive clauses in (3) show a pattern similar to those in (2); each clause consists of a predicate with verbal morphology and a single argument, yet the arguments in (3) have a semantic core which is verbal. In order for an event to be an argument, the grammar requires that the clause be nominalized. In the above examples, the nominalizer is circled and the boldface in the gloss line shows its use as a nominalizer. This strategy is common in Salish and appears to go back to Proto-Salish (Kroeber 1999). In these nominalizations, the possessive morphology corresponds to the subjects of the nominalized clauses.

The examples in (3) only partially describe the picture because the arguments are grammatically intransitive. Transitive subjects of nominalizations look virtually the same as intransitive subjects, but objects are curiously marked by a proclitic from the intransitive subject set. N. Mattina (1996) labels the intransitive set the 'absolutive paradigm', presumably because these forms mark subjects of intransitive clauses and objects of nominalized clauses.

(4a) uɬ my - p - nú - s t incàˀ kʷ= i -s-kʷúl-st-əm (gw127)
DM know-MUT-NC-3ERG OBL I 2 1POSS-NOM-send-LTR-INTR
He'll know that I sent <u>*you*</u>

(4b) cù - nt - s way' kʷ= i - s - k'íw - lx - st - m (gw865)
say-TR-3ERG DM 2 1POSS-NOM-climb-motion-LTR-INTR
She'll say to you, 'I'll take <u>*you*</u> upstairs…'

In both examples above, the subject of the nominalized clause is marked by the possessive morpheme *i(n-)* and the object by the intransitive subject morpheme *kʷ=*. Example (4a) also nicely illustrates the argument structure of the transitive predicate *mypnús* ‘He’ll know X’ in that X must be interpreted as the nominalized clause. The emphatic pronoun, *incàʔ* ‘I,’ is necessarily marked as oblique. Thus the two core arguments are the nominalized clause and the third person subject which is marked on the predicate by the ergative *-s*.

Nominalization is also employed in some Okanagan negation strategies. Payne (1997) explains that languages may have several different strategies for negation and one common type is nominalization.

(5a) lut t’= i - k - s - ən - xʷst - ítkʷ (gw289)
NEG EVID 1POSS-IRR-NOM-LOC-walk-water
I'm not going to wade in the water.

(5b) lut kʷ= t’= i - k - s - xʷíc’ - əɬt - əm (gw289)
NEG 2 EVID 1POSS-IRR-NOM-give-DITR-INTR
I'm not going to hand it to you,

The negative, *lut*, has a single nominalized argument where the subject of the nominalized clause is marked by the possessive set and the object by the intransitive set.

During the diachronic progression of the grammar, the nominalization schema seems to have been further extended in Okanagan as a means of signaling interclausal dependencies. Kroeber (1996) proposes for Thompson Salish that nominalization of this type may be thought of as a clause-chaining device where the nominalizer is seen as marking a predicate as non-finite. The example in (5), which is taken from A. Mattina (1985), suggests that this may be a useful notion for Okanagan as well. It consists of a sequence of three clauses. In the third clause, Mattina’s gloss certainly indicates that the clause is pragmatically and perhaps syntactically dependent, yet the clause has no overt marker of dependency other than the nominalization.

(6) yəyʕà - t stìm’ x̌mìnk - ənt - p (gw10)
all - STA what want - HTR - 2pERG
Anything you want

uɬ xʷíc’ - ɬm - ən,
DM give - 2pObj-1ERG
I give to you

uɬ kʷu= a - s - c - ənpətʔíls - əm
DM 1p 2POSS-NOM-ASP-respect.feelings - INTR
<u>because</u> you respect my feelings.

This example represents a different use of possessive morphology than was seen in (4). Here the nominalized clause is not an argument of any verb; rather it provides adverbial information to the discourse. The use of free-standing non-finite constructions to add circumstantial or adverbial information to a discourse is a common strategy seen in clause-chaining languages.

The final step in the process of evolution of nominalization seems to be the use of nominalized clauses as seemingly independent clauses.

(7) cù - s lut aɬ'iˀ kʷ= t'= i - k - s - ɬwn - íkst - əm (gw362)
say-3ERG NEG because 2 EVID 1POSS-NOM-leave-hand-INTR
He said: "No, I'm (not) going to let you go,

way' kʷ= i - s - c - ƛ'ˀá - m
DM 2 1POSS-NOM-ASP-look.for-INTR
I came to get you."

The gloss in (7) does not indicate a dependency relationship like the one in (7), although a dependent interpretation might not be infelicitous. The second clause provides the reason for doing the first. The many examples like the one in (7) help explain the motivation for the synchronic category label 'genitive subject', but a diachronic approach sheds greater light on the complexity of the history and its use in discourse.

To summarize, what began as a simple kernel in the morphology of the language has apparently been exploited by speakers of Okanagan for several different grammatical purposes. That is, nominalization and possessive marking has spread from the level of morphology to the syntax and pragmatics of the language as a grammatical device which provides a means of accomplishing several different functions. Synchronically we find the possessive morphemes operating with nominalized clauses to indicate not only possession, but also arguments of predicates, interclausal dependencies, and even seemingly independent clauses.

2. Second Person Singular Object

The second piece of the puzzle involves the second person singular object forms. They are represented by two, phonologically unrelated allomorphs: *-s* and *-m*. The choice between the two allomorphs is sensitive to the transitivizer of the predicate. It can easily be shown that *-s* is found with the higher-transitivity predicates and *-m* with the lower-transitivity predicates. Newman (1979b), in his reconstruction of Salish object forms, proposes two paradigms which he calls the 'neutral object' and the 'causative object.' For second persons, the neutral object forms in the Salishan family usually have an *-s* and the causative objects usually have an *-m*, except where the forms have been 'replaced.'

Newman's analysis is complementary to the one presented here in that neutral object paradigms pair with higher transitive predicates and causative objects correspond to lower transitivity predicates. However, his reconstruction suggests that this patterning could go back to Proto-Salish, a claim which may be debatable and one that would obfuscate any internal motivation which may exist. In what follows, I present an analysis of the internal restructuring of the Okanagan pronominal system, and I elaborate on the functional motivations behind this reorganization. This analysis will not dispute the shape of the pronominal forms, but it may begin to answer questions like how or why they have come to have their shape.

To begin, it may be helpful to note that the order of morphemes in the Okanagan transitive predicate is:

(8)

(aspect)	stem	transitivizer	object	subject

Mattina and Montler (1990) present the following data illustrating the sensitivity of the second person object morpheme to the transitivizer.

(9)	**-nt transivizer**		**-st transitivizer**	
subject	strong root	weak root	strong root	weak root
I	wík-nt-s-n	c'lx-nt-s-ín	wík-st-m-n	c'lx-st-úm-n
	I saw you	I grabbed you	I usually see you	I usually grab you
he	wík-nt-s	c'lx-nt-s-ís	wík-st-m-s	c'lx-st-úm-s
	he saw you	he grabbed you	he usually sees you	he usually grabs you

In the above examples, the morpheme in question is circled and its gloss is in boldface. Before proceeding with the discussion of the pronominal forms, a few points about the data are worthy of attention. Firstly, the forms given under *-st transitivizer* are presented for illustrative purposes, but are unattested without the aspect marker ac-/c-.[4] (N. Mattina 1996:72) Secondly, the two transitivizers discussed here have reflexes in all of the Salish languages and have received various labels. Following Hopper and Thompson (1980), I refer to clauses with -*nt* as higher transitivity because of their association with telic volitional actions and *-st* clauses as lower transitivity because of their non-punctual interpretation.[5]

The examples in (9) clearly show that there are two distinct morphemes referring to second person objects. Perhaps more interestingly, it seems that the current alternation is the result of two different paths, neither of which originated with second person morphemes. Although from two different diachronic sources, the stimulus for development was the same: a politeness ploy to defocus second person.

The likely source of the *-s* is the third person marker. The use of third person pronominals for second person is well known in many languages including

[4] For a full description of Okanagan aspect, see N. Mattina (1996).

[5] Transitivity in Okanagan has received a good deal of attention. See A. Mattina (1978a/b), A. Mattina (1982), and Hebert (1982b). For a broader description of the family, see Thompson (1985).

German, French, Spanish and Hungarian. This type of replacement tends to happen for formal second person forms. Over the diachronic development of Okanagan grammar, in formal situations, speakers may have used third person markers as a means of deflecting from the more salient speech act participants' role to the third person. Eventually the formal forms replaced the familiar terms and they became the only means of reference. This process is analogous to the situation in German where the pronoun *sie* 'they' has come to be used for formal second persons. On this continent, Frachtenberg (1922) describes a similar pattern in Siuslaw where third person plural is also used for second person singular.

The source of the second alternant *-m* is actually the pan-Salishan marker of intransitivity. Detransitivization served as a means of defocusing the agent. In Okanagan this change occurred in lower-transitivity clauses with second person agents. That is, the clauses which were semantically lower transitivity, *á la* Hopper and Thompson (1980), were marked as grammatically intransitive with the morpheme *-m*. Historically, this may be due in part because the *-st* transitivizer implicated an agent that was low in potency. However, synchronically the only evidence of detransitivization is the morpheme *-m*. Presumably the resulting intransitive clause could have had an optional oblique second person. Eventually these intransitive constructions were reanalyzed as transitive (when the detransitivizing *-m* was interpreted as referring to a second person patient) and the clauses marked with the transitive subject suffixes.

3. First Person Plural Transitive Subject

The third piece of the puzzle also involves detransitivization, but in this case it began as a strategy for what A. Mattina (1973) calls 'third-person indefinite forms'. Interestingly, Newman (1979a) does not discuss the allomorphy of first person plural transitive subject markers *-t/-m*. Based on comparative evidence we know the *-t* represents the older form, but where the *-m* came from may not be immediately obvious. Again, there is good motivation for the detransitivizing *-m* to play a role.

A. Mattina (1973) observed the homophony of third-person indefinite forms and transitive clauses with first person agents and third person patients. The diachronic path linking the two may look something like the following. As a first pass the detransitivized construction resembled a passive, as in 'The cake was eaten.' This form was used to highlight the patient and defocus the agent. In time this was interpreted as a third person indefinite form, as in 'Somebody ate the cake.' Eventually this third person indefinite construction was reanalyzed as a first person plural form in a scheme that is analogous to the facts surrounding the French *on*, as in *on arrive* = 'We're here!' But the plan was not adopted wholesale and today we are left with the present allomorphy of *-(í)m/-t*.

(10)	**-nt transivizer**		**-st transitivizer**	
	strong root	weak root	strong root	weak root
You	wík-nt-s -t	c'lx-nt-s -it	wík-st-m -t	c'lx-st-úm -t
	We saw you	We grabbed you	We usually see you	We usually grab you
3rd person	wík-nt -m	c'lx-nt -ím	wík-st -m	c'lx-st -ím
	We saw him	We grabbed him	We usually see him	We usually grab him
indefinite	wík-nt -m	c'lx-nt -ím	wík-st -m	c'lx-st -ím
	Somebody saw you	Somebody grabbed you	Somebody usually sees you	Somebody usually grabs you

Immediately we notice the homophony of the Okanagan forms involving third person objects and indefinites, shown above on the second and third lines. The first person plural *-m* crept into the grammar along the aforementioned path, but, as seen on the first line of Okanagan forms, the morpheme remained *-t* with second person objects. The first person plural allomorph *-m* is reserved for use only in constructions which also involve third person objects.

4. First Person k^wu=

The final piece of the puzzle involves the first person form *k*^{w}*u*= which marks singular and plural objects as well as plural and some singular intransitive subjects. Newman (1979a) claims that this form is a borrowing from Kutenai *ku*=, analyzable as *k*-'subject marker' + *hu*= 'first person.' The status of *k*- is refined by Morgan (1991), which is certainly the most comprehensive investigation of Kutenai to date. Morgan shows that *hu*= marks first person subjects regardless of number. The morpheme *k*= is a subordinating proclitic which coalesces with the initial laryngeal consonant to yield *ku*=.

The fact that a pronoun is borrowed is startling enough, but the extent of restructuring of the pronominals it caused is equally striking. Thomason and Kaufman 1988 explain that this type of structural borrowing most likely points to intense contact. It does seem reasonable to view *k*^{w}*u*= as a borrowing into Okanagan because its distribution does not match the rest of the pronominals. It is the only pronominal that is truly absolutive in its distribution, that is, it marks subjects of intransitive and objects of transitive clauses. In addition, as an object marker, its form and position differs from the rest of the paradigm: it is a proclitic, not a suffix. Most likely *k*^{w}*u*= entered the Okanagan grammar as a marker of intransitive subject and then spread to other paradigms. The examples in (11) show that *k*^{w}*u*= fits in nicely with the rest of the intransitive paradigm.

(11a) way' kən= n - wìs - əlx (gw336)
DM 1 LOC-high-DST
I jumped up

(11b) way' k^w= c - kìc - x (gw849)
DM 2 ASP-reach.someone-IMP
You got here

(11c) way' k^wu= ɬ - x^wùy (gw371)
DM 1p back-go
We will go back

(11d) uɬ way' p= x^wùy (gw11)
DM DM 2p go
Ok, (you all) go on

The examples in (11) show *k^{w}u=* occurs in the same slot as the rest of the markers and it bears some phonological similarity with the first and second singular forms which are also velar-initial. Apart from the fact that these clauses are not subordinate, the use of *k^{w}u=* in Okanagan matches its usage in Kutenai as proclitic marking first person.

Apparently *k^{w}u=* did not replace the older singular form, *kən=*, but N. Mattina (1996:172) reports that *k^{w}u=* is used for first person singular in predicate nominal constructions involving kin or affinal terms. Compare (12a) and (12b) with (12c).

(12a) k^wu= lˀíw - s (nm1996:174)
1 father-3POSS
I am his father

(12b) k^wu= kɬ - ylmíx^{w}əm (nm1996:180)
1 IRR-chief
I will be a chief

(12c) kən= kɬ - naˀník'mn (nm1996:175)
1 IRR - knife
I will be a knife

These data are based on her fieldwork and are presented in her dissertation. The data seem to illustrate an apparent split in the first person singular subject marking, shown in (12). The predicate in (12c) is not a kin or affinal term and therefore is marked with the older subject form *kən=*. As a result, *k^{w}u=* can be used to mark both singular and plural first person subjects, but its use with singular referents is limited. The fact that Okanagan preserves the older form *kən=* is probably attributable to frequency.

Once this pronoun was borrowed, it would, of course, be used for the other function of the intransitive set: namely, marking objects of nominalizations, as was shown in examples (5) and (6). In the following examples, we find that when *kʷu=* is used to mark an object, the form does not specify number. This ambiguity with respect to number matches its use in Kutenai.

(13a) uɬ kʷa x̌əl stìm' kiˀ kʷu= əc - ən - q'aˀ - ìls - əm - st - xʷ (gw369)
DM DM for what that 1 ASP-LOC-business-feelings-INTR-LTR-2ERG
Why is it your business

kʷu= ɬ= a - k - s - kʷnì - m
1 DIR 2POSS-IRR-NOM-take-INTR
to take me

(13b) síw - ənt - s- t kʷu= a - k - s - qʷaˀm-ənwíxʷ-st-əm (gw696)
ask-HTR-2OBJ-1pERG 1 2POSS-IRR-NOM-accustomed-RCP-LTR-INTR
We ask you to give us an introduction

(13c) lut kʷu= a - k - s - ən - k'ʷíxʷ - kn' - əm (gw862)
NEG 1 2POSS-IRR-NOM-LOC-unravel-back-INTR
Don't take the saddle off me

(13d) lut kʷu= t'ə= k-s-c-ən-kcn-íkən'-(n)t-əm (gw223)
NEG 1 EVID IRR-NOM-LOC-overtake-back-HTR-INTR
She will never overtake us

In the examples above, the clauses in (13a) and (13b) show that *kʷu=* marks first person objects of nominalizations regardless of their number. Similarly, (13c) and (13d) show that the strategy is also employed in negative constructions.

The important difference from the rest of the intransitive paradigm is that *kʷu=* has permeated the transitive paradigm. That is, in addition to marking objects of nominalizations, it has also become the standard way mark all first person objects.

(14a) kʷu= k - ˀamútm - ənt - xʷ (gw465)
1 RES- sit - HTR-2ERG
(You) stay with us

(14b) uɬiˀ kʷu= ɬ= málx̌aˀ - nt - xʷ (gw305)
DM 1 DIR lie - HTR - 2ERG
And then you lied to me

The examples above show that *kʷu=* is used to mark first person objects in main clauses regardless of whether the referent is singular or plural. The fact that *kʷu=* is the only form from the intransitive set that marks objects in main clauses strengthens the hypothesis that *kʷu=* is borrowed.

One question remains: how did *kʷu=* actually come to mark objects in main clauses as well as nominalizations? The answer may lie in clauses with a third person agent acting on a first person plural patient.

(15a) kʷu= ɬ= málx̌aʔ - s (gw302)
I DIR lie - 3ERG
He lied to me

(15b) uɬ kʷu= síw - ənt - əm (gw911)
DM I ask - HTR - INTR
he'll ask us (lit. we were asked)

The examples in (15) show that in clauses with third person agents and first person patients a distinction has been grammaticized in that clauses with first plural patients are detransitivized. The predicate resembles the indefinite forms seen in example (11). By detransitivizing, speakers are left with an intransitive clause, as in (15b), which will take the intransitive proclitic *kʷu=*. The ambiguity of number may have led to the spreading of *kʷu=* to mark all first person objects, as in (15a). Eventually *kʷu=* spread to mark first person objects with second person subjects, as in (15), but the motivation to detransitivize did not exist.

5. Conclusion

In the end, what initially appear to be anomalous forms in an already complex grammar turn out to be well-motivated internal restructuring of the language. Rather than being an arbitrary pattern of person marking, the formal similarities of the morphemes represent a realignment according to functional pressures in the diachronic development of the grammar of Okanagan. Detransitivization followed by reanalysis, as well as borrowing, helps to explain the allomorphy found in the person markers of Okanagan.

References

Chafe, Wallace. 1987. Cognitive constraints on information flow. In Russell S. Tomlin (ed.) *Coherence and grounding in discourse*, 21-51. Amsterdam: John Benjamins.

Comrie, Bernard and Sandra A. Thompson. 1985. Lexical Nominalization. In Timothy Shopen (ed.) *Language typology and syntactic description*, 349-398. Cambridge: Cambridge University Press.

Frachtenberg, Leo J. 1922. Siuslawan (Lower Umpqua). In Franz Boas (ed.) *Handbook of American Indian Languages*, vol. 2, 431-629. Washington DC: Bureau of Ethnology

Hebert, Yvonne. 1982a. Transitivity in Nicola Lake Okanagan. PhD dissertation, University of British Columbia.

Hebert, Yvonne. 1982b. Aspect and transitivity in (Nicola Lake) Okanagan. In Paul Hopper and Sandra A. Thompson (eds.) *Syntax and Semantics* **15**:195-215. New York: Academic Press.

Hopper, Paul and Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* **56**:251-299.

Kroeber, Paul. 1996. In the papers for the *Thirtieth-first International Conference for Salish and Neighboring Languages,* 197-202. Vancouver: UBC.

Kroeber, Paul D. 1999. *The Salish language family: reconstructing syntax.* Lincoln and London: University of Nebraska Press.

Mattina, Anthony. 1973. *Colville Grammatical Structure.* University of Hawaii Working Papers in Linguistics 5. Honolulu: University of Hawaii.

Mattina, Anthony. 1978a. The Colville Transitive System. In the papers for the *Thirteenth International Conference for Salishan Languages.* University of Victoria.

Mattina, Anthony. 1978b. Parallels between the Colville Transitives and Pseudo-Intransitives. In D. Malsch, J. Hoard and C. Sloat (eds.) *Proceedings of the Eighth Annual Meeting of the Western Conference on Linguistics*, Linguistics Research Inc., Edmonton, Alberta.

Mattina, Anthony. 1982. Colville-Okanagan transitive system. *International Journal of American Linguistics*, **48**:421-435.

Mattina, Anthony. 1985. *The Golden Woman: the Colville narrative of Peter J. Seymour.* Tucson: University of Arizona Press.

Mattina, Anthony. 1987. *Colville-Okanagan Dictionary.* University of Montana Occasional Papers in Linguistics 5. Missoula: University of Montana.

Mattina, Anthony. 2001. In the papers for the *Thirtieth-first International Conference for Salish and Neighboring Languages,* 215-235. University of British Columbia Working Papers in Linguistics, 6. Vancouver: UBC

Mattina, Anthony and Timothy Montler. 1990. Towards a Reconstruction of the Proto-Salish Transitive System. Paper presented at the Third Spring Workshop on Theory and Method in Linguistic Reconstruction. Pittsburgh.

Mattina, Nancy. 1996. *Aspect and Category in Okanagan Word Formation.* PhD dissertation, Simon Fraser University.

Morgan, Lawrence R. 1991. *A description of the Kutenai Language.* PhD dissertation, University of California, Berkeley.

Newman, Stanley. 1979a. A history of Salish possessive and subject forms. *International Journal of American Linguistics*, **45**:207-223.

Newman, Stanley. 1979b. The Salish object forms. *International Journal of American Linguistics*, **46**:155-167.

Thomason and Kaufman 1988. *Language Contact, Creolization and Genetic Linguistics.* Berkeley:University of California Press.

Thompson, Laurence. 1985. Control in Salish grammar. In Frans Plank (ed.) *Relational Typology*, 391-428. Berlin: Mouton de Gruyter.

Paul Barthmaier
Department of Linguistics
University of California, Santa Barbara
Santa Barbara, CA 93106

ptb0@umail.ucsb.edu

Effects of Signal-Independent Factors in Speech Perception

JOSÉ R. BENKÍ*
University of Michigan, Ann Arbor

0. Introduction and Background

In human speech recognition, listeners use sensory information from the speech signal to match a stimulus with an internal representation. The accuracy of that process is affected by many factors, including, but not limited to, the acoustic-phonetic properties of the stimulus, whether the stimulus is a familiar lexical item, its frequency of usage, and whether the stimulus is confusable with other words. Boothroyd and Nittrouer (1988) quantified the advantage of lexical status afforded to listeners by comparing the recognition of familiar CVC words with CVC nonsense (though phonotactically English) syllables using the j-factor model. While they did not evaluate contextual effects due to usage frequency or neighborhood density, the design is well suited to quantification of those effects as well. The present study is an investigation of how lexical status, frequency, and neighborhood density affect the speech recognition in noise, through a replication and an extended analysis of the first experiment in Boothroyd and Nittrouer. The j-factor model, proposed as a metric of context effects insensitive to overall performance level, is used to quantify the effects of these factors. All of these factors are measured with the j-factor model. The effect of neighborhood density is particularly interesting because it is primarily due to the first two segments.

Boothroyd and Nittrouer propose two measures of context effects that are relatively insensitive to the degree of signal degradation or overall performance level. The present study uses the second of these measures, the j-factor, which quantifies the recognition of a whole as a function of the recognition of its parts. From probability theory, the recognition probability of a whole is the product of the marginal recognition probabilities of its parts. For CVC syllables, $p(\text{syll}) = p(C_1)\ p(V)\ p(C_2)$. Assuming the recognition probabilities of individual segments or phonemes in CVC syllables are statistically independent and

* Special thanks to Katherine Cassady and Sherene Flemmings for running the experiment. A preliminary version of this work was presented a Cognitive Science Cognitive Neuroscience workshop in August 2001 at the University of Michigan, organized by Julie Boland and Rick Lewis. I am grateful to Pam Beddor, John Kingston, Terry Nearey, and other audience members for helpful comments. All errors are my own.

approximately equal (Fletcher, 1953), $p(\text{syll}) = p(\text{seg})^j$, where j represents the number of independently perceived segments in the syllable. The j-factor can be empirically determined by calculating the logarithms of recognition probabilities of whole syllables and segments in an identification task, yielding $j=\log(p(\text{syll}))/\log(p(\text{seg}))$. A finding of j=n (where n=3 for CVC stimuli) is consistent with independent recognition of the segments and implies that listeners are not exploiting contextual information. The reduction of j below n is a measure of the effect of context. At the limit of j=1, the recognition of any one segment is all that is needed to recognize the whole.

In the first experiment of their study, Boothroyd and Nittrouer measured j-factors for CVC words (j=2.46±0.08) nonsense syllables (j=3.07±0.14), concluding that j=3.07 for nonsense targets is consistent with perception of three independent units. The finding of j=2.46 for word targets is interpreted by Boothroyd and Nittrouer as a measure of the contextual advantage for words.

The j-factor reduction indicates that the higher recognition probabilities of meaningful syllables are due in part at least to the higher predictability of words relative to nonwords (cf. Allen, 1994). The j-factor quantifies this lessening of statistical independence among the segments of meaningful syllables.

On the basis of a computational simulation of Boothroyd and Nittrouer's experiment, Nearey (1998) suggests that the j-factor effects could be reproduced in a Luce choice model of word recognition as a bias that favors words over nonsense syllables. If the j-factor measures bias, then manipulations of bias in a word recognition task should affect the j-factor. If facilitation for high frequency words is the result of a bias (Broadbent, 1967; Norris 1986), then high frequency words are predicted to have lower j-factors than low frequency words.

Potential confusors to a given stimulus can affect recognition (Savin, 1963). The neighborhood activation model (NAM; Luce & Pisoni, 1998), quantified in (1), proposes that phonetic neighbors compete with the actual target for activation in a Luce choice model. Degree of phonetic overlap between the neighbor and the target stimulus determines the degree of competition. The log usage frequency is used as a weight for both the target and its neighbors.

$$(1)\quad p(ID_S) = \frac{p(S|S)\log(freq_S)}{p(S|S)\log(freq_S) + \sum_j p(N_j|S)\log(freq_j)}$$

The probability of identifying a stimulus S is $p(ID_S)$; $p(S|S)\log(freq_S)$ is the frequency-weighted stimulus word probability of S given S (FWSWP), and $\sum_j p(N_j \mid S)\log(freq_j)$ is the sum of the frequency-weighted probabilities of each neighbor N_j of S given S (FWNP).

For empirical evaluation of the model, Luce and Pisoni use the Kucera-Francis (Kucera & Francis, 1967) usage frequencies, and their own confusion matrices of nonsense syllables in noise. The conditional probability of an item is estimated by multiplying the conditional marginal probabilities of the constituent segments obtained from the confusion matrices.

Accuracy should be positively correlated with FWSWP, the stimulus probability based on acoustic-phonetic salience weighted by frequency, but negatively correlated with FWNP, the frequency-weighted probability of competitors. These qualitative predictions as well as the quantitative predictions of (1) are borne out in experiments reported by Luce and Pisoni.

The j-factor model can be applied to the parts of (1) to measure the contextual advantages of words with high and low values of FWSWP and FWNP. In the case of the FWSWP, only variation from usage frequency and not stimulus probability should be measurable with the j-factor. It may be that the stimulus probability factor dominates the frequency weight, in which case the FWSWP should not have any context effect, as measured by the j-factor.

On the other hand, neighborhood density, quantified by the FWNP, should be correlated with the j-factor if the j-factor is inversely related to bias, as suggested by Nearey. Consider the case of a listener perceiving partial phonetic information of a target word in a dense phonetic neighborhood. Given the partial phonetic information, the probabilities of non-target potential responses are large, so any bias in favor of the target will be reduced. If the partial phonetic information delimits a sparse phonetic neighborhood, the probabilities of the non-target competitors are low, and bias for the target should be high. Under this account, words with high values of FWNP (low bias) should have high j-factors, while words with low values of FWNP (high bias) should have low j-factors.

Boothroyd and Nittrouer's design offers an opportunity to test these predictions of frequency and neighborhood density, since the words span a range of usage frequencies, and are phonetically balanced with the nonsense syllables.

1. Method

The procedure for Boothroyd and Nittrouer's Experiment 1, in which participants identified CVC nonsense and word syllables at different noise levels, was followed as closely as possible, except that stimulus presentation and response collection was done online. Proportion correct of phonemes and whole syllables of different subsets of the test items were subsequently used in j-factor analyses.

Forty-three young adults were recruited from an undergraduate linguistics course at the University of Michigan and participated for course extra credit. All were native speakers of English and reported no known hearing problems.

The same lists of CVC syllables developed by Boothroyd and Nittrouer, consisting of 120 words and 120 nonsense items were used for this study. Both the word and nonsense syllable lists were phonetically balanced such that the phonemes in the sets of 10 initial consonants /b p d t k s h m l r/, 10 vowels /i ɪ eɪ ɛ u oʊ ɔ æ ɑ aɪ/, and 10 final consonants /p d t g k s z m n l/ were evenly distributed in the word and nonsense syllable lists.

Each item was read by the author, a native speaker of American Midwest English, in the carrier phrase "You will write ... please" in a sound-treated room and was recorded to DAT with a Realistic Highball microphone and a Tascam DA-30 digital tape deck at a sampling rate of 48 kHz. The recording of each item embedded in the carrier phrase was converted to a WAV file at the same sampling rate and stored on computer disk. The overall level of each stimulus was adjusted so that the peak amplitudes of all stimuli were matched.

The experiment was run using software running in the Matlab (version 6.1) environment on four Windows NT laptop computers in an anechoic chamber. Signal-dependent (though uncorrelated) noise (Schroeder, 1968) was added online at one of four S/N ratios (−14 dB, −11 dB, −8 dB, −5 dB). The resulting stimuli in their carrier phrases were presented for identification binaurally via AKG headphones with the volume set to a comfortable listening level, presented in 24 random blocks of 10 random targets, each block containing all words or nonsense syllables.

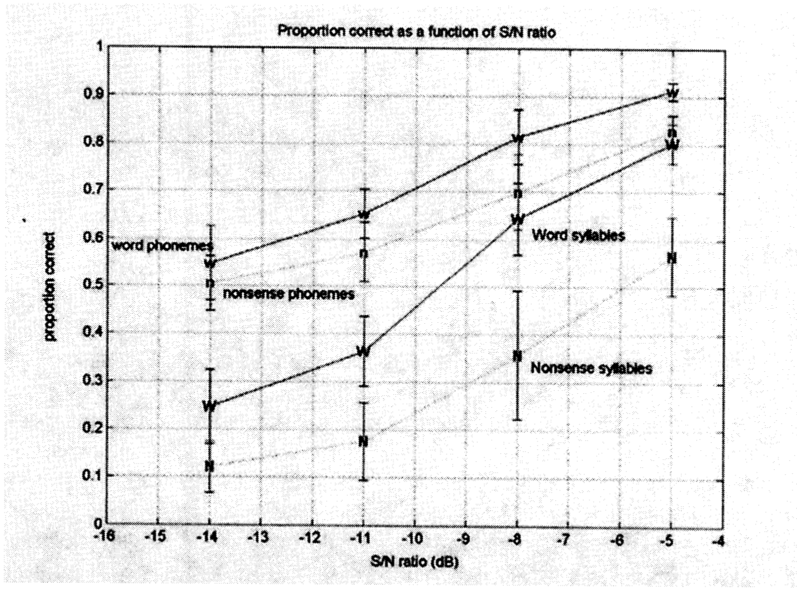
Thirty-seven participants were randomly assigned to one of the four S/N ratios (11 participants at −14 dB, 9 at −11 dB, 9 at −8 dB, and 8 at −5 dB). Six participants at the beginning of the study were assigned to other S/N ratios (1 at −10 dB, 2 at −9 dB, 2 at −7.5 dB, and 1 at −4 dB) in order to determine a range of S/N ratios for performance levels approximate of those in Boothroyd and Nittrouer. All participants were instructed in writing that they would be listening to real and nonsense consonant-vowel-consonant syllables of English presented in a carrier phrase with noise, and were to type what they heard using standard English orthography for both the words and nonsense items. A brief list of examples of English orthography for spelling nonsense items was provided.

2. Results and Analysis

Each phoneme response was scored as correct if it matched the corresponding stimulus phoneme, and incorrect otherwise, with the following adjustments. First, /ɑ/ and /ɔ/ were counted as matching vowels. In their stimulus list preparation and response analysis, Boothroyd and Nittrouer regarded the vowels /ɑ/ and /ɔ/ as distinct phonemes, and these distinctions were maintained in the preparation of the stimuli for the present study. However, these vowels are merged in the English spoken by many of the participants, and were therefore counted as the same vowel for scoring purposes. Missing consonants were scored as "null" responses and incorrect. If a cluster was reported for one of the consonants, it was scored as "other" and incorrect, unless one of the elements was a correct response and the epenthetic consonant occurred between the vowel and the correct consonant. In those cases, half were scored as vowel errors ("other") and half as consonant errors ("other").

Using the above criteria, the observed probabilities of correct recognition of nonsense phonemes, word phonemes, nonsense syllables, and word syllables for each participant were calculated and are plotted in (2).

(2)



The range and pattern of performance is quite similar to those reported by Boothroyd and Nittrouer. The S/N ratios are about 11 dB lower in the present study, which may be due to differences in the quality of the stimuli, type of noise (they used spectrally-shaped white noise that was the same level for all of the stimuli of a given S/N ratio, instead of the signal-correlated noise used here), or the experimental procedure.

Proportions correct of phonemes and syllables for each condition were converted into j-factors for each participant and averaged for an estimate of the j-factor for each condition. Because measurement errors for probabilities near zero or unity have a large effect on the estimate of the j-factor, if either the phoneme or syllable probability was less than 0.05 or greater than 0.95, the resulting j-factor was not included in the calculation of average j-factor or subsequent statistical tests, following Boothroyd and Nittrouer.

3.1. Lexical status

The j-factors averaged across participants for the high and low context condition for each comparison are shown in (3). There is no significant difference between the nonsense syllable j-factor j=3.07 and n=3 (t(40)=1.69, p=0.0991), as predicted by independent perception of phonemes in nonsense syllables and consistent with Boothroyd and Nittrouer. A paired comparison of words and nonsense syllable j-factors shows the word j-factor mean, j=2.35, to be significantly less than the nonsense syllable j-factor mean (t(40)=11.196, p<0.00001), diagnostic of phonemes in words not being perceived independently of each other, or of a bias in favor of words.

(3)
The *j*-factors for high and low context conditions of lexical status, frequency, FWNP, and FWSWP averaged over participants.

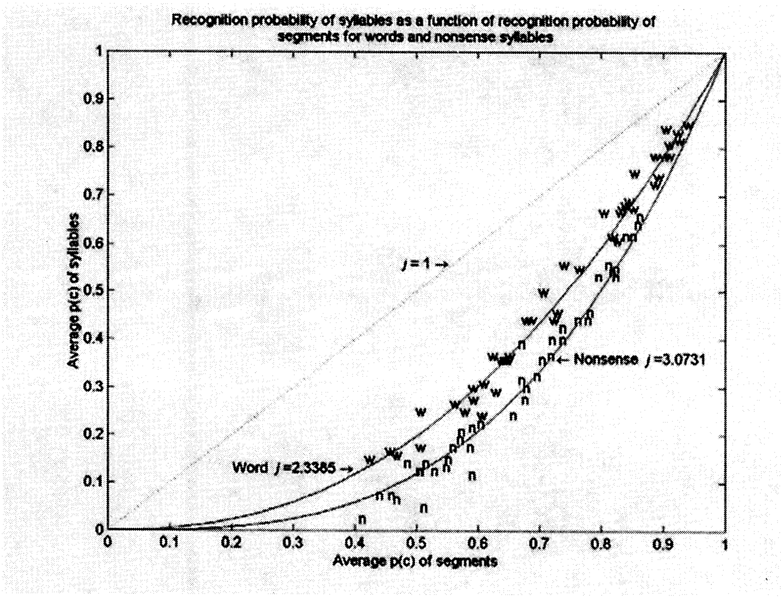
Comparison	*j*-factors with 95% confidence intervals	
	High context	Low context
Lexical status: word/nonsense syllable	2.34 ± 0.08 (B&N 2.46 ± 0.08)	3.07 ± 0.08 (B&N 3.07 ± 0.14)
High freq./Low freq.	2.25 ± 0.10	2.46 ± 0.10
Low FWNP/High FWNP	2.11 ± 0.09	2.61 ± 0.11
High FWSWP/Low FWSWP	2.46 ± 0.13	2.39 ± 0.09

Values for each participant are plotted in (4), with best-fitting curves for words and nonsense syllables. Each point represents average syllable recognition probability as a function of average phoneme recognition probability for words or nonsense syllables of a single individual. The j-factors are not significantly correlated with phoneme recognition probability for either nonsense syllables (R^2=0.0650, F(1,40)=2.7104, p=0.1077) or for words (R^2<0.00001, F(1,42)=0.0002, p=0.9998). The lack of correlation with phoneme recognition probability and the good fit across the range of recognition probability supports the use of the j-factor as an index of context effects independent of recognition probability.

3.2. Word Frequency

Word frequency effects were measured by dividing word trials into high and low frequency groups using the median log Kucera-Francis frequency of all the words (3.29) as a cutoff. The high frequency words have a mean log Kucera-Francis frequency of 4.90, while the low frequency words have a mean log Kucera-Francis frequency of 2.46. Average phoneme and syllable recognition probabilities were calculated for high and low frequency words for each participant, and converted to j-factors as shown in (3). As expected, the high frequency words have a lower j-factor (j=2.25) than the low frequency words (j=2.46), consistent with the prediction of a bias in favor of high frequency words, with the magnitude being about a third of the size of the lexical status effect. A paired comparison indicates that the difference between the mean high and low frequency j-factors is significant (t(42)=3.809, p=0.00045). The difference is also significant if a familywise α=0.05 error rate is maintained using a Bonferroni criterion for four tests (the cutoff for familywise α=0.05 for four tests is t=2.4949 (Hays 1994:1007)).

(4)



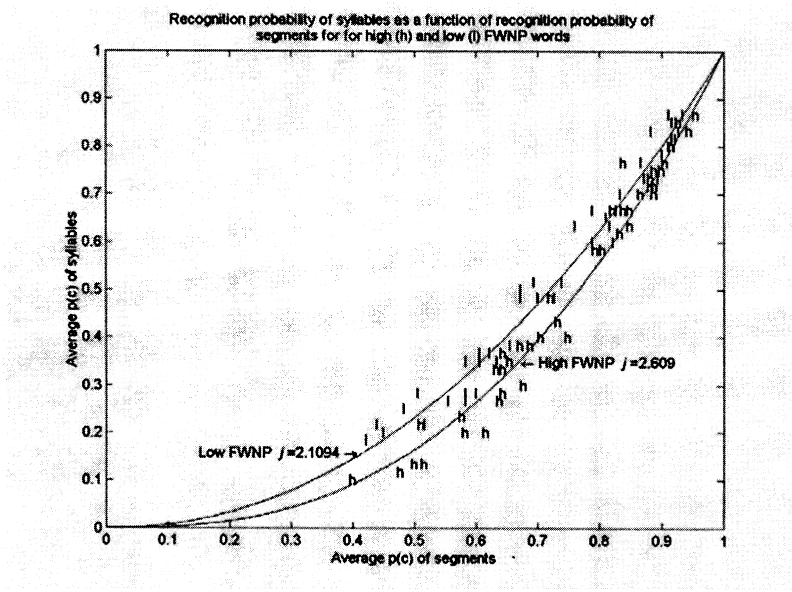
3.3. Neighborhood Density and Stimulus Word Probability

An online version of Webster's Pocket Dictionary (*Webster's Seventh Collegiate Dictionary,* 1967; Nusbaum, Pisoni, & Davis, 1984) was used to determine the neighbors for each target word. All neighbors differed with the target by one segment, with a substitution or a deletion for the third non-matching segment, but no insertions. In order to compute conditional phoneme probabilities for computing the FWNP and FWSWP, confusion matrices were calculated for C_1, V, and C_2 by collapsing the nonsense syllable responses across all participants. The cells of each confusion matrix were used to calculate the conditional probabilities needed to compute FWNP and FWSWP for each target word.

The median FWNP and FWSWP values were used as cutoffs to divide the target words into high and low FWNP groups and high and low FWSWP groups. A j-factor was calculated for each participant for the high and low groups of both FWNP and FWSWP; the average j-factors are in (3).

A paired comparison of the mean j-factors low and high FWNP words indicates that the difference is significant ($t(41)=8.691$, $p<0.00001$). The magnitude of the effect is nearly as large as that between words and nonsense syllables, confirming the expectation that targets with sparse phonetic neighborhoods (j=2.11) have a large contextual advantage over targets with dense phonetic neighborhoods (j=2.61). Data for individuals for words with high and low values of FWNP, and average j-factors are plotted in (5).

(5)



The effect of FWSWP is nonsignificant (t(37)=1.050, p=0.3004), with j=2.46 for high FWSWP words and j=2.40 for low FWSWP words. This nonresult is consistent with the phoneme probability overwhelming the frequency weighting, resulting in accuracy differences but not in j-factor differences.

The contextual advantage afforded to words with sparse phonetic neighborhoods can be further investigated by subdividing the neighborhood of a given target into those neighbors sharing C_1V with the target, those sharing VC_2, and those sharing C_1C_2. The FWNP can be calculated for these three different neighborhoods to divide the target words into high and low FWNP groups for j-factor analyses. The results of these analyses are in (6). The C_1V neighborhood shows a significant difference (t(39)=6.507, p<0.00001) between low (j=2.15) and high (j=2.52) FWNP targets in favor of words with low density neighborhoods. The difference in j-factor for the C_1C_2 neighborhood is slight and in the opposite direction as might be expected (low density, j=2.41; high density j=2.29) but significant (t(39)=2.386, p=0.0220). However, the difference for the C_1C_2 neighborhood is *not* significant if a familywise α=0.05 error rate is maintained using a Bonferroni criterion for three tests (the cutoff for familywise α=0.05 for three tests is t=2.3954; Hays 1994, p.1007). The difference for VC_2 neighborhoods is not significant (t(41)=0.004, p=0.9966). The vast majority of the difference in the j-factors for low and high FWNP words seems to arise from C_1V neighborhood structure.

It is unlikely that the neighborhood analysis results are because of an excessive number of C_1V neighbors relative to the other two types of neighbors.

Column 1 of (6) shows the mean number of neighbors per target. The average 20.8 neighbors per target are roughly equally divided between the three types of neighbors.

(6)
Neighborhood *j*-factor analysis. The following values are reported for each type of neighborhood: the average number of neighbors (with S.D.), the average *j*-factors with 95% C.I. for low FWNP (low density neighborhood), high FWNP words (high density neighborhood) words.

Neighborhood type	Mean number of neighbors	*j*-factor for low FWNP words	*j*-factor for high FWNP words
All neighbors	20.8 (4.8)	2.11 ± 0.09	2.61 ± 0.11
C_1V neighbors	7.1 (2.4)	2.15 ± 0.11	2.52 ± 0.10
VC_2 neighbors	5.7 (2.2)	2.35 ± 0.08	2.34 ± 0.08
C_1C_2 neighbors	8.0 (3.5)	2.41 ± 0.08	2.29 ± 0.11

4. Discussion

The interpretation of the result j=n, as was found for nonsense CVC syllables with n=3 segments, is consistent with the hypothesis that the constituent segments of syllables are perceived independently. But what does the result j<n mean? Boothroyd and Nittrouer suggest it is a measure of the reduction of independent perceptual units. Words are perceived as consisting of j=2.35 independent units, with each phoneme consisting of about 0.78 units.

Nearey (1998), proceeding from a computational simulation of Boothroyd and Nittrouer's results, suggests that small reductions (around 1 or less for n=3) in the j-factor could arise from a bias in favor of particular items in a Luce choice model. Results of 2<j≤n are consistent with independent perception of *n* segments, and reduction of *j* below *n* quantifies the amount of bias involved for those items. The present results for word frequency, that high frequency words have lower j-factors than low frequency words, are entirely consistent with this interpretation and would support bias accounts of word frequency effects in word recognition. However, it is important to note that the j-factors reported here are averages over groups of words. Under certain situations of high context, it is possible that a gestalt model of word recognition, in which words are perceived as wholes, would be a more appropriate interpretation for results of j=1.

The neighborhood density results suggest that the implementation of bias must be understood with reference to temporal distribution of the acoustic-phonetic information. Recall that, as predicted, words with sparse neighborhoods had a lower j-factor than words with dense neighborhoods, consistent with a bias favoring words from sparse neighborhoods. Importantly, this result largely holds true for neighborhoods defined by CV neighbors but not VC neighbors.

Correct perception of the beginning of a word in a sparse phonetic neighborhood delimits a small set of potential candidates. A listener can then focus attention on just those phonetic features that distinguish the members of this small set to achieve correct recognition despite reduced acoustic-phonetic information present at the end of the syllable. This account of how listeners use bias, based on a dynamic analysis of the effects of neighborhood density, offers support for the dynamic aspects (but perhaps not the strict autonomy) of the cohort theory of word recognition (Marslen-Wilson, 1989).

The lack of any significant j-factor effect for VC neighborhood indicates that in open-response identification, contextual information from correct recognition of syllable-final material is not used in order to reevaluate or sharpen the perception of earlier-occurring syllable-initial material in the same way that contextual knowledge guides perception of upcoming material (but cf. Salasoo & Pisoni, 1985). What is measured by the j-factor for words in sparse C_1V neighborhoods does not just narrow the set of possible word candidates so that guesses can be more effective, but seems to have genuine perceptual effects. If the role of bias were merely to narrow such a set, along the lines of what Broadbent (1967) calls the sophisticated guessing model, then one would expect a significant reduction of the j-factor for words in sparse VC_2 neighborhoods as well.

This asymmetric effect is consistent with claims for more robust acoustic-phonetic information in the speech signal for onsets than for codas. Wright (2001), for example suggests some active compensatory strategy on the part of listeners, since syllable-final consonants must be correctly identified in languages that have them, such as English. If the strategy takes place according to the account outlined above, with listeners focusing attention whenever expected neighborhood density permits, then recognition rates for codas should be low when expected neighborhood is dense, and high when the expected neighborhood is sparse.

A preliminary comparison of the average recognition rates for C_1 and C_2 is consistent with these predictions. Nonsense syllables, whose segments are highly unpredictable, show $p(C_1)\text{-}p(C_2)= 0.2124$, providing a baseline for the advantage of onsets over codas. Words, with $p(C_1)\text{-}p(C_2)=0.0798$, show lower differences than the nonsense syllable difference. The recognition rates for C_1 and C_2 of low C_1V FWNP words—sparse expected neighborhood—are nearly equal at $p(C_1)\text{-}p(C_2)= 0.0226$, while the comparable rate for high FWNP words—dense expected neighborhood—is $p(C_1)\text{-}p(C_2)=0.1704$.

Frequency of usage and neighborhood density effects could be explained here in terms of a criterion bias shift, supporting feedforward models of top-down effects in word recognition such as Merge (Norris, McQueen, & Cutler, 2000) or FLMP (Massaro, 1998) over feedback models such as TRACE (McClelland & Elman, 1986). However, given the dynamic nature of the neighborhood density effect as reported here, the operationalization of bias appears to narrow the difference between feedforward and feedback models. The current findings indicate that bias appears to improve the efficiency of perception of phonological structure in the coda when acoustic-phonetic information is impoverished. This interpretation of bias as measured by the j-factor may represent attentional priming effects (Grossberg & Stone 1986).

5. Conclusion

Support has been provided for Boothroyd and Nittrouer's j-factor as a robust and replicable measure of the effects of context in human speech perception. The j-factor represents the number of perceptually independent parts within a whole, and can be interpreted as a bias in favor of words over nonsense syllables, words with higher usage frequencies over words with lower usage frequencies, or of words from sparse neighborhoods over words from dense neighborhoods. The neighborhood density effect is dynamic, such that the neighborhood is primarily determined by the first two segments of a CVC word. This dynamic effect appears to improve perception of codas of CVC words in sparse CV neighborhoods.

Future work will use the j-factor model to investigate context effects in other types of stimuli besides English CVC monosyllables, such as longer words, and other languages with more different syllable structure. Investigation of the model's assumptions is also planned, such as approximately equal recognition rates of all phonemes, and whether segments (as opposed to syllables or features) are the proper units of analysis.

References

Allen, J. 1994. How do humans process and recognize speech? *IEEE Transactions on Speech and Audio Processing* **2**: 567-577.

Boothroyd, A. & S. Nittrouer. 1988. Mathematical treatment of context effects in phoneme and word recognition. *Journal of the Acoustical Society of America* **84**: 101-114.

Broadbent, D. E. 1967. Word-frequency effect and response bias. *Psychological Review* **74**: 1-15.

Fletcher, H. 1953. *Speech and hearing in communication.* New York: Van Nostrand.

Goldinger, S. D. 1997. Words and voices: Perception and production in an episodic lexicon. In K. Johnson & J. W. Mullennix (eds.) *Talker variability in speech processing.* San Diego, CA: Academic Press.

Grossberg, S. & G. Stone. 1986. Neural dynamics of word recognition and recall: Attentional priming, learning, and resonance. *Psychological Review* **93**: 46-74.

Grossberg, S. 2000. Brain feedback and adaptive resonance in speech perception. *Behavioral and Brain Sciences* **23**: 332-333.

Hays, W. L. 1995. *Statistics.* 5th ed. Fort Worth, TX: Harcourt Brace.

Johnson, K. 1997. Speech perception without normalization: An exemplar model. In K. Johnson & J. W. Mullennix (eds.), *Talker variability in speech processing*. San Diego, CA: Academic Press.

Kucera, F., & W. Francis. 1967. *Computational analysis of present day American English.* Providence, RI: Brown University Press.

Luce, P. A., & D. B. Pisoni. 1998. Recognizing spoken words: The neighborhood activation model. *Ear and Hearing* **19**: 1-36.

Luce, P. A., S. D. Goldinger, & M. S. Vitevitch. 2000. It's good...but is it ART? *Behavioral and Brain Sciences* **23**: 336.

Marslen-Wilson, W. 1989. Access and integration: projecting sound onto meaning. In W. Marslen-Wilson (ed.) *Lexical representation and process* Cambridge, MA: MIT Press.

Massaro, D. W. 1998. *Perceiving talking faces: From speech perception to a behavioral principle.* Cambridge, MA: MIT Press.

McClelland, J. L., & J. L. Elman. 1986. The TRACE model of speech perception. *Cognitive Psychology* **18**: 1-86.

Nearey, T. 1998. On the factorability of phonological units in speech perception. Paper presented at the Sixth Conference on Laboratory Phonology, University of York, July 1998.

Norris, D. G. 1986. Word recognition: Context effects without priming. *Cognition* **22**: 93-136.

Norris, D. G., J. M. McQueen, & A. Cutler. 2000. Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences* **23**: 299-370.

Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. 1984. Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words. *Research on Speech Perception Progress Report No. 10*. Bloomington: Speech Research Laboratory, Psychology Department, Indiana University.

Salasoo, A., & Pisoni, D. B. 1985. Interaction of knowledge sources in spoken word recognition. *Journal of Memory and Language* **24**: 210-231.

Savin, H. 1963. Word-frequency effects and errors in the perception of speech. *Journal of the Acoustical Society of America* **35**: 200-206.

Schroeder, M. R. 1968. Reference signal for signal quality studies. *Journal of the Acoustical Society of America* **44**: 1735-1736.

Webster's Seventh Collegiate Dictionary. 1967. Los Angeles, CA: Library Reproduction Service.

Wright, R. 2001 Perceptual cues in contrast maintenance. In E. Hume & E. Johnson (eds.) *The role of speech perception in phonology*, 251-277. San Diego, CA: Academic Press.

José R. Benkí
Department of Linguistics
1076 Frieze Bldg.
University of Michigan
Ann Arbor, 48109-1285

benki@umich.edu

Herds of Wildebeest, Flasks of Vodka, Heaps of Trouble: An Embodied Construction Grammar Approach to English Measure Phrases

ELLEN DODGE and ABBY WRIGHT
University of California, Berkeley

1. Introduction

In this paper we examine English Measure Phrases (EMPs), for example, *herd of zebras, cup of coffee, square of cloth,* and *miles of beach.*[1] Although EMPs internally resemble other phrases of the same form (a noun followed by *of* followed by a noun phrase), they interact with other elements in the clause in unique ways. On the basis of the phenomena studied in this paper, it will be necessary to distinguish EMPs from other *N-of-NP* phrases. First, EMPs show unusual modificational patterns. Consider the following corpus examples:

(1) The only other colours are provided by a snaking blue-black <u>ribbon</u> of tarmac… (BNC)[2]

(2) Ronni blinked down at the glittering sea that was divided from the craggy landscape by a broad silken <u>ribbon</u> of glittering pale gold sand. (BNC)

In (1) *blue-black* is understood as modifying *tarmac*, not *ribbon*, which serves instead to specify the dimensional boundaries of the tarmac. In (2) *broad* refers to the dimension of *ribbon*, while *silken* describes a quality of the sand.

EMPs also exhibit varying agreement patterns. Consider the following:

(3) To ease the pressure, a <u>truckload</u> of Commandos were taken to the rear, where they could relax for a couple of days. (BNC)

(4) A <u>herd</u> of zebras, hence, produces about a quarter to a third of its weight in prey carcasses per year. (BNC)

---

[1] We would like to thank Charles Fillmore, the FrameNet Project, George Lakoff and Jerry Feldman, the Neural Theory of Language Group, and Andreas Kathol.

[2] For our corpus data we used the British National Corpus (BNC).

In example (4) *a herd* is construed as singular and shows singular agreement. In sentence (3) the opposite is the case: *a truckload of Commandos* is construed as plural and shows plural agreement.

Finally, the EMP construction allows for the noun phrase nested inside the EMP to satisfy the selectional restrictions of the verb.

(5) A toddler was fighting for life last night after he swallowed a bottle of lethal acid at a doctor's surgery. (BNC)
(6) The bottle of champagne took five attempts to break... (BNC)

In sentence (5) *a bottle of lethal acid* is an *amount* of acid; it is unlikely that the child swallowed both the acid and its container. In (6), however, it is clear that the noun phrase, *a bottle of champagne,* is about a bottle since *break*'s selectional restrictions require that its direct object be solid or a functional item, which can not be satisfied by *champagne.* In these examples the verb's selectional restrictions can be satisfied by either of the elements in the phrase. Furthermore, EMPs can be nested and the verb can select for the innermost noun. Consider example (7) and note also the plural agreement on the verb; *fall* agrees with *bits*, an intervening measure element between *load* and *paper.*

(7) When I open it a load of bits of paper fall out and flutter to the ground. (BNC)

We will argue that EMPs can be insightfully analyzed using the fledgling framework of *Embodied Construction Grammar* (Chang et al. 2001), which we will outline in Section 4. In this form of construction grammar, linguistic constructions map phonological forms to embodied conceptual representations.

2. English Measure Phrases and Their Function

As mentioned above, *herd of zebras, cup of coffee, square of cloth,* and *miles of beach* are all examples of EMPs. They are of the form *X of Y*, where *X* is a count noun and *Y* is a mass noun or plural count noun. We will refer to plural count nouns as multiplex since this captures the undifferentiated nature of the individuals. So for *cup of peas*, *cup* will be described as the *X-element* and *peas* as the multiplex *Y-element.* We will assume a traditional constituent analysis for EMPs. The EMP is a nominal element which consists of a noun modified by a prepositional phrase. We are not making syntactic claims, *per se*, about the internal form of the phrase (see Akmajian & Lehrer 1976 and Gawron 2002 for proposals about the internal structure).

What we are calling English Measure Phrases are part of a larger group of constructions sometimes called *partitive constructions* (Quirk et al. 1985). In this paper, we will not be looking at quality partitive nouns (such as *type of font*), temporal measures (*second of boxing*), or phrases that describe part-whole relations (*sip of the wine*). Actual lexical items in English may appear in more than one type of phrase. For example, while *hunk of cheese* is an EMP, the phrase *a hunk of that cheese log* describes a part-whole relationship.

We will now propose a typology of EMPs (see Lehrer 1986 and Svensson 1998 for related descriptions). Our typology is constructed on the basis of differences in both the *X*- and *Y-elements*; types vary as to whether mass and/or multiplex *Y-elements* are allowed, and also as to what sort of additional classifying information the *X-element* provides. Due to the graded nature of these distinctions, the different types are intended to indicate relevant areas along a continuum rather than being strictly disjoint and distinct types.

2.1 Types of EMPs

Container-Measures are phrases such as *glass of tea, busload of children*, or *cupful of wine*, and can have a *Y-element* that is either a mass or a multiplex. They constrain and individuate fluids and other entities that must be physically contained to be measured. Information about the typical size of each kind of container gives inexact information about the amount of the *Y-element*. In this group are also those *X-elements* formed by adding *–ful* or *–load* to objects. Objects that are not typically conceptualized as containers usually require this suffix: for example, *a handful/*hand of peanuts.*

Standard-Measures such as *yard of cloth*, *cup of beans*, or *gallon of milk* are measurements on a standardized scale. These can have either a mass or multiplex entity as their *Y-element*, though a multiplex can only occur in measures of volume. Standard-Measures provide exact information about the length, area, volume or mass of a substance.

Dimensional-Boundaries are phrases like *stick of butter*, *sheet of paper* or *ribbon of sand.* The *Y-elements* which appear in this type are all internally cohesive and thus are only mass nouns, not plural count nouns. The *X-element* may provide relative size information; for a given substance, *a chip* is generally smaller than *a hunk.* Additionally, in the Dimensional-Boundary type of EMP an *X-element* such as *stick* tells us that the *Y-element* is a fairly rigid, predominantly one-dimensional object (or *a long thin thing*, as is familiar from classifier languages), thus providing information about shape, dimensionality and rigidity.

The Configuration type has members such as *line of trees*, *circle of crows, stack of papers* or *heap of stones*. In this type, individuals are arranged in a particular configuration, so the *Y-element* must be a multiplex noun. The *X-element* provides information about the configuration's shape and orientation.

The final type is Collection-of-Members which has examples such as *herd of zebras*, *swarm of bees*, or *team of soccer players* and requires a multiplex *Y-element*. In this case the *X-element* provides information about the social or functional relationships between the individuals.

Some lexical items may occur in more than one of the EMP types. For example, *a circle of crows* is a Configuration since it describes the arrangement of a set of objects in terms of the shape of the boundary, similar to *ring of crows*;

however, *circle of cloth* is of the Dimensional-Boundary type since it indicates that the cloth is a flat object that extends to a circular boundary.

Figure 1: Corpus Examples from the BNC

Container-Measure
The court also heard from medical experts who found traces of half of a bottle of whisky, and cocaine, in Kin's body. (BNC)
Standard-Measure
A foot of snow fell in the Scottish Highlands yesterday, bringing chaos to road, rail, and air services. (BNC)
Dimensional-Boundary
She folded the silver square of wrapping paper in half. (BNC)
Configuration
A man was sprinkling petrol on a heap of sprouts to help them burn. (BNC)
Collection-of-Members
You may fatten a couple of pigs to kill and keep a small flock of laying hens or ducks on household scraps. (BNC)

2.2 The Function of EMPs

EMPs measure, individuate, and give classificatory information about substances. The *X-element* provides the structure by which an unmeasured and internally undifferentiated *Y-element* is individuated. Although in any given instance a mass noun may be bounded and have a particular size and shape, this information is situational, rather than being an inherent feature of that type of entity. When a substance is simply a mass without any measure or quantification, it is not countable. EMPs serve to make mass nouns countable. Multiplex items are easily quantified through counting, but the count does not supply information about the shape or configuration of the group, the relationships between elements of the group, or even whether or not the multiple entities form a single group at all. In this case the EMP serves to make one countable instance of a multiplex of entities, indicating the relationships between them. Once we have a countable instance of a mass or multiplex, we can describe interactions with that object; this is something that count nouns get for free, but mass nouns do not.

EMPs show significant similarities to the classifier systems of prototypical classifier languages, such as Yucatec Maya (Lucy 1992). Properties provided by the *X-element* in the EMP such as shape, dimensionality, extent, orientation and consistency have been considered universal parameters of classification in studies of classifier systems (Aikhenvald 2000). Additionally, the functions performed by EMPs and classifiers are similar. Lucy (1992, 2000) and others have suggested that a classifier form indicates the individuation status (unit or quantity) of a lexical noun indicating type or quality. Under such a view, the nouns in classifier languages are all of the English mass noun type, and require classifiers to be individuated.

Although similar, there are morphosyntactic differences between prototypical classifiers and EMPs. Prototypical classifiers are closed class items with no independent status, and are used obligatorily. In English there are no individual lexical items dedicated as classifiers; instead, certain lexical items take on a classifying function when present in the EMP (see Lehrer 1986 and Svensson 1998). Furthermore, in English the classifying *X-element* must appear within a noun phrase with the *Y-element*, which is not necessarily true in classifier languages. By using a construction grammar approach we can readily recognize the function the *X-element* takes on by becoming a part of this construction.

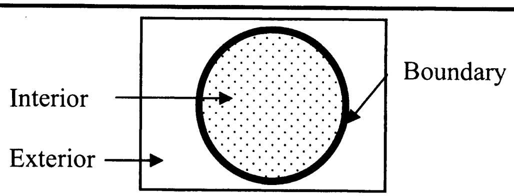
3. The Semantics of English Measure Phrases

As seen in Section 2, our criteria for defining and analyzing EMPs are chiefly semantic. To incorporate these semantic properties into our analysis of EMPs, a more structured account of the meanings involved is needed. In this section we look both at the prototypical scene which motivates EMPs, as well as the underlying cognitive regularities by which we structure such experience.

As noted in the last section, the *Y-elements* in EMPs and the nouns in classifier languages are structured along parameters which appear to be universals. This is no surprise in an embodied concept of meaning; all human beings share the same perceptual, motor, and cognitive systems and engage in many of the same functional interactions with the world. We suggest that one reason these classification parameters seem largely universal is that classifiers and EMPs both exploit the same basic image schematic structures. We define image schemas as representations of regularities in our perceptual, motor and cognitive systems which structure our experiences and interactions with the world. EMPs exploit a variety of image schemas to provide different sorts of structure. Below is a representation of the *Boundary* schema, an image schema which provides the structure EMPs utilize to individuate substances.

Figure 2: Schematic Representation of Image Schema

Boundary Schema
<u>Roles</u>:
Interior
Boundary
Exterior



When a substance is individuated, it can be conceptualized as a whole bounded object, with both an interior area and a boundary. The boundary separates the substance from its surroundings, the exterior. Other schemas may combine with the Boundary schema, allowing, for example, the specification of the relative dimensional extent and/or the orientation of the bounded object.

In addition to image schematic structure, the prototypical *scene* or *frame* of measurement which motivates EMPs needs to be described in a structured way. Frame semantics provides such a structure. According to frame semantics, "words represent categorizations of experience, and each of these categories is underlain by a motivating situation occurring against a background of knowledge and experience." (Fillmore 1982: 112). The semantic frame of measurement, or the prototypical measuring scene, includes an agent who partitions and separates a quantity of a substance (mass or multiplex) from an undifferentiated mass in order to handle, transport or otherwise interact with the substance. This is prototypically done using a container, although a physical container need not be present, especially with an internally cohesive mass or arrangement (as in *bouquet of flowers* or a *chunk of cheese*). English speakers use *in* or *out* with many of these *X-elements*, indicating that they are often construed as metaphorical containers.

(8) We cut the bread **into** chunks.
(9) We put the leaves **in** a pile.
(10) He tried to get **out of** the gang, but was held **in** by threats of violence.
(11) He divided the old farm **into** acres to sell to developers.

The idea of containers is a complex one. Most physical containers are artifacts whose primary function is to contain other entities. The container typically acts as a rigid boundary which restrains the contents located in its interior. Fullness plays a role in these containers; flasks can be full of vodka and glasses can be half filled with milk. When one adds more content to the container it becomes fuller but usually not larger. However, some containers have flexible boundaries; they are always completely filled with their contents and grow larger upon the addition of more contents. The boundary of the contained substance is always the same size and shape as the container itself. The metaphorical containers found within this construction are all of this type. One cannot say *the full square of cloth*, because the square is always completely filled by cloth. Similarly, a gang or herd is completely filled by its members; the addition of more members will increase *the size* but not the fullness of the group.

Our understanding of EMPs depends on the background knowledge of a measurement frame. Although most previous work on frame semantics has involved verbs which evoke frames that structure entire clauses, frame analyses are also possible for nouns or noun phrases which evoke events and scenes. For most frame-evoking verbs, one or more of the scene's human participants are typically present within the clause. EMPs, however, evoke the measurement frame but do not profile the agents or measurers who are doing the measuring; these measurers are completely absent within the phrase, and may also be absent from the larger clause in which the EMP occurs.

4. Embodied Construction Grammar

We will incorporate both frame semantics and image-schematic structure in our representation of EMPs, to follow in section 5. The framework in which our analysis will be couched is Embodied Construction Grammar (ECG), a grammatical framework being developed by the Neural Theory of Language group at the University of California, Berkeley (Chang et al). ECG builds on earlier work in construction grammar (Goldberg 1995, Kay & Fillmore 1999, Ginzburg & Sag 2000) and shares key commitments of other types of construction grammar. As Kay & Fillmore (1999) describe, "The construction grammarian is required to develop an explicit system of representation, capable of encoding economically and without loss of generalization all the constructions (or patterns) of the language" (p. 2). In construction grammar, then, all patterns of a language are considered constructions and even things as "peripheral" as idioms are worthy of study and have a place within the grammar. Croft (2001) writes: "...construction grammar has generalized the notion of a construction to apply to any grammatical structure, including both its form and meaning" (p. 17).

As described in the introduction, ECG constructions pair form and meaning with the requirement that the meaning component be embodied. This requires that meaning is grounded in the body's perceptual, motor, and non-linguistic conceptual systems, along with world experience. Image schemas and frames are based on our perceptual and motor systems and our experience and are therefore grounded. ECG provides us with a representation that allows us to utilize these concepts and their roles and constraints. Our representation will be chiefly concerned with the semantic properties of EMPs, and we will show that some of their interesting grammatical properties are the result of their semantics. We will analyze EMPs as an instantiation of the EMP Construction, a pairing of form and meaning. The semantic properties of the construction license the internal noun phrase, or *Y-element*, to contribute the semantic category and number of the entire construction.

5. The English Measure Phrase Construction

As shown in Section 1, EMPs behave unusually with respect to their interactions with other elements of the grammar. Our ECG account of the EMP Construction provides crucial insight into this behavior. The semantics of the EMP are represented within the construction as a series of constraints that enforce the relationships between the two elements of the constructions. As seen in our representation of the EMP Construction in Fig. 3, below, this construction evokes the *measures* relation, which has two roles. The Measure is a subcase of Container, as described in Section 3. The Substance must be either a mass or a multiplex. The *measures* relation, the central relation of the construction, evokes two other relations, *fills* and *contains*. The Substance *fills* the Measure and extends to the boundaries of the Measure as understood through the measurement scene where an agent uses *X* to measure *Y*. At the same time the Measure *contains* the Substance. These relations taken together allow the Measure to provide shape and amount information for the Substance. In the *Constraints* section, the *X-element* is bound to the Measure role and the *Y-element* to the Substance role. The Category of the EMP is supplied by Substance; because the phrase describes an amount of a Substance, the phrase is about that Substance. The relations between Measure and Substance are made explicit by showing the bindings of each of their roles to one another. For example, Measure has a Boundary role, referred to as Measure.Boundary, and it is bound to Substance.Boundary.

Figure 3: Semantics of the EMP Construction.

EMP Construction
Evokes: *measures* Relation (which in turn evokes *contains* and *fills* relations)
Roles: Measure is a **subcase of** Container
Substance is a mass or multiplex
Relations: Substance ***fills*** Measure
Measure ***contains*** Substance
Constraints: Measure ← X-element Substance ← Y-element
Category of EMP ← Substance
Measure.Contents ↔ Substance (because of *contains* relation)
Measure.Boundary ↔ Substance.Boundary (from *measures*)
Measure.Area|Volume ↔ Substance.Amount (from *fills*)
Measure.DimExt ↔ Substance.DimExt

Thus our construction consists of two entities, the Measure and the Substance, and the relationship between them. Both of the entities play a crucial role in understanding the semantics of the construction. Because the Substance is the element that provides substance properties, it is what the entire construction is "about", which leads to the unusual circumstance where the *Y-element* becomes the category determinant of the entire noun phrase. The *X-element* contributes purely schematic, classificatory information.

6. Grammatical Repercussions

Next we will examine what insights the above constructional analysis affords us regarding the unusual grammatical behaviors of the EMP: the unusual selectional restrictions, the modificational patterns and the agreement properties described in Section 1.

There appear to be two form-identical constructions with different grammatical properties (see examples (5) and (6)). Some of these phrases made our work at the FrameNet Project[3] difficult. For example, *drink* might be followed

[3] <http://framenet.ICSI.berkeley.edu/~framenet> See Fillmore & Sato and Baker & Ruppenhofer, this volume, for more information on the FrameNet Project.

by *a bottle of champagne*; thus the semantic argument of *drink* was instantiated within the noun phrase, *a bottle of champagne*. This construction, as well as the one in example (5), is what we have described as the EMP Construction. The second, seen in (6), is at least superficially a more straightforward prepositional phrase modificational construction, which describes a container and its contents. In order to highlight the grammatical peculiarities of the EMP Construction, it will be instructive to compare these two constructions.

First we will turn to the selectional restrictions. The phrase *bottle of champagne* can describe either a bottle that is filled with champagne, or a specified amount of champagne. Consider the nearly minimal pair:

(13) The partygoers drank a bottle of champagne.
(14) The partygoers broke a bottle of champagne over the prow of the ship.

Here *drink* selects for *champagne*, and *break* selects for *bottle*. In (13) *champagne* is within the noun phrase that serves as the direct object, yet it contributes the semantic category of the entire phrase. In example (14), however, the semantic category of the phrase is contributed by *bottle*, not *champagne*, and it thus falls outside our delineation of EMPs. We found no attested examples in the British National Corpus that require *bottle of champagne* (or any other similar container phrase) to simultaneously describe a bottle and champagne. We do not find, for example, *He broke and then wiped up a bottle of champagne*. Furthermore such sentences seem to be unacceptable for native speakers. Thus, we can assert that the EMP Construction is a separate construction, which licenses the *Y-element* to contribute the semantic category of the entire phrase. This is a property of this particular construction and is motivated by the semantic properties described in Section 3. As stated in the previous section, the *X-element* serves to provide a boundary for the *Y-element* and the *Y-element* fills the *X-element*, as seen in the prototypical measurement scene; thus the phrase describes an amount of *Y*. This licenses *Y* to contribute the semantic category of the entire phrase. Thus, the *Y-element* is what predicates will be sensitive to when selecting their arguments.

Secondly, there are interesting alternations between the EMP and other *N-of-NP* constructions that involve modificational patterns. Consider the following sentences:

(15) *the Winesap growers of our apples/the growers of our Winesap apples
(16) I ate a delicious can of peas/I ate a can of delicious peas.
(17) I smashed a ?*delicious/rusty can of peas.
(18) *I ate a rusty can of peas.

It appears that modifiers of the *Y* (such as *delicious*) are licensed to appear at the periphery of the phrase only in the EMP construction (examples (15) and (16)). This property of the EMP is reducible to the same constructional description that explains the selectional restriction facts. The grammar employs a modificational construction (since all syntactic patterns, idiomatic or regular, are constructions), which will interact appropriately with the EMP Construction. In examples (16) and (17), *delicious* is therefore an acceptable modifier in the EMP construction which is about peas, but not in the prepositional phrase modificational construction, which is describing a can. Furthermore, as argued in Section 5, the *X-element* of the EMP provides amount and shape information, but cannot contribute its own physical object properties. Hence, modifiers of physical properties of the container are unacceptable, as shown in example (18).

The final interesting grammatical fact of EMPs is their agreement properties. Recall examples (3) and (4). In (3), *truckload of commandoes* triggers plural agreement and is the antecedent of a plural pronoun. However in (4), *herd* triggers singular agreement on the verb and is the antecedent of a singular pronoun. Thus, the agreement is sensitive to something besides the head noun (the *X-element* and what might be called a "syntactic head") of the EMP, since in both cases the *X* is singular but the EMP is conceived of as plural in one case.[4]

An initial hypothesis might be that the *Y-element*'s contribution of the semantic category of the whole referent will account for the plural number; however, this proves insufficient. It is not the case that singular agreement requires that the singular noun, *X*, be contributing the category of the EMP. In (19), the phrase is about zebras.

(19) The herd of zebras is/are grazing.

However, when the head noun determines the category of the phrase, it requires singular agreement; this results in the usual case in English grammar where agreement, syntactic head, and category determinant all rely on the same element of the phrase. Plural agreement can be licensed by the EMP construction. In (20) for the predicate *are growing* to be acceptable requires the EMP construction, in which the phrase is "about" zebras.

(20) The herd of zebras is/*are growing. (where *growing* refers to *herd*)

Furthermore, the particular predicate and the nature of the measure seem to make a difference in ease of plural construal for the entity. When the actions are easily understood as applying to each individual, plural agreement is more readily acceptable for most speakers than in cases where the action affects the entire group, as illustrated in (21). When *X* describes a collection, plural agreement is more readily acceptable as well, as illustrated in examples (22) and (23).

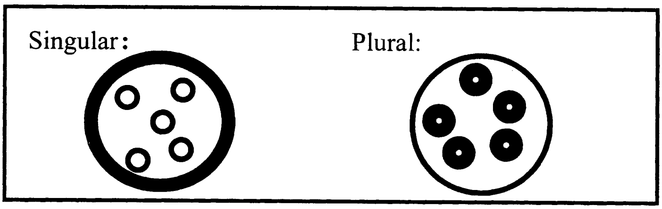
(21) An entire herd of elephants was/were shot.

[4] Since our data are from the British National Corpus, some readers might suspect the agreement facts are an artifact of British English's acceptance of plural agreement in sentences that American speakers find clearly unacceptable, as in *The government were relieved to balance the budget.* However, as examples (3) and (4) show, both patterns of agreement were found in the corpus.

(22) A herd of zebras was/were grazing
(23) A bag of peas was/?were spilled.

Consider Figure 4. The varying conception of the EMP as a singular or plural entity can be understood as a matter of construal sensitive to two properties of the EMP construction. The EMP construction serves to make an individual entity out of a mass or multiplex on the one hand, while on the other hand allowing the phrase as a whole to have the semantic category of the *Y-element*. Thus the individuation properties of the EMP construction and the licensing of the increased importance of *Y* (to the extreme of the phrase now being about *Y*) conflict in determining which entity involved in the construction will determine the number of the entire phrase. This licenses the speaker to construe the EMP with either number. The entity described by the EMP can be construed as collection of undifferentiated individuals or as salient individuals in a group. The latter construal occurs when the *X* has enough collection properties to make a plural construal possible and, of course, when the *Y* is a multiplex[5].

Figure 4: Number Represented as Construal[6]



7. Conclusion

In conclusion, we have argued that English Measure Phrases instantiate a particular grammatical construction: the EMP Construction. This construction licenses the *Y-element* to contribute the category determinant for the entire phrase, resulting in the case where the *Y-element* satisfies the predicate's selectional restrictions, and modifiers of the *Y-element* are licensed to appear at the periphery of the phrase. Furthermore, agreement was argued to be a matter of construal, licensed by the two facets of the EMP Construction: its function of individuating mass substances and its possessing the category of the *Y-element*, which can consist of a multiplex. We couched our analysis in Embodied Construction

[5] It might be suggested that plural agreement is the result of adjacency facts, with the closest noun phrase determining the number agreement on the verb. In (3), however, the appearance of *they* later in the sentence indicates that the entity *a truckload of Commandoes* is being *conceived* as a plural entity. We would not argue, of course, that adjacency effects play no role in agreement; simply that they cannot be responsible for all unusual agreement facts.

[6] Here we follow Langacker 1991, where nouns designate regions and count nouns designate bounded regions.

Grammar and argued for the necessity of a framework that recognizes the primacy of constructions and the embodiment of semantics.

References

Aikhenvald, Alexandra. 2000. *Classifiers: A typology of noun categorization devices.* Oxford: Oxford University Press.

Akmajian, Adrian and Adrienne Lehrer. 1976. NP-like quantifiers and the problem of determining the Head of an NP. *Linguistic Analysis* **2(4)**: 395-413.

Bergen, Benjamin K. and Nancy C. Chang. 2002. Embodied Construction Grammar in simulation-based language understanding. Technical report TR-02-004, International Computer Science Institute. ,www.icsi.berkeley.edu.

British National Corpus. <http://www.hcu.ox.ac.uk/BNC/>

Croft, William. 2001. *Radical Construction Grammar.* Oxford: Oxford University Press.

Fillmore, Charles. 1982. Frame Semantics. In *Linguistics in the Morning Calm.* Seoul: Hanshin.

Gawron, Mark. 2002. Two kinds of quantizers in DP. Linguistic Society of America, Annual Meeting, San Francisco, California.

Ginzburg, Jonathan and Ivan Sag. 2000. *Interrogative investigations: the form, meaning, and use of English interrogatives.* Stanford: CSLI Publications.

Goldberg, Adele. 1995. *Constructions: A Construction Grammar approach to Argument Structure.* Chicago: The University of Chicago Press.

Kay, Paul and Charles Fillmore. 1999. The *what's x doing y* construction. *Language* **75(1)**: 1-33.

Langacker, Ronald. 1991. *Concept, Image and Symbol: The cognitive basis of grammar.* Berlin: Mouton de Gruyter.

Lehrer, Andrienne. 1986. English classifier constructions. *Lingua* **68**: 109-148.

Lucy, John A. 1992. *Grammatical Categories and cognition.* Cambridge: Cambridge University Press.

Lucy, John A. 2000. Systems of nominal classification: a concluding discussion. In Gunter Senft (ed.) *Systems of nominal classification.* Cambridge: Cambridge University Press.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartik. 1985. *A comprehensive grammar of the English language.* New York: Longman Inc

Svensson, Patrick. 1998. *Number and countability in English nouns: an embodied model.* Uppsala, Sweden: Swedish Science Press.

1203 Dwinelle Hall
University of California, Berkeley
Berkeley, CA 94720-2650

edodge@socrates.berkeley.edu, acwright@socrates.berkeley.edu

Transparency and Building Lexical Dependency Graphs

CHARLES J. FILLMORE[i] and HIROAKI SATO[ii]
[i]International Computer Science Institute, Berkeley and [ii]Senshu University, Japan

0. Introduction

The authors are participants in a project, called "FrameNet"[1], which is aimed at building a large computer lexicon of contemporary written English and making it accessible through the World Wide Web for both computational and lexicographic interests. In the process of designing this resource, we have to keep in mind how it can serve its intended applications. A prerequisite to most imaginable NLP applications is *word sense disambiguation* (WSD), the automatic process by which a word in a linguistic context can be (probabilistically) assigned its locally intended meaning. This paper will introduce FrameNet and will characterize a facility, based on its tools and data, that could in principle be directed to WSD efforts, and will suggest how both technical (engineering) and linguistic considerations must be called on to build it.

Briefly (a more detailed account can be found below), the manner in which the FrameNet work is carried out, one word at a time, involves the annotation of sentences exemplifying each word taken from a large corpus[2], where each resulting annotation provides an example of the use and the essential combinatorial properties of one *lexical unit*[3] (LU). The overall goal is to offer for each LU a representative collection of annotated examples upon which automatic processes can operate to display a *valence description*[4] of this particular word, in

[1] FrameNet is an NSF-sponsored resource-building effort supported under Grant No. ITR/HCI – 96132: "FrameNet ++": An On-Line Lexical Semantic Resource and Its Application to Speech and Language Technology. PIs are C Fillmore and S Narayanan (ICSI), D Jurafsky (U Colorado), and M Gawron (San Diego State U). Project Manager is C Baker (ICSI). Current funding continues through August 2003. Information at http://framenet/icsi/berkeley/edu/~framenet.

[2] The current FrameNet Corpus is the British National Corpus, which we are using courtesy of Oxford University Press and with software tools made available through the Institut für Maschinelle Sprachverarbeitung of the University of Stuttgart. Information about the BNC can be found at http://info.ox.ac.uk/bnc/

[3] A lexical unit is a lexeme in one of its senses; we follow the usage of D. A. Cruse (1986).

[4] A FrameNet valence description is a display of the "frame elements" (FEs, roughly participant roles) associated with a lexical unit in a given frame, and the manner of their syntactic realization,

terms of the semantic roles, grammatical functions and phrase types of those of the word's phrasal companions that fill slots associated with the word's semantic frame.

Manual annotation of large corpora in the FrameNet manner would be prohibitively expensive, and it is not efficient to let the number of annotations reflect the relative frequency of words, word senses, or valence patterns, and so attempts are being made to automate the process in a way that will allow at least some of the tagging to be done on a large scale using both machine learning and human-written apriori rules. The ability to conduct automatic semantic/syntactic annotation would be enhanced by the study of lexical collocations within the clusters of information that represent the FrameNet annotation of individual LUs (because it could assist in disambiguation of polysemous words), and at the same time the automatic annotation of new documents could help us discover new collocations. The automated ability to find such collocations could make it possible: 1) to recognize collocations for their own sake, accumulating collocation frequencies across a wide variety of corpora[5]; and 2) to detect the specific lexical collocations that occur in a given document, which could be a clue to the topic of the individual passages in the document. The FrameNet database as presently constructed aims at coverage of the distributional varieties, but cannot directly yield information about relative frequencies. If such software can be created, we hope to compile frequency evidence for lexical collocations and/or semantic type selection for each predicating LU.[6]

The collocations will be collected in the form of what can be called *kernel dependency graphs* (KDGs)[7]. These are small packages of information that associate the lexical head (governor) of a set of related dependents, the lexical heads of the constituents that are dependent on that governor, and the frame-specific semantic relations by which the dependent elements are related to the governor. A large corpus-based registry of KDGs would permit queries about, say, the most common dependents of given governors (in given semantic roles), the most common governors of particular nouns, and so on.

---

this expressed in terms of grammatical functions and phrase types if they are present in the sentence, and type of omission (depending on what licenses their absence in the sentence) if they are not realized in the sentence.

[5] The registry of lexical collocations that FrameNet is able to accumulate will be a skimpy but semantically and syntactically richer variety of the Dependency Resource of Dekang Lin (http://www.cs.ualberta.ca/~lindek/demos.htm). The differences are that the collocates will be chosen according to semantic role, rather than merely syntactic relations, and they will differ across different senses of the target word. (Further differences will be described in the present paper.)

[6] We are only minimally able to carry out such studies at the present time; our current efforts are to create the needed conceptual and software infrastructure.

[7] For present purposes the assumed grammatical model is of a Dependency type (Tesnière 1959), but the term "kernel" is borrowed from early transformationalist writing (Chomsky 1957) referring to the structure of a predication with a single verb and its accompanying grammatical partners. The actual characteristics of our KDGs differ from both, precisely because of the treatment of the transparency structures among other things.

The purpose of the present paper is to argue that for such KDGs to be semantically relevant, we need to be able to recognize certain kinds of discrepancies between syntactic and semantic structure, and this involves recognizing the role of 1) *support verbs* in the case of nominal governors—nouns that have their own frame structure—and 2) various classes of *transparent nouns*[8], so that we can spot instances of patterns such as [Noun+of+Noun] in which it is the semantic head (here, the second noun), not the syntactic head (the first noun) that is most related to the context of the NP as a whole in respect to collocations, selection or agreement phenomena.

1. Acquisition of an Inventory of Kernel Dependency Graphs

The objects we intend to accumulate in the KDG project are clusters of lexical items, each serving as a lexical head[9], seen in a quasi-dependency grammar representation. One such cluster could be a simple transitive verb in association with the head nouns of the constituents serving as its subject and its direct object. However, we wish to give more information than can be found in such a minimally structured cluster of words. We also wish to identify 1) the semantic frame within which the governor serves as an LU (the frame that provides the elements of its conceptual structure), 2) markers of the syntactically oblique arguments in the cluster, and 3) the semantic roles (called *frame elements*) of each dependent. The purpose of the present paper is to lay out the grammatical and technical issues in the effort to achieve this.

Each of the clusters of information that we hope to assemble as the work of the KDG project will contain 1) a lexical head, i.e., a predicator of some kind, an LU that evokes a semantic frame with one or more arguments[10], and 2) the lexical heads of the constituents that express the arguments of the head predicator. For example, one such cluster (reduced to the lexical heads), based on the verb *steal*, would be derived, as the display in (2), from sentence (1). (The head word is the governor, the indented words are the lexical heads of the dependents.)

(1) The boy we noticed entering the store behind us was caught stealing a fish.

(2)
```
steal
        boy
        fish
```

From the same sentence, another such cluster would be that seen in (3).

[8] The term is said to be due to Naomi Sager, but we have not found the source.

[9] In general, established compounds will be treated as single lexical heads. Named entities, such as personal names, place names, dates, etc., will be converted to names of types in the collocation registry.

[10] Since the information we seek involves constituents that fill out some of the details of the frame of the governor, we needn't begin by establishing criteria for an argument vs. adjunct distinction.

(3)
```
catch
      SOMEONE
      boy
      stealing
```

The displays in (2) and (3) are incomplete. Versions that include frame identity, and the marking and semantic role information, can be seen in (4) and (5). The verb *steal* belongs to, or evokes, a Theft frame, and the elements of that frame include the *perpetrator* and the *loot*.[11] The diagram[12] identifies the governor, the frame, and those participants of the frame realized in the sentence.

(4)
```
<KDG rdf:ID="1137864">
   <governor>steal</governor>
   <frame rdf:resource="Theft"
    <perpetrator>boy</perpetrator>
    <loot>fish</loot>
   </frame>
</KDG>
```

The governor *caught* in that same sentence evokes what we might call a Spotting frame, which requires an understanding of an observer, the person observed, and the act whose performance has been observed.

(5)
```
<KDG rdf:ID="46823">
   <governor>catch</governor>
   <frame rdf:resource="Spotting2"
    <observer>SOMEONE</observer>
    <observed>boy</observed>
    <act>stealing</act>
   </frame>
</KDG>
```

It can be noticed (1) that *boy* was taken as the "perpetrator" of the *steal* governor, though this is through indirect syntactic evidence (the NP headed by *boy* is the direct object of *caught*, but construed through familiar control structures, as satisfying the subject role of *caught*'s complement VP), and (2) that the agent of the *catch* governor is missing, but understood as an indefinite someone, as part of the interpretation of the passive voice in *was caught*.

Thus we see, a part of the task of deriving KDG clusters from a corpus involves getting around the lexico-syntactic structures that can conceal the "deep" syntactic arrangements within the predication or that can intervene between a governor and its arguments.

---

[11] For present purposes we assign no special significance to the names of frames or the names of the frame elements. It is only important that the frame names are unique and carefully defined, and that the frame element names are defined relative to the frames to which they belong.

[12] Since the FrameNet data will be presented to the world in both XML and DAML+OIL framework, the display here is only to give an idea what the formal representation will look like. The ID numbering and the resource names suggested here are not actually a part of our present database.

Something approaching a complete semantic analysis of each sentence could be built by carrying this through with each predicator in each clause, showing all possible semantic links of predication, modification, coordination, etc. (The "act" in KDG (5) is headed by the *steal* of KDG (4), so a representation of the whole complex, at this level, could be built by binding arguments of one KDG with those of another, or by binding one argument with another entire KDG.) But here we are defining a narrower task, allowing ourselves to ignore certain syntactically higher predicates and find the "deepest" dependency graphs in a clause, in the hope that this would make it possible to zero in on a semantic structure that could reliably indicate something about the topic of the passage. The intuition is that if we are interested in detecting the core idea in a passage, from a sentence like *Authorities have revealed that witnesses claimed that an accountant embezzled an entire month's payroll from Procter and Gamble*, we would be more interested in claims about what somebody has *stolen* than in claims of what some people have *said*.

For verb-headed KDGs, the identification of controlled arguments is fairly straightforward, given famliliar accounts of syntactic control. If FrameNet annotations were accompanied by complete successful parses of the sentences, it would not be necessary for FrameNet annotators to identify those constituents external to the VP which realize FEs of the head-verb's frame. But for the sake of being able to collect collocational and selectional information, i.e., for the sake of being able to answer the questions, which words [collocations] and which classes of words [selectional preferences] occur in particular FE functions with given frames, we seek to label all constituents in the selected sentences which stand for, on the basis of obligatory syntactic relationships, the various FEs of each target verb. This means not only recognizing constituents which are "extracted" (e.g., through topicalization, interrogation, or relativization) but also the arguments of embedding predicates (verbs, nouns, or adjectives) which also serve as arguments the VPs of their complements. For the KDG study, in fact, an assignment that we could justify would have us ignore the control and raising predications themselves, except for the sake of finding KDG arguments, wherever the semantic contribution of the embedding predicates is limited to such matters as Appearance (*seem*), Probability and Ability (*likely*, *may*, *can*, *ability*), Intention (*try*, *want*, *eager*, *willing*, *intention*, *decision*), Report (*claim, say, announcement*), Belief and Expectation (*believe*, *expect*), Phase or Aspect (*begin*, *continue*, *stop*), Obligation (*must*, *ought*), Causation and Inducement (*cause*, *persuade*); in other words, almost all of the familiar control and raising verbs and adjectives, including the modals, plus complement-taking nouns indicating any of the listed categories.

1.1 Nouns as Frame Evokers

One way in which the KDG project goes beyond the simplest notion of kernel structure is that we are equally interested in the frame structure of nouns as well as verb. The simplest case of an argument structure centered in a noun is that of nominalizations of verbs that provide those same structures. The FEs that go with *decision* are the same as those that go with *decide*, as suggested by the examples in (6).

(6) a. we decided that we would storm the palace
a'. the decision that we would storm the palace
b. I decided to confess everything
b'. my decision to confess everything

We see from the examples in (6) that clausal and phrasal direct complements of the noun are essentially those of the source verb. However, it is possible to realize the subject of the corresponding verb, standing for the one who makes the decision, as a postnominal oblique complement of the noun, i.e., with *of* or *by*, as in (7a) and (7b), or as a prehead noun modifier as in (7c).

(7) a. the decision of the committee
b. the decision by the committee
c. the White House decision

A Topic FE can also be recognized with this noun, as postnominal oblique, as in (8a), or as a pronominal noun or adjectival modifier, as in (8b).

(8) a. a decision concerning her retirement
b. a retirement decision

But there is an additional device for identifying a participant in a noun-evoked frame: the *support verb*[13] (SV). Support verbs in the case of event nouns are analogous to the control verbs that provide VP-external arguments for nonfinite verbs, in that one of their arguments is construed as a participant in the situation associated with the head noun in their direct object. In the case of *decision*, the support verb is *make* (in U.S. English) or *take* (in U.K. English). In sentence (9), it is understood that the subject of *make* is the decision-maker in the situation identified by the word *decision.*

(9) I *made* the premature decision that it was time to introduce myself to her father.

Obviously, not every verb that can take a nominalization as its object has this function. Those italicized in (10), for example, do not.

[13] The term appears to have been first used by the late Maurice Gross, but the concept has been around for a long time. An earlier German name was Funktionsverb. Our use of the term is broader than the traditional concept, including a number of verbs that have semantic content of their own.

(10) a. We just have to *accept* your decision.
b. They tried to *force* a decision.

The criterion for being a support verb is that of linguistic necessity the subject of the verb must be understood as a participant in the event designated by the supported noun.

The traditional SVs (especially those called *light verbs*) have little semantic responsibility beyond allowing the frame-structure of the noun to be expressed in a verb-headed clause: in this limited function we find in English mostly the verbs *make*, *have*, *give*, *take*, and *do*. But there are many other SVs that have semantic functions analogous to those of control verbs. Most, or perhaps all, of these, fit various categories of lexical functions elaborated by Igor' Mel'cuk (1996).

For some there is not much more to say than that the subject of the verb is the Actor of the noun's event type: *say a prayer*, *sing a song*, *tell a lie*, plus the light verbs. Different SVs can select different perspectives on events, such as the Actor versus Undergoer perspective seen in *give* vs. *take an examination*, *submit* vs. *receive an application*; *inflict* versus *sustain an injury*; *perform* versus *undergo an operation*. In some cases the SV selects one or another sense of a word: thus *have an argument* selects the 'quarrel' meaning of the noun; *make an argument* selects the 'reasoning' sense. Some SVs differ subtly in register, as in *make a complaint* versus *register*, *lodge* or *enter a complaint*. Some SVs go with nouns in particular semantic classes: names of sins and crimes, as well as the words *sin* and *crime*, seem to prefer *commit*. Some SVs require their subjects to be participants in the larger scenario than the event that is directly associated with the verbal source of a noun: thus *make a promise* is like the verb *promise*, but *keep a promise* and *break a promise* do not themselves refer to the act of promising; similarly, while *give an exam* and *take an exam* are related to the event expressed by the verb *examine*, *pass an exam* and *fail an exam* are not. But they are still bound to arguments of the event associated with the noun.

The KDG project, insofar as it aims to record KDGs built around noun-expressed frames, clearly needs access to information about SVs appropriate to each frame-evoking noun.

1.2 Transparent Nouns

The second departure from the simplest version of our task involves situations in which there is a discrepancy between the syntactic head of a phrase and its semantic core. In many structures of the type [N *of* N], the first noun has some quantifying or typing function and it is the second noun[14] that bears collocational relations with something in the surrounding text. Those instances of first-nouns in this pattern that have this property we will call *transparent nouns*. Their semantic

[14] A tighter formulation would acknowledge the possibility of recursion: in *I could never drink more than a pint of this kind of beer*, we have, with *pint* and *kind*, one [N *of* N] structure inside another, and it is the deepest one that serves as a collocate with *drink*.

functions include those of Quantities, Aggregates, Parts & Portions, Types, Classifiers, and Evaluators. The words that serve this function are not grammatically dedicated to serving in this way, and some of them can stand on their own with *of*-phrases as their complements. A contrast between the two situations can be seen in item (11) where the word *number* can be seen as having either function.

(11) a. We tried to estimate the *number* of apples the horse ate.
b. The horse appears to have eaten a *number* of poisoned apples.

In (11a) an important collocational relation holds among *we*, *estimate* and *number*; in (11b) among *horse*, *eat* and *apples*, the latter being transparent to the presence of the word *number*. In (11b) the word *number* is paired with *lot*, *bunch*, etc.; in (11a) its partners include *quantity*, *size*, *magnitude*, etc.

The intended idea of transparency can be seen by considering how one might derive a semantically relevant KDG from a sentence like (12):

(12) A majority of tobacco producers use a form of asbestos in this kind of filter.

If we took the lexical heads of each of the NP constituents in this sentence, the KDG, using the words underlined in (13), would be (14)[15] (majority uses form in kind).

(13) A majority of tobacco producers **use** a form of asbestos in this kind of filter.

(14)
```
<KDG rdf:ID="11786424">
 <governor>use</governor>
 <frame rdf:resource="use3">
   <agent>majority</agent>
   <ingredient>form</ingredient>
   <product>in: kind</product>
 </frame>
</KDG>
```

By skipping the transparent nouns of quantity and type, choosing instead the nouns underlined in (15), we derive KDG (16), which we take consider more informative as to the point of the passage (tobacco-producers use asbestos in filter).

(15) A majority of tobacco producers **use** a form of asbestos in this kind of filter.

[15] We take the frame associated with *use* in this sentence to involve an agent who consuming ingredients (resources, material) in creating from them some sort of product.

(16)
```
<KDG rdf:ID="11786424">
 <governor>use</governor >
 <frame rdf:resource="use3">
   <agent>tobacco producers</agent >
   <ingredient>asbestos</ingredient>
   <product>in: filter</product>
 </frame>
</KDG>
```

Some examples of collocational relations obtaining through various kinds of transparent nouns are given in Table 1.

Table 1

Prep & favored object	on the shelf in the closet	on part of the shelf in part of the closet
Nouns & required possessors	her husband my wife	her idiot of a husband my gem of a wife
Subjects & favored VPs	water would quench my thirst	a bit of water would quench my thirst
Verb & favored object	hard to iron this shirt	hard to iron this kind of shirt
Adjective & favored complement	guilty of murder	guilty of three counts of murder
Noun & favored support verb	made a good impression	made that kind of impression

Notice that the sentence in (17), involving a noun derived from the verb *use*, would yield essentially the same structure as the KDG in (16), although this time the governor is a noun, and the agent is connected with the frame through a support verb those subject is controlled through an embedding predicate of saying.

(17) A number of tobacco producers are said to have made use of some form of asbestos in this kind of filter.

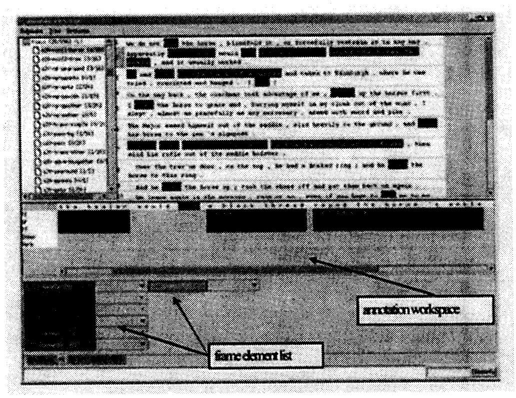
2. Introducing FrameNet

There are several aspects of the FrameNet data which are capable of serving the purposes of KDG derivation discussed in this paper.

The FrameNet project is producing a dictionary *cum* lexicon, for both human and computational uses, that documents the combinatorial properties of English lexical items, in semantic and syntactic terms, based on attestations in a very large corpus. Basically, the work of FrameNet is to discover and describe those semantic frames which provide the conceptual background for groups of semantically related words; to analyze the situational props, participants and components that conceptually make up part of the frame, to give them names (as "frame elements", FEs); to come up with a list of words—verbs, nouns, adjectives—whose meanings evoke this frame; to explore attestations of some of these words in the Corpus to get a preliminary idea of the ways in which the various FEs combine with the word and how they are syntactically realized; to create a concordance of instances of each word in a way that is sure to include all varieties of the syntactic contexts in which the word occurs (in the intended meaning); and to submit these sentences to a software tool that can be used for selecting sentences that typify uses of the word (in the intended sense) and labeling the constituents that express or identify the FEs.[16]

Figure 1 gives a glimpse of the annotation tool used in the project. Shown is a set of sentences with the verb *tie* as analyzed in an Attachment frame. (The part of the sentence visible in the workspace is *the healer would tie a black thread round the horse's ankle.*)

Figure 1



The search for the support verbs that accompany the noun *complaint*, and the syntactic roles of its FEs, yields information like what is seen in Figure 2. (The inclusion of *afterwards* as a support verb is, of course, a mistake.) We can notice that the frame structure of this noun has separated Speaker as the person submitting the complaint and Addressee as the person toward whom the complaint is addressed, and the verbs *face* and *get* can have Addressee subjects, while *express* and *lodge* and have Speaker as subject. (The verb *have* seems to allow both possibilities. *I have a complaint* may mean that I wish to express a complaint; *we have had numerous complaints about our prices* refers to complaints we have received.)

[16] In practice the process zigs and zags. Proposals are made for the structure of a new frame, and wordlists are compiled. Examination of concordance lines for some of the proposed words may require us to revise the frame description, or to add FEs, or to consider whether the initial list should be divided into more than one frame. After such decisions are made, some words get added to the lists and others get set aside.

Figure 2

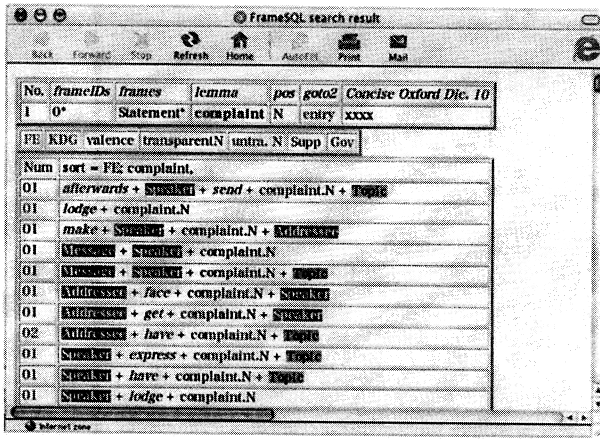
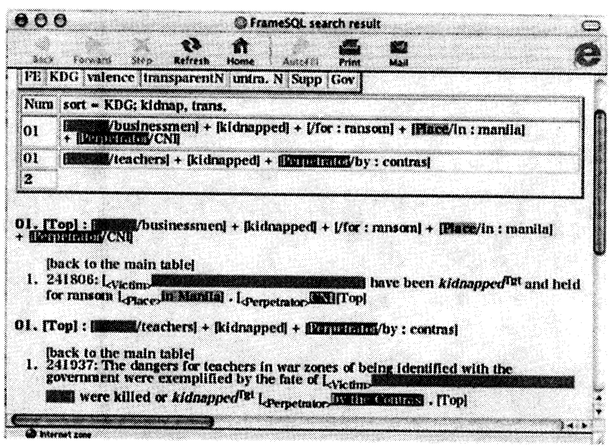


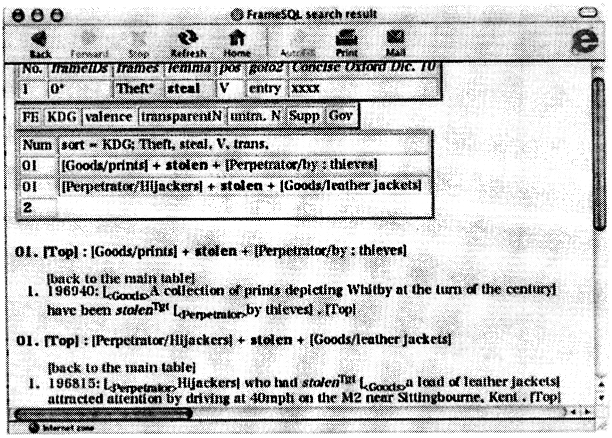
Figure 3 shows two KDGs found in attestations of the verb *kidnap* displayed with the actual sentences on which they are based, and demonstrates that the quantity expression *number* was successfully ignored.

Figure 3



Similarly, in Figure 4 the system has correctly identified *prints*, through the transparent noun *collection*, and *jackets* through *loads*.

Figure 4



3. Summary

The database-viewing tools demonstrated so far are not yet able to generate a full registry of KDGs from the FrameNet annotations, and are certainly not capable of extracting KDGs from raw text. But they do show that if a parser can recognize not only the familiar embedding structures for filling out the missing components of nonfinite VPs but also support verbs in our sense (which can be seen as having a similar function) and transparent nouns, the data needed for constructing a KDG registry are at hand.

Acknowledgements

The authors are indebted to Srini Narayanan and Collin Baker for help at various points in the preparation of this paper.

References

Akimoto, Minoji. 1989. *A Study of Verbo-Nominal Structures in English.* Shinozaki Shorin. Tokyo.

Chomsky, A. Noam. 1957. *Syntactic Structures*. The Hague: Mouton.

Cruse, D. Alan. 1986. *Lexical Semantics.* Cambridge, U.K.: Cambridge University Press.

Davis, Anthony R. & Leslie Barrett. 2001. Support verbs, nominalizations and participant roles. In Pierrette Bouillon & Kyoko Kanzaki (eds.) *Proceedings of the First International Workshop on Generative Approaches to the Lexicon.* Geneva.

Fontenelle, Thierry. 1999. Semantic resources for word sense disambiguation: a sine qua non?. *Linguistica e Filologia*, **9**: 25-41.

Mel'cuk, Igor'. 1996. Lexical functions: a tool for the description of lexical relations in the lexicon. In Leo Wanner (ed.) *Lexical Functions in Lexicography and Natural Language Processing*, 37-102. Amsterdam/Philadelphia: John Benjamins.

Napoli, Donna Jo. 1989. *Predication Theory: A Case Study for Indexing Theory.* Cambridge, U.K.: Cambridge University Press.

Svensson, Patrik. 1989. *Number and Countability in English Nouns*. Uppsala: Swedish Science Press.

Tesnière, Lucien. 1959. *Syntaxe Structurale.* Paris: Klincksieck.

Charles J. Fillmore
International Computer Science Institute
1947 Center St., Suite 600
Berkeley, California 94704

fillmore@icsi.berkeley.edu

Hiroaki Sato
Senshu University, Japan

hiroaki@isc.senshu-u.ac.jp

Gender Stereotype in Prosody: Japanese Interactional Particles *Ne* and *Yo*[1]

MIE HIRAMOTO-SANDERS
University of Hawai'i at Mānoa

0. Introduction

In this paper, I will focus on how prosody affects the interpretation of gender-neutral interactional particles *ne* and *yo* in Japanese. The interactional particles are potential sites for stereotyping, and "males have flatter intonation patterns than females do in pitch range and pitch heights" (McConnell-Ginet 1983: 73). Based on these assumptions, a purpose of the paper is to find possible answers for the following research questions: 1) Will gender stereotype characteristics appear on some of the most frequently used particles such as *ne* or *yo*? and 2) Will female-style speech, in general, will be more emphasized than male-style speech?

Interactional particles are discourse markers, and by their nature "loosely placed at various points of utterances in conversation" (Reynolds 2000:88). Previous research tried to explain the different functions of the interactional particles in a categorical way. However, there are limitations to this kind of analysis because such particles are emphatic and vague in meaning when used as discourse markers. In other words, it is very difficult to say that *ne* marks an affirmation or *yo* marks a suggestion the way one can say that *ga* marks nominal case or *o* marks accusative case. For example, the following sentence can be interpreted in different ways according to the prosodic value on the interactional particles.

(1)	*Ashita*	*ginkoo*	*ni*	*itte*	*ne.*
	tomorrow	bank	to	go	IP[2]

'(I) will go to the bank tomorrow, and . . . .' (↗)
'(Please) go to the bank tomorrow.' (↘)

---

[1] I would like to express my sincere gratitude to Benjamin Bergen, Ryoko Hattori, Vincent Kiste, Tomoko Kozasa, Hsiu-Chuan Liao and everyone at UHM who gave me helpful suggestions and comments or participated in this study.

[2] Abbreviations used in this paper are: IP = Interactional Particle, COP = Copula, INT = Interrogative marker, GEN = Genitive marker, NEG = Negative marker

When interpreting the meanings of Japanese interactional particles, prosodic features (e.g., intonation, pitch, duration) cannot be ignored, because prosodic features change their meanings.

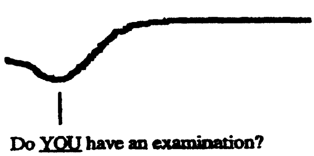
1. Intonational Characteristics of Japanese

Ladd (1996:6) defines intonation as "the use of *suprasegmental* phonetic features to convey 'postlexical' or *sentence-level* pragmatic meanings in a *linguistically structured way*." Japanese is a pitch language, and it is important to clearly separate the notions of pitch and stress from intonation. For a pitch language like Japanese, the following minimal pairs are possible.

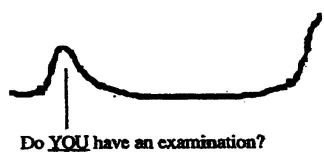
(2) *kame* 'vase' / *kame* 'turtle' (3) *hashi* 'bridge' / *hashi* 'chopsticks'

Intonation, but not a pitch accent, is relevant to prosody in Japanese interactional particles. In Japanese, the intonation patterns are most likely to change at phrasal boundaries (such as before a long pause or at a sentence-final position). This point is important since Japanese interactional particles are most likely to occur at phrasal-ending positions. Ueyama and Jun (1996:110) compared how native English speakers and Japanese ESL speakers produce the rising intonation patterns in an English interrogative sentence, and found that the English speakers raise the intonation gradually (as in (4)), whereas the Japanese ESL speakers raise their intonation abruptly at the end of the sentence (as in (5)).

(4) English native speakers



(5) Japanese speakers



This suggests that the intonational features are condensed at the phrasal boundaries in Japanese, whereas they are stretched through longer suprasegmental material in English. The intonation pattern in (5) shows a good example of a prosodic contour being placed at the phrasal boundaries by the Japanese ESL speakers. The same is true for Japanese speakers speaking Japanese (Eda 2000), as seen in the following English and Japanese sentence patterns.

(6) It's cold, isn't it?

(7) *Samui desu ne.*
cold COP IP
It's cold, isn't it?

(8) It's cold, you know.

(9) *Samui desu yo.*
cold COP IP
It's cold, you know.

2. Gender and Prosody

Interactional particles (except *ne*) are often regarded as discourse markers that typically occur at sentence-final positions and signal speakers' attitude toward the content of the utterance or toward the addressees. Some interactional particles are used exclusively by speakers of a particular gender, while others are used as gender-neutral forms. Although Shibamoto (1985:176) suggests that *ne* is a feminine particle, others consider it to be gender-neutral (McGloin 1990, Eda 2000, Reynolds 2000). This paper follows the latter assumption on the treatment of *ne*. As for the other interactional particle *yo*, it is widely recognized as gender-neutral.

Some interactional particles are gender-exclusive (i.e., linguistic forms that are exclusively used by one gender) in Japanese. For example, '*ze*' and '*zo*' are gender-exclusive interactional particles for male speakers in Japanese, regardless of prosodic features, as in (10).

(10)	*Sandoicchi*	*taberu*	*ze/zo.*	
	Sandwich	eat	IP	
	'(I will) eat a sandwich.'		(↗ or ↘)	(masculine, *feminine)

'*Wa*' is often defined as a feminine interactional particle, but specifying gender depends on prosody, as in (11).

(11)	*Sandoicchi*	*taberu*	*wa.*	
	Sandwich	eat	IP	
	'(I will) eat a sandwich.'		(↗)	(feminine, *masculine)
	'(I will) eat a sandwich.'		(↘)	(masculine or dialectal[3], *feminine)

Thus, prosody is important when interpreting gender in Japanese interactional particles. There are different interpretations of sentence-final intonation patterns that are independent from each sentence's inherited grammatical or lexical features. (For English intonation patterns, see Pierrehumbert & Hirschberg (1990).)

"Variation" of intonation, as defined by McConnell-Ginet (1983:72) is "an alternative way of uttering the 'same' linguistic unit." This can be demonstrated by the following pairs of examples.

(12) Female speakers

(13) Male Speakers

[3] Kansai Dialect spoken in Western Japan.

As shown in (12) and (13), although both males and females utter this word in the same rising intonation pattern, 'hello' of females is, in general, higher in pitch and longer in duration than 'hello' of males. These characteristics are probably due to a general rule that female voices have higher pitch as well as wider pitch range than those of males. In other words, males have flatter and lower intonation patterns than females do (McConnell-Ginet 1983:73).

Japanese interactional particles function as discourse markers in a form of "audible gestures" (Bolinger & Sear 1981:110) to convey paralinguistic meanings. That is, Japanese interactional particles are potential stereotyping sites, including gender stereotyping.

3. Methodology

To examine gender-stereotyping assumptions among Japanese speakers, sets of gender-neutral sentences (sentences that could be naturally articulated by both male and female speakers) that contain *ne* and *yo* were formulated. Three male and three female native speakers of Japanese in their 20's and 30's participated in a production test. To avoid dialectal variation in the use of particles and intonation, all of the informants[4] were native speakers of Tokyo Standard Japanese. PitchWorks program was used to digitize and analyze the data. The data collection took place in a recording studio at the University of Hawai'i in April 2001.

The informants were asked to read the sentences in three different ways: first, naturally (natural speech, or NS); second, male-like (masculine speech, or MS); and lastly, female-like (feminine speech, or FS). For the MS and FS parts, the informants were told to pretend that they were going through auditions for a theater play, acting male and female roles using the same lines. The same six sentences listed below were used for all three speech styles (NS, MS, and FS), thus, a total of 18 sentences were collected from each informant. The sentences used for the production test are:

(14) Sentence-external *ne*
Ne, bideo demo miru?
IP, video how about watch
'Say, do you want to watch a video?'

(15) Sentence-medial *ne*
Anta[5] *ne, dooshite itsumo okurete kuru no?*
you IP why always late come INT
'You, why do you show up late all the time?'

[4] The informants' names used in this paper are pseudonymous.
[5] *Anta* is a casual form of *anata* 'you' in Japanese. It is frequently used in informal conversations.

(16) Sentence-final *ne* (channeling)
Ashita no dyinaa tanoshimi da ne.
tomorrow GEN dinner exciting COP IP
'Aren't you excited about tomorrow's dinner?'

(17) Sentence-final *ne* (back-channeling)
Yokatta, butsuri-no tesuto heikin-ten wa atta yo.
good physics-GEN test average-score at least there is IP
'Thank goodness, I made average on my physics test.'
Sorenara anshin da ne.
in that case no worries COP IP
'(You) do not have to worry about it anymore, then.'

(18) Sentence-final *yo* (channeling)
Nee nee, sore-ja nai yo.
say say that-COP NEG IP
'Say, it's not that one.'

(19) Sentence-final *yo* (back-channeling)
Kono jikan Jinbo-san moo neteru kana?
this time Jinbo-san already sleeping INT
'Do you think Jinbo-san is sleeping already at this time?'
Un, moo juuji da yo.
yeah already ten o'clock COP IP
'Yeah, it's ten o'clock already.'

The criteria that I used for selecting the utterances for the data are 1) a total of four possible positions that *ne* may occur (one sentence-external, one sentence-medial, and two sentence-final positions of channeling and back-channeling functions), and 2) a total of two possible positions that *yo* may occur (two sentence final positions of channeling and back-channeling functions).

4. Results

For this study, the results of the production test were divided into different categories according to their tonal structures, which are rising (➚), falling (➘), rising with a scoop (➘➚), and falling with a scoop (➚➘). I analyzed the male group and female group data separately; first, let us discuss the results from the male group.

All of the male speakers shared the same category (falling with a scoop ➚➘) in all three speech patterns (NS, MS, and FS) for the sentence in (15). In the tables below, 'Initial Pitch' means the beginning pitch of the interactional particle, "Rising Pitch Range (RPR)' and "Falling Pitch Range (FPR)' indicate how much the pitch went up and down based on the peak of the contour tones, and 'Ending

Pitch' indicates when the interactional particle ended in the utterance. 'Duration' refers to the duration of the particle.

(20) Falling with a Scoop (↗↘) in NS of the Male Group in Sentence (15)

	Yuuji	Takuya	Norio	Average
Initial Pitch	139	99	118	118.7
RPR	12	6	6	8
FPR	31	11	31	24.3
Ending Pitch	120	92	93	101.3
Duration	461.1	303.3	266.9	343.7

(21) Falling with a Scoop (↗↘) in MS of the Male Group in Sentence (15)

	Yuuji	Takuya	Norio	Average
Initial Pitch	117	96	92	101.7
RPR	30	14	21	21.7
FPR	26	24	35	28.3
Ending Pitch	121	86	78	95
Duration	366.8	247.9	337	317.2

(22) Falling with a Scoop (↗↘) in FS of the Male Group in Sentence (15)

	Yuuji	Takuya	Norio	Average
Initial Pitch	123	117	98	112.7
RPR	52	13	10	31
FPR	49	20	24	31
Ending Pitch	126	110	84	106.7
Duration	708.9	261.2	398	456

I compared the durations of the interactional particles. The overall average of the durations in FS were longer than MS or NS, as shown in (20)-(22) as well as in (23).

(23) Male Group's Duration of *Ne* and *Yo* (ms.)

		NS	MS	FS	Average
Ne	External	201.7	223.6	227.6	**218.3**
	Medial	343.8	317.2	456.0	**372.3**
	Final C[6]	232.8	230.9	272.1	**245.3**
	Final BC[7]	199.2	194.8	356.9	**250.3**
Average		**244.3**	**242.1**	**328.2**	**271.6**
Yo	Final C	176.6	197.6	240.3	**204.8**
	Final BC	127.8	130.2	208.5	**155.5**
Average		**152.2**	**163.9**	**224.4**	**180.2**
Total Average		**198.3**	**203.0**	**267.3**	**225.9**

This means that the male informants were extending the duration (except for the external use of *ne*) when articulating FS to add feminine features in their speech.

[6] C =Channeling
[7] BC =Back-Channeling

On the other hand, MS did not differ much from NS. For both *ne* and *yo*, FS carried the longest duration, while NS and MS did not show much difference in their duration. Now, let us consider the distributions of the intonation patterns.

(24) Male Group's Distributions of the Intonation Patterns

		NS				MS				FS			
		↗	↘	↗↘	↘↗	↗	↘	↗↘	↘↗	↗	↘	↗↘	↘↗
Ne	External		3				3				2	1	
	Medial			**3**				**3**				**3**	
	Final C	2		1		1	1	1		3			
	Final BC	3				3				2		1	
Yo	Final C	2	1			1	2			1	1		1
	Final BC	3				3				1	1		1
	Total	10	4	4	0	8	6	4	0	7	4	5	2

As shown in 24, among the four types of intonation, the contour tones were less frequently used than the non-contour tones by both male and female speakers, and also in all three speech styles. However, the male informants increased the use of contour tones for FS more than they did in NS and MS. This is partly due to the extension of duration in articulation (the longer the duration is, the more time the speakers have to manipulate contour tones on the interactional particles). Nevertheless, from the data, it is clear that the speakers were using contour tones to emphasize their femininity in FS, while the differences between NS and MS were not so obvious. The boldface numbers in the table indicate that the three male in formants uttered sentence (15) in the same intonation pattern. (PitchWorks images are available in the Appendix, Figures (1)-(3).)

Now, let us look at the results from the female group. Generally speaking, prosodic characteristics of the female speakers, compared to the males, are higher pitch levels, wider pitch range, and longer durations. Naturally, the female informants' NS were articulated with higher pitch levels (about 100 Hz higher on average) than the male group. The female informants shared the same falling (↘) tone of *yo* in the sentence (18).

(25) Falling (↘) in NS of the Female Group in Sentence (18)

	Rika	Michiko	Sayuri	Average
Initial Pitch	174	246	248	222.7
Ending Pitch	149	177	151	159
Pitch Range	25	29	97	50.3
Duration	197.7	223.4	438.3	286.5

(26) Falling (↘) in MS of the Female Group in Sentence (18)

	Rika	Michiko	Sayuri	Average
Initial Pitch	165	200	190	185
Ending Pitch	145	167	149	153.7
Pitch Range	20	33	41	31.3
Duration	150.2	250.3	185.9	195.5

(27) Falling (↘) in FS of the Female Group in Sentence (18)

	Rika	Michiko	Sayuri	Average
Initial Pitch	174	211	255	213.3
Ending Pitch	157	117	153	142.3
Pitch Range	13	94	102	69.7
Duration	278.3	296.8	491.6	355.6

As shown in (25)-(27), the informants' overall pitches are lower, pitch ranges are narrower, and durations are shorter in MS, whereas the opposite prosodic characteristics were evident in FS. The same results were observed in the duration of other sentences as seen in (28).

(28) Female Group's Duration of *Ne* and *Yo* (ms.)

		NS	MS	FS	Average
Ne	External	225.3	249.8	383.2	**286.1**
	Medial	281.6	254.2	302.0	**279.3**
	Final C	271.2	310.5	318.6	**300.1**
	Final BC	207.5	239.3	455.3	**300.7**
Average		**246.4**	**263.5**	**364.8**	**291.6**
Yo	Final C	286.5	195.5	355.6	**279.2**
	Final BC	160.4	100.4	351.6	**204.1**
Average		**223.5**	**148.0**	**353.6**	**241.7**
Total Average		**234.9**	**205.7**	**359.2**	**266.6**

Again, the longest duration is seen in FS. The duration of *yo* is especially short in MS compared to that of NS in this group. This suggests, first, that the female informants were trying to articulate *yo* in MS shorter than NS (it is not caused by some idiosyncratic characteristics of one speaker in the data). Second, the female informants also lengthened their duration in FS to add femininity in their speech. The average duration of *yo* was shorter than *ne* among the male speakers by about 100 ms. For the female speakers, the difference in the duration is only evident in MS. It suggests that the female informants were trying to articulate *yo* in MS shorter to add masculine feature. Table (29) below shows the overall occurrences of the intonation patterns in the female group. The boldface numbers in the table indicate that the three female informants produced sentence (18) in the same intonation pattern (↘), as shown in (25)-(27). (PitchWorks images are available in the Appendix, Figures (4)-(6).)

(29) Female Group's Distributions of the Intonation Patterns

		NS				MS				FS			
		↗	↘	↗↘	↘↗	↗	↘	↗↘	↘↗	↗	↘	↗↘	↘↗
Ne	External		3			1	2				1	1	1
	Medial		1	2			3			3			
	Final C		2	1			2	1			2	1	
	Final BC	2	1				3				2	1	
Yo	Final C		**3**				**3**				**3**		
	Final BC	3				1	2			3			
	Total	5	10	3	0	2	15	1	0	6	8	3	1

In the female group, the speakers used more falling tones in MS than in other speech styles. The falling tone in Japanese has a more assertive connotation than does the rising tone. It gives a less vague and firmer feel to utterances. The female informants used the falling tones in MS to increase masculinity in their speech when asked to add gender-stereotypical features in their speech. At the same time, these informants made use of 'less assertive' rising tones more frequently in FS.

5. Discussion

According to the results of the production tests from the male group, gender stereotyping is more evident in FS than MS. Male speakers lengthened the interactional particles and used more contour tones in FS, whereas their NS and MS did not show many differences in the use of prosody. For the female speakers, MS was characterized by shorter duration and falling tones. They showed longer durations and more contour tones when adding stereotyping features in FS than the male informants did. The average pitch of NS in male informants was about 150Hz, while it was about 250Hz in the female group. As for FS, both groups exaggerated the femininity in unnaturally high pitch intonation, as high as 400 to 500Hz for some samples. These results suggest that the prosodic features expressed in the interactional particles demonstrate gender-stereotyping.

In addition to the anatomical reasons for the prosodic differences between the male and female speakers, there seem to be language-specific reasons that contribute to the stereotyping of FS as it appeared in the data. In Japanese, male and female speakers are culturally expected to show differences in their use of prosody. Endo (1995:37) states that "women are apt to adopt a distinctive tone of voice, carriage, and behavior" in addition to "a high-pitched voice." Furthermore, Japanese women have been traditionally disciplined to speak with those vocal manners from the *Heian* period (9^{th} to 12^{th} centuries) (Endo 1995:37), and Japanese male audience still expect women to speak in "a soft, gentle, nonthreatening tone" (Endo 1995:39). In addition, Loveday (1986) and Ohara (1994) point out that the socio-cultural factors in the Japanese language encourage women's use of higher pitches and longer duration since they are often connected with politeness to a certain extent.

6. Conclusion

In this study, stereotypes in Japanese interactional particles *ne* and *yo* were examined. My research questions were, 1) will gender stereotype characteristics appear on some of the most frequently used particles such as *ne* or *yo*, and 2) will female-style speech uttered by speakers, in general, be more emphasized than male-style speech? A production test was conducted to find answers for the questions, and the results suggest that the prosodic features expressed in the interactional particles *ne* and *yo* do show gender-stereotyping.

From a socio-cultural perspective, historically speaking, Japanese language seems to favor less-assertive manners in the use of prosody in females' speech. This cultural factor highlights some of the stereotypical features in the results of the production test. As strongly suggested by Eda (2000), in order to understand how these particles function in the language, one must pay attention to the use of prosody. She describes "the most serious deficit of the works mentioned so far" as a lack of understanding in "apparent variation in meaning under variation in sentence final intonation" that relates to interactional particles as well as intonation patterns (Eda 2000:170-171). This study is based on a limited amount of data; therefore, more studies on the relations between interactional particles and prosody are needed in the future.

References

Bolinger, Dwight Le Merton, & Donald A. Sear. 1981. *Aspects of language*. New York: Harcourt Brace Jovanovich.

Eda, Sanae. 2000. Sentence final particles *ne* and *yo*: An interface between prosody and pragmatics. In M. Nakayama & C. Quinn, Jr. (eds.) *Japanese/Korean Linguistics* 9: 167-180. Stanford: CSLI.

Endo, Orie. 1995. Aspects of sexism in language. In K. Fujimura-Fanselow & A. Kameda (eds.) *Japanese woman: New feminist perspectives on the past, present, and future,* 29-42. New York: The Feminist Press.

Ladd, D. Robert. 1996. *Intonational phonology*. Cambridge: Cambridge University Press.

Loveday, Leo. 1986. *Explorations in Japanese sociolinguistics*. Amsterdam and Philadelphia: John Benjamins Publishing Company.

McConnell-Ginet, Sally. 1983. Intonation in a man's world. In B. Thorne, C. Kramarae, & N. Henley (eds.) *Language, gender, and society*. Rowley: Newbury House Publishers.

McGloin, Naomi Hanaoka. 1990. Sex and sentence-final particles. *Aspect of Japanese Women's Language*, 23-41. Tokyo: Kuroshio Shuppan.

Ohara, Yumiko. 1994. Koe no takasa kara ukeru inshoo ni tsuite. *Kotoba* **14**: 14-19.

Pierrehumbert, Janet, & Julia Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In P. Cohen, J. Morgan, & M. Pollack (eds.) *Intentions in communication*, 271-311. Cambridge: MIT Press.

Shibamoto, Janet S. 1985. Japanese women's language: As spoken by women, as spoken by men. In S. Bremner, N. Caskey, & B. Moonwomon (eds.) *Proceedings of the First Berkeley Women and Language Conference*, 171-182. Berkeley: Berkeley Women and Language Group.

Reynolds, Katsue Akiba. 2000. A sociohistorical perspectives: Implications of a quantitative analysis of Japanese interactional particles. *Kotoba* **21**: 88-103.

Ueyama, Motoko, & Sun-Ah Jun. 1996. Focus realization of Japanese English and Korean English intonation. In I. Maddieson (ed.) *UCLA Working Papers in Phonetics* **94** (December): 110-125.

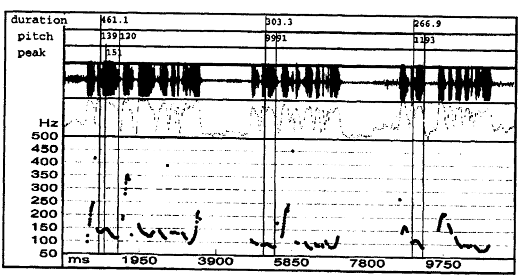
Mie Hiramoto Sanders
569 Moore Hall
University of Hawai'i at Mānoa
1890 East-West Road, Honolulu, HI 96822

mies@hawaii.edu

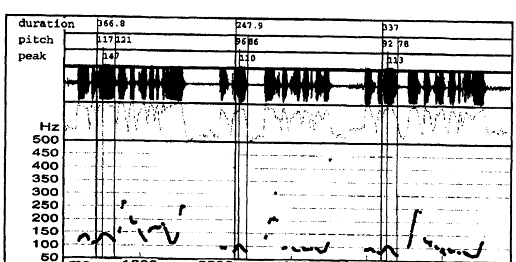
Appendix

PitchWorks Figures of the sentence (15) and (18).

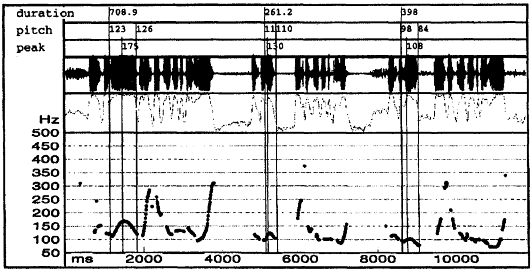
(1) Falling with a Scoop (↗↘) in NS of the Male Group



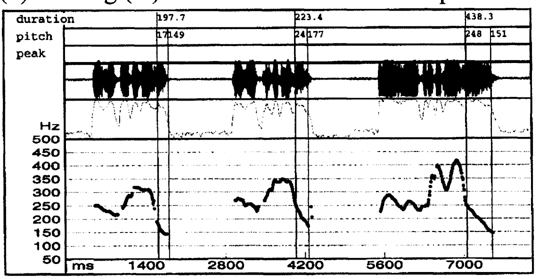
(2) Falling with a Scoop (↗↘) in MS of the Male Group



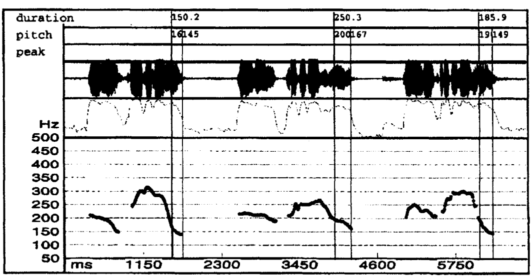
(3) Falling with a Scoop (➚➘) in FS of the Male Group



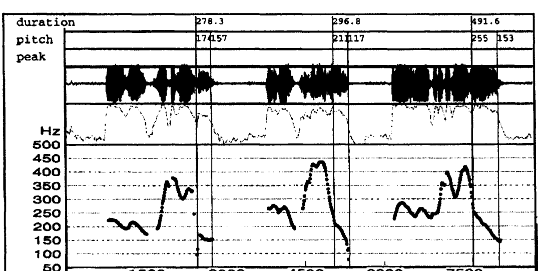
(4) Falling (➘) in NS of the Female Group



(5) Falling (➘) in MS of the Female Group



(6) Falling (➘) in FS of the Female Group



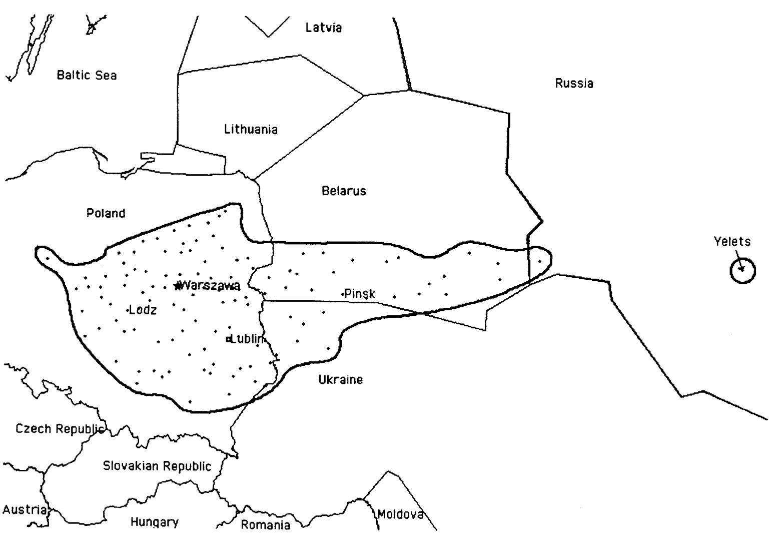
Variationist Dialectology: Chain Shifts and Mergers in Yiddish*

URI HORESH
University of Pennsylvania

0. Introduction

The purpose of this paper is to illustrate the use of dialectological methodology to explain language change and account for variation. The data for this study have been extracted from *The Language and Culture Atlas of Ashkenazic Jewry* (*LCAAJ*).

Map I: The Polish Yiddish dialect area



While the *LCAAJ* encompasses the entire pre-WWII Yiddish-speaking world, this current study focuses on the area marked "P" – for (Northern) Poland. As

* Thanks are due to Hannah Chervitz for collaborative work at some early and crucial stages of conducting this research. I also take pleasure in expressing gratitude to Sharon Ash, Sarah Bunin Benor, Daniel Ezra Johnson, Ronald I. Kim and William Labov for their insightful remarks and fruitful discussions. All errors are, as usual, my own.

easily visible in Map I[2], this area does not coincide with the contemporary political boundaries of the Republic of Poland. Rather, it is the area considered by the editors of the atlas to represent the Polish dialect area of Yiddish. It is also the region in which most of the interviews that led to the compilation of this part of the atlas were conducted according to the Polish Regional Abridgement of the Standardized Master Questionnaire (SMQ). This area includes parts of today's Poland, Belarus, the Ukraine, Moldova and Russia. For the sake of this study, the cluster of dialects in this region will be referred to as Polish Yiddish (PY).

1. Chain Shifts in PY: Overview

In our initial attempt to comprehend some of the data presented in the *LCAAJ* and adapt it into a concise, coherent statement about processes in the language, we used primarily the maps in Volume I ("Historical and Theoretical Foundations"). In this volume, unlike many other linguistic atlases, the data are presented in the form of "typical developments." Rather than providing the reader with proto-words (e.g., in some standardized form of the language, and their local reflexes, this atlas presents the general pattern typical of a particular phoneme of Proto-Yiddish, providing several examples for each such phoneme.

The *LCAAJ* uses the following shorthand for reference to the vowels of Proto-Yiddish (motivated by the phonemes assumed for Middle High German (MHG), but encompassing words of all etymological origins): a capital letter represents the general "color" of the vowel; it is followed by a subscript numeral, usually from 1 to 4, but in one case from 1 to 5, which provides more information regarding the vowel. Subscript 1 refers to short monophthongs; subscript 2 and 3 refer to long monophthongs, and subscript 4 (and 5) refer to diphthongs.

Labov (1994:118) defines a minimal chain shift as "a change in the positions of two phonemes in which one moves away from an original position that is then occupied by the other." Upon examining the maps in Volume I depicting sound changes in the vocalic inventory, we can determine that several shifts have occurred in PY. They are presented here assuming three subsystems:

Short monophthongs:
(1) u > i

Long monophthongs:
Within the subsystem:
(2) u :> i:
(3) a: > u:

Across subsystems:
(4) o: > oy
(5) e: > ay

Diphthongs:
Within the subsystem:
(6) ey > ay

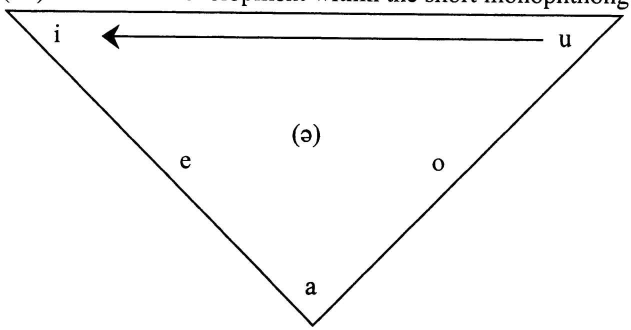
[2] In this paper I will refer to maps included within the paper itself using Roman numerals. Maps from the *LCAAJ* will be referred to by their original numbers, using Arab numerals.

(7) au > ou ~ oy (but note (8) below)

Across subsystems:
(8) au > o: (but note (7) above)
(9) ay > a:

The short monophthongs are thus fairly stable, with only one shift to report. This seems, however, to be a significant shift, as it is in fact a merger of /u/ and /i/.

(10) Historical development within the short monophthong subsystem

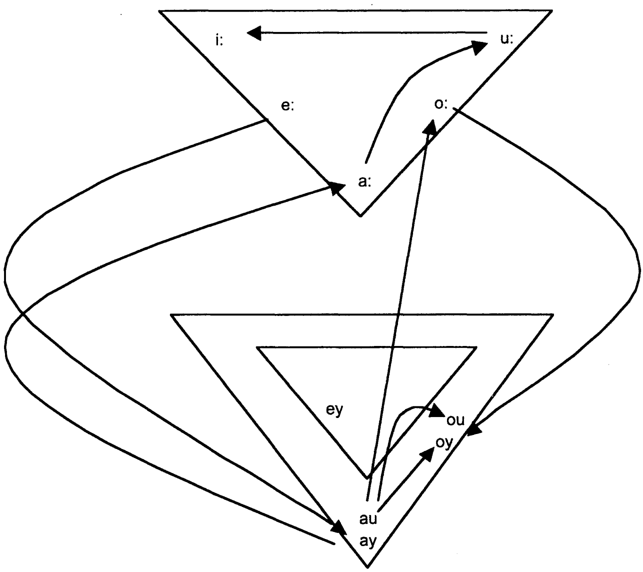


A similar shift occurred with the long counterpart, resulting again in fronting and ultimately unrounding. Once the /u/ space became vacant, /a/ could take its place. These two changes in the long monophthongal subsystem constitute a minimal chain shift as defined in Labov 1994. The shifts described thus far correspond to Principle I in Labov (1994:116) (*long vowels rise*: a: > u:) and Principle III (*back vowels move to the front*: u[:] > i[:]).

An interesting interchange of phonemes across subsystems is also worth noting. The long monophthong subsystem has gotten rid of /o:/ for the sake of the diphthong /oy/. However, one of the reflexes of original /au/ is the long monophthong /o:/. Note, however, that /au/ shows some degree of variation, and may also be realized as either [ou] or [oy], keeping the shift within its original subsystem.

Another instance of chain shifting involves both intra-subsystem and inter-subsystem changes. In accordance with Principle IIa in Labov (1994:116), the nucleus of the diphthong /ey/ falls to /ay/. Historic /ay/ shifts into the long monophthong subsystem taking the space vacated when /a/ shifted upward toward /u/. Figure (11) represents the phonological space for long monophthongs and for diphthongs, marking the shifts within each subsystem and across the two subsystems.

(11) Historical development within and across bimoraic subsystems

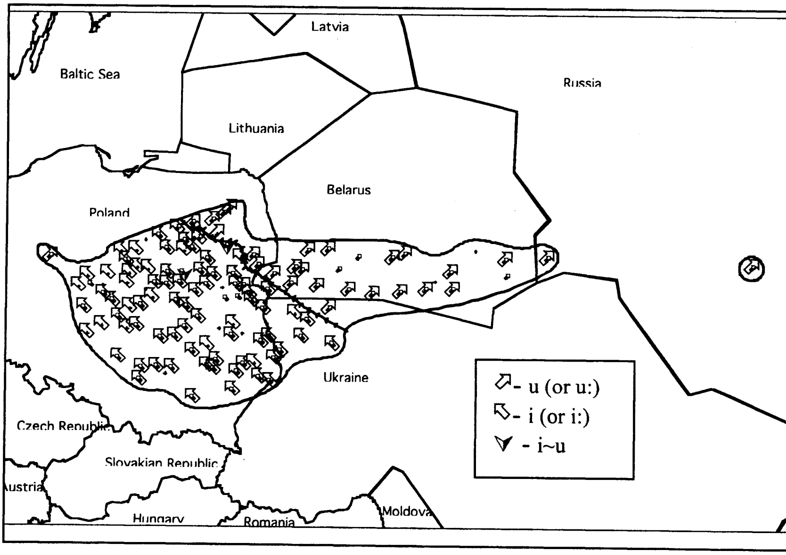


Note that the shifts in (11) are reminiscent of, but not identical to those depicted in (19) of Labov (1994:286), viz., chain shifting across subsystems in Central Yiddish -- based on Herzog's 1965 work on Yiddish in northern Poland.

2. Chain Shifts and Variation: The Case of *u~i*

Having established the existence of such chain shifts, further data were extracted from some of the maps in Volume III of the atlas ("The Eastern Yiddish – Western Yiddish Continuum"). This portion of the atlas resembles some of the traditional linguistic atlases available for other languages, in that each map typically represents the reflexes of one particular word, serving as one of a few examples for a linguistic phenomenon. We used maps pertaining to roughly the same vocalic phonemes as those we had used from Volume I for our initial study. Using the cartographic software package MapInfo at the Linguistics Laboratory at the University of Pennsylvania, the data for the specific points on the various maps were entered into a spreadsheet, so that it would be possible to superimpose data from different maps and compare the distribution of variants. Examining the specific data from Volume III revealed that the PY dialect area is by no means free of variation. There is a clear u~i distinction, with only a few points exhibiting either inter-lexical or intra-lexical variation.

Map II: The u~i isogloss



Map II illustrates how data from two sources can be combined into one single, more systematic representation. In this map, the larger arrows denote the reflexes of Proto-Yiddish $U_{2,3}$ in the word *dUx(ə)nən* 'perform the priestly benediction', and the smaller ones are reflexes of U_1 in the word *sUkə* 'tabernacle' or *sUkəs* 'Sukkoth, Feast of Tabernacles'. Superimposing the symbols of one map on those of the other (originally Maps 20 and 19 in Vol. III of *LCAAJ*, respectively), opens a window for comparison. In this case, most small "i" arrows fit into corresponding large "i" arrows, and most "u" arrows show the same correlation between the two proto-phonemes. The only inconsistencies are in close vicinity to what we have determined to be the general isogloss for this historical change.

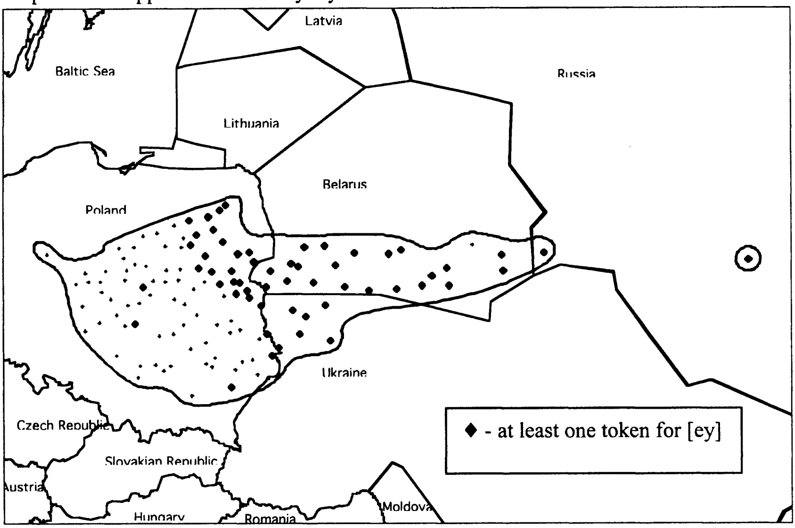
Our account of this part of the PY chain shift has now evolved from that of a solid, across-the-board sound change (see (1) above) to a more complex yet systematic view of the shift as being conditioned by a geographic boundary.

3. Chain Shifts: When Diphthongs Merge

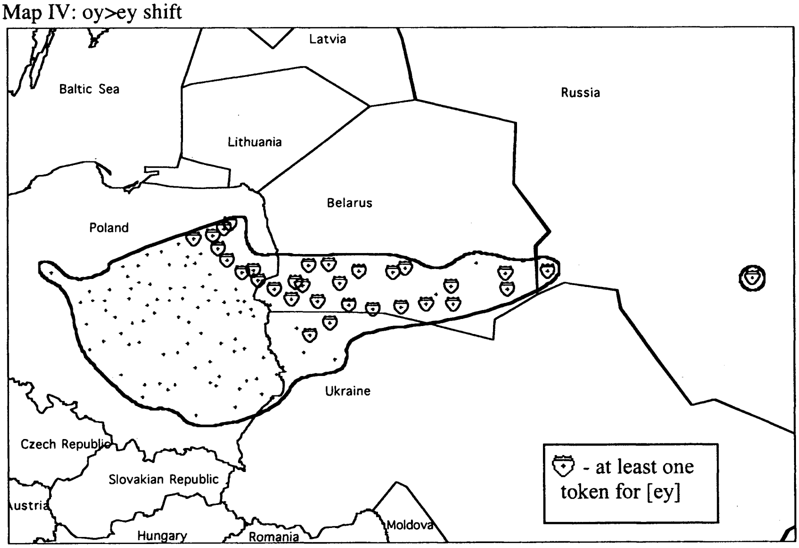
From our expository discussion of chain shifts (see 1 above), it may be inferred that the diphthongs /ay/ and /oy/ are typically results of shifts either across subsystems, as in (4) and (5), or within the subsystem of diphthongs, as in (6) and (7). This account of the diphthong inventory of PY suggests that the historic diphthong /ey/ (presumably related to the long monophthong /e:/, as they both result in the same reflex /ay/) no longer exists in contemporary PY dialects. Yet several maps in the *LCAAJ* suggest that in a large area within the broader PY dialect area, /ay/ is the prevalent reflex of Proto-Yiddish phonemes E_4 and $E_{2,3}$, while /ey/ is the equivalent elsewhere. Map III shows the distribution of /ay/ and /ey/. It is a compilation of data from four maps in the original *LCAAJ* (Vol. III): Map 6 for *miʃtEns gəzógt* 'expression of pity or content', Map 7 for *lE(ə)nən* 'to read', Map 8 for *bə(h)Eməs* 'cattle; cows', and Map 9 for *sExl* 'brains; sense'.

The diamonds represent points where at least one of these words was pronounced with an /ey/ diphthong. The rest are points where there was either no data, or (for the most part), where the data showed that these local dialects have undergone the ey>ay shift.

Map III: Non-application of the ey>ay shift



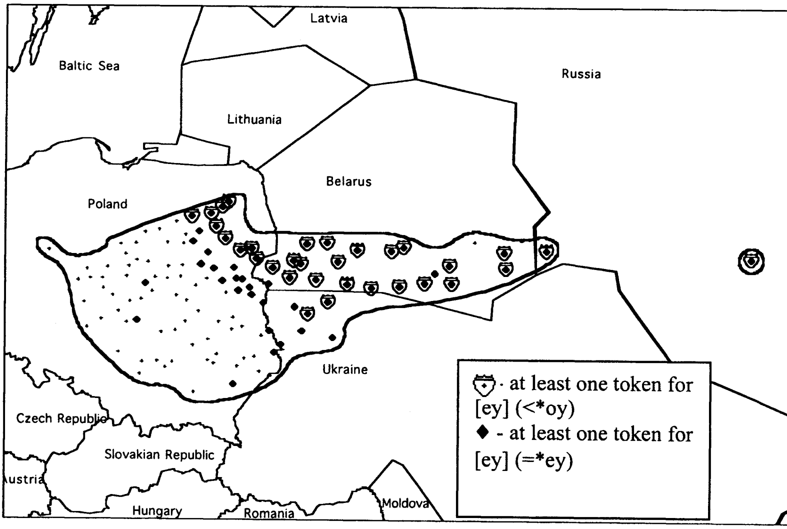
Map IV: oy>ey shift



As in the case of the u>i historic rule, superimposing information from different maps enables us to broaden the scope of our linguistic inquiry and refute some of the generalizations that may have resulted from coincidental data.

Three maps in the *LCAAJ* present the distribution of regional variants for Proto-Yiddish $O_{2,3}$. Map 10 deals with the word *brOgəz* 'angry', Map 11 with the word *jontOvim* 'Jewish festivals', and Map 12 with *xàləmO(ə)d* 'period between first two and last two days of Passover or Sukkoth'. Data combined from these three maps constitute Map IV. In this map are shown the towns and villages in whose dialects at least one instance of /ey/ for a historical /oy/ occurs.

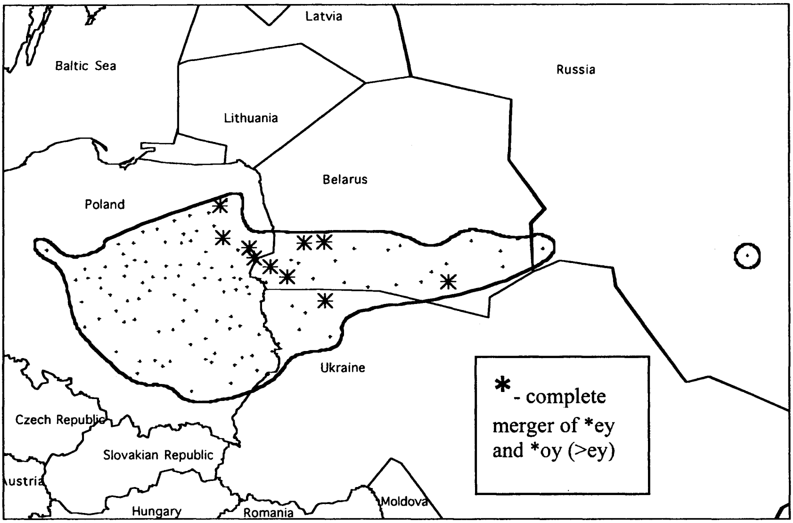
Map V: Initial evidence for oy/ey merger



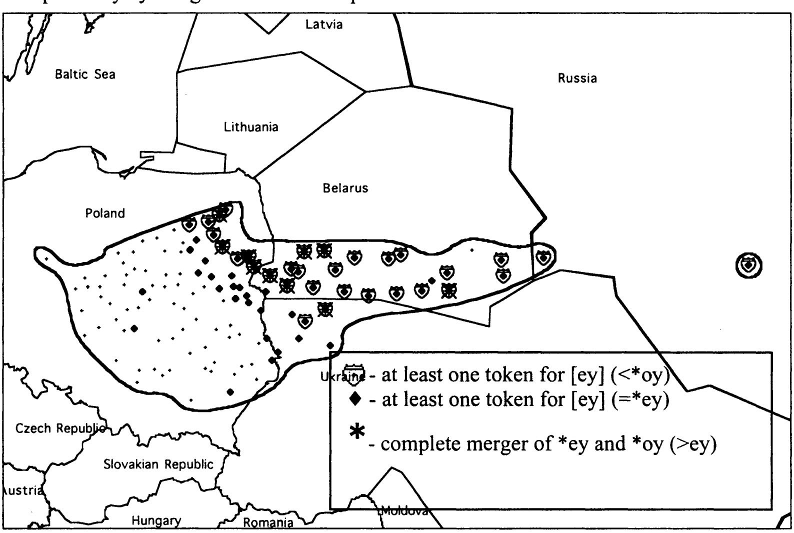
Next, we can superimpose Map IV on Map III and the results, visible in Map V are quite interesting. In all the locales where at least one historical /ey/ did not surface as [ay], there was at least one historical /oy/, which was pronounced [ey]. This leads to the hypothesis that there might be a merger of two diphthongs in some of these local dialects. This hypothesis can be confirmed by inspecting our database. In a few cases, a given locality was coded as having only [ey] reflexes for all proto-E and proto-O phonemes. In other words, in a subset of the area defined by Map IV, where at least one token of each set of proto-phonemes coincided, there is in fact a full merger.

This full merger is the subject of Map VI . Stars indicate locales in which a full merger is apparent from the data. Map VII is an illustration of the proportion of full mergers in respect to those cases in which the merger is incomplete.

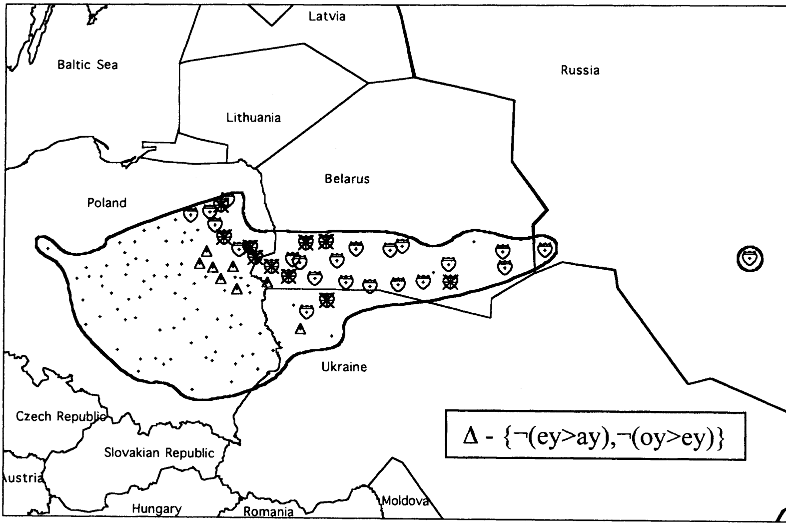
Map VI: oy/ey merger



Map VII: oy/ey merger: full vs. incomplete



Map VIII: Sound change(s) vs. stability



4. Conclusions

This paper combined data collected within the framework of structural dialectology with a variationist analysis according to the principles of chain shifts and mergers laid out by Labov (1994). Applying the features of the mapmaking software MapInfo to linguistics has proven fruitful in accounting for sound changes, some of which had been previously reported, and enhancing our understanding of such changes and of the variation in their occurrence across geographic regions. Examining chain shifts and mergers and whether or not these processes have reached completion becomes more tangible when different maps are combined and conclusions are drawn from them. Using *LCAAJ* as a preliminary source of data has been illuminating, yet not unproblematic. Some of the original recordings of the atlas interviews are likely to soon become available, allowing further research of these and other changes in PY and other dialects.

References

Herzog, Marvin et al. (eds.). 1992. *The Language and Culture Atlas of Ashkenazic Jewry* (Prepared and published under the aegis of an editorial collegium). Tübingen: Max Niemeyer; New York: Yivo Institute for Jewish Research.

Labov, William. 1994. *Principles of Linguistic Change*. Vol. 1: Internal Factors. (Language in Society; 20). Oxford, Eng. & Cambridge, MA: Blackwell.

Department of Linguistics
University of Pennsylvania
619 Williams Hall
Philadelphia, PA 19104-6305

urih@babel.ling.upenn.edu

Consistency and Creativity in First Language Acquisition

MICHAEL ISRAEL
University of Maryland, College Park

1. Creativity in a Usage-Based Grammar

Creativity poses a problem for linguistic theory (Braine 1971; Bowerman 1988, 1996; Pinker 1989). The problem is that speakers do not just reuse linguistic formulae they have heard before, but creatively manipulate language to form and interpret novel utterances. To do this, speakers require relatively abstract knowledge of linguistic structures, and children must somehow acquire this knowledge. Children, however, not only learn how to produce novel grammatical utterances, but also how to avoid ungrammatical ones. To do this, they must somehow discover the limits on the productivity of grammatical patterns. The problem of how children avoid constructing an over general grammar (Braine and Brooks 1995) thus turns on a basic tension between a need to use language creatively, and a need to conform to the usage one hears: the former requires relatively abstract linguistic representations, the latter requires specific knowledge about what sorts of things people actually say.

In this paper I argue that a general PRINCIPLE OF CONSISTENCY both motivates and constrains the process of analogical learning in a usage-based grammar. The basic idea is that children learning a language, and speakers in general, represent linguistic units in ways that maximize their motivation and emphasize their commonalities. Two units are consistent with each other to the degree that they match in their formal and semantic specifications. LOCAL CONSISTENCY applies to linguistic units activated online in usage events, and requires these to be as consistent as possible with entrenched utterance types. GLOBAL CONSISTENCY applies to the repertoire of constructions as a whole, and requires that units be represented in ways which maximize their consistency with each other. Local consistency favors a massive inventory of low-scope constructions to represent the rich details of experienced usage events: it thus fosters arbitrariness in the grammar, but also makes on-line processing easier by offering conventional units for every occasion. Global consistency favors the development of abstract representations and recurrent inheritance links across constructions: it thus increases motivation in the grammar, but also makes processing harder as the schematic units it favors are farther removed from the details of actual usage.

Global consistency motivates the emergence of schematic linguistic units which can license novel utterances; local consistency constrains the use of such units by encouraging conformance to familiar patterns of usage. The effects of consistency should be apparent in the ways children do—or systematically do not—use grammatical constructions creatively, and the bulk of this paper will therefore focus on the varieties of creativity observable in young children's use of a set of constructions featuring non-finite complement clauses. However, before we examine the consistency principle's empirical teeth, it may help to consider its theoretical roots in the usage-based theory of Cognitive Grammar (Langacker 1987, 1991, 1999).

This theory starts with the assumption that a grammar is a system of internalized cognitive routines which emerge to meet the exigencies of language use. More precisely, a grammar represents a speaker's knowledge of linguistic convention, and consists of a structured inventory of linguistic units, including phonological, semantic, and symbolic structures, represented with varying degrees of internal complexity, and at varying levels of abstraction (Langacker 1987: 73). The inventory is structured by a complex network of inheritance links marking part-whole, type-token, and based-on relations between units. The relation between grammar and usage is one of categorization. Linguistic units activated in usage serve to categorize, and so to sanction usage events. Usage events (or aspects thereof) which for some reason cannot be categorized, or which diverge too sharply from the specifications of a sanctioning structure, are experienced as either ungrammatical or uninterpretable. Typically, linguistic units will compete to be activated as the sanctioning structure for a given usage event: factors favoring selection include (i) priming (recency of prior activation), (ii) entrenchment (frequency of prior activation), and (iii) specificity (closeness of match to target structure). Often, an entire utterance may be sanctioned by a single, frequently recurring complex unit (e.g. *how do you do*); other times, more abstract constructions may be required to sanction the composition of two or more units in a complex whole.

Given these basic assumptions, language acquisition involves nothing more than the accumulation of linguistic units and the discovery of relations between them. All units are learned from experience, either as (parts of) overtly occurring expressions, or else as schematizations over previously mastered units. Acquisition begins with simple and concrete units (e.g. *bye-bye, all gone, mine, what dat?*) which the child can employ directly in the performance of specific linguistic acts. The emergence of flexible and creative use depends on the discovery of regular and recurring patterns found across a range of familiar utterance types. Such patterns are entrenched as (relatively) abstract constructional schemas capable of licensing a (potentially) open-ended set of utterance tokens. At first, such schemas will be very concrete, involving simple slot and frame structures based on specific uses of specific lexical items. Later, these may form the basis for further abstractions, as schemas are built on schemas to capture increasingly higher-order grammatical generalizations.

In principle, there is no limit to how abstract constructions can become, but there is a real question as to how much abstraction is necessary (cf. Tomasello 2000a, Fisher 2002). Recent work on child language points increasingly to the role of concrete, item-based units and frequency driven learning in the development of grammatical competence (e.g. Pine, Lieven & Rowland 1997; Tomasello 2000b); at the same time, work in connectionist modeling, language processing, and corpus linguistics has been converging on a view of linguistic structure which is massively usage-based and frequency driven (see, for example, recent papers in Barlow & Kemmer 2000, and Bybee & Hopper 2001). As a consequence, some researchers have become skeptical about the importance—or even the existence—of abstract linguistic representations. But there is no reason grammars should not include both concrete and abstract representations. Indeed, the consistency principle suggests that the two should in principle be inseparable, and that abstract schemas emerge in tandem with and as a function of an expanding repertoire of item-specific constructions. In this paper I will argue that while early grammatical representations are indeed massively item-based, from an early age children also seem to be sensitive to relatively abstract similarities across construction types.

2. Consistency and Non-finite Complement Clauses

The operation of consistency as a developmental principle makes four basic predictions about the development of abstract representations in child language: (i) Constructional Grounding—complex constructions will tend to be based on simpler, previously entrenched routines; (ii) Early Agrammaticality—early creative combinations may be (in some ways) the least constrained, since early on children will lack consistent patterns on which to model their own; (iii) Interference Effects—constructions with superficial similarities may be linked in ways that trigger performance errors or other confusions; (iv) Persistence —the more motivations an innovation has, (i.e. the more consistent it is with a child's overall repertoire of constructions), the more likely it is to persist.

Basically, we want to know how abstract child grammar gets, and how it gets as abstract as it does. Crucially, this means we need some way of observing abstractions in the concrete behavior of children's spontaneous usage. In this paper, I will be less concerned with the precise representation of children's linguistic structures, and more with the ways children's spontaneous productions may provide evidence for abstract representations in general. To this end I examine the emergence of non-finite complement clause (NFCC) constructions in seven children between the ages of 1;6 and 5 years, and identify four types of creative combinations typical of children's spontaneous usage.

Data comes from the ReVerb project (Israel, Brooks, & Tomasello, in progress), which coded the inflection, argument structure, and complement array of every instance of every verb in the spontaneous speech of seven children from the

Childes database (MacWhinney 1995). The particular constructions under study, NFCCs, form a family of related utterance types, all of which feature a matrix verb, a surface direct object and a non-finite complement of any category (X-Comp) predicated of the direct object. The examples in (1a-c) illustrate typical NFCC constructions. The schema in (2) captures some of the basic structural features which unite them.

(1) a. that [make$_{matrix}$ [it dark outside]$_{NFCC}$]. Eve 2;0
b. [put$_{matrix}$ [man on the shelf]$_{NFCC}$]. Nina 2;1
c. I'm gonna [get$_{mat}$ [the cow to drink some milk]$_{NFCC}$]. Peter 2;8

(2) [(NP$_1$) [V$_{matrix}$ [NP$_2$ X-Comp]$_{NFCC}$]]

A motley array of constructions fits this pattern: the schema abstracts over adjectival, prepositional, and verbal X-Comps, and ignores differences between raising, control, and small clause structures, among others. But while these structures may not form a single, coherent category in the grammar of English, they are, in many ways, globally consistent with each other: they share the same basic word order, reflecting a common, general strategy for encoding complex propositions in a single, finite clause. The question is how children learn to differentiate these constructions without being misled by their surface similarities. Given these similarities, this extended family of constructions offers fertile ground for grammatical innovations and overgeneralizations, and so provides an ideal laboratory in which to observe early linguistic creativity.

In the rest of this paper, I will concentrate on some of the more interesting ways the ReVerb children use NFCC constructions, and I will identify four classes of creative usage which bear out the consistency principle's four predictions about linguistic representation: the gradual emergence of flexible routines illustrates the process of constructional grounding; children's groping patterns attest to the role of early agrammaticality; various mixed constructions illustrate interference effects; and finally, two examples of novel constructional blends attest to the persistence of well-motivated innovative forms.

3. Constructional Grounding and Flexible Routines

Creativity comes in a variety of forms. We are concerned here with the productive use of familiar expressions in novel combinations—creativity which depends on the schematic representation of complex patterns across utterance types. Such representations are most evident when children say things they are unlikely to have heard elsewhere. When Eve at 2;2 says *I falled that down*, we can attribute this novel usage to the child's recognition of the abstract relation between causative and inchoative uses of other verbs like *break, open* and *grow* (cf. Bowerman 1996). Children may also produce utterances which appear thoroughly conventional, but which are just as novel from the child's perspective. This kind of creativity is harder to observe, but given the formulaic nature of early child language, it can be seen in the ways particular item-based routines gradually take on more and more flexible usages. Indeed, I will argue that it is precisely the gradual process of loosening up formerly rigid utterance routines that allows children to develop accurate and increasingly abstract representations of sentence structure. Local consistency limits the degree to which novel utterances may diverge from established patterns, and so ensures that new, more abstract patterns should emerge as extensions from or elaborations of simpler and more concrete linguistic units. This is what Johnson (1999:8) calls constructional grounding—a process whereby complex linguistic signs may be learned as special cases of other, simpler signs with overlapping distributions and similar constructional properties. In this case, the simple signs are unanalyzed low-scope formulae and the complex signs are the abstract constituency and valence constructions which these formulae instantiate. In this sense, low-scope, item-based constructions are not just a sign of children's underdeveloped grammatical abilities; they may in fact be the foundation on which more sophisticated and abstract grammar is built.

There is no shortage of evidence that many NFCC constructions start off as low-scope formulae. Adam's use of the verb *want* is a dramatic example. Most children use *want* mainly to express indirect requests (e.g. *I want X*) and to form indirect suggestions (e.g. *do you want X?*). In Adam's case, the tendency is strikingly exceptionless: over five months, between the ages of 2;9 and 3;2, all 179 instances of Adam's *want*+NFCC utterances are, like those in (3), 2nd person, present tense questions, used for the most part to seek approval for a proposed course of action Adam would like to pursue. Around 3;3, Adam expands his usage to include 1st person, present tense assertions like those in (4), which directly report his current desires. By the age of 5;2, with 301 examples in the entire corpus, Adam has only 8 instances of *want*+NFCC which do not conform to one of these narrow usages.

(3)	wan(t) me open it ?	2;9.4	Adam 13
	want me get out ?	2;9.18	Adam 14
	do want me ride it ?	2;10.30	Adam 17
	d(o) you want me drink hot coffee ?	2;10.30	Adam 17
	d(o) you want me # put hole in ?	2;11.13	Adam 18
	do want he talk ?	2;11.13	Adam 18
(4)	I don't want you to take it out # Mommy .	3;3.4	Adam 26
	I want Paul to drink .	3;3.18	Adam 27

The consistency of Adam's usage here clearly suggests that he is relying on complex, pre-compiled formulae. He appears to have both a [(do) (you) want NP VP?] "suggestion" schema and a [I (don't) want NP VP] "request" schema. In fact, the prototypical form of these schemas is even more narrowly defined: for instance, 168 of Adam's total 214 "suggestions" feature *me* as the direct object. Similarly, Adam's unadult-like omission of the complementizer *to* is disproportionately correlated with the "suggestion" schema, accounting for 194 (95%) of the 205 such omissions with *want*+NFCC; conversely, only 13 (14%) of the 90 correct uses of [*want to* +NFCC] occur in the "suggestion" schema.

The use of *see* tends to be more varied than that of *want,* though again there is a bias for uses with 1st and 2nd person subjects. The verb is often used to monitor and maintain joint attention between child and addressee, and most of the children use *see* regularly, if not exclusively, for this purpose. Sarah uses *see*+NFCC largely to focus attention on herself: 15 of her 24 *see*+NFCC utterances occur as part of a [*wanna see me* VP?] schema, as in (5).

(5)	want to see me roller+skate ?	4;0.5	Sarah 87
	wan(t) (t)a see me make an f@l ?	4;2.28	Sarah 98
	wan(t) (t)a see me write # make water ?	4;2.28	Sarah 98
	you wan(t) (t)a see me make a house ?	4;3.19	Sarah 101
	you wan(t) (t)a see me make a # straight line ?	4;3.26	Sarah 102
	you wan(t) (t)a see me write Sarah ?	4;4.25	Sarah 106

Adam's use of *want* and Sarah's use of *see* are compelling examples of narrow scope formulae; however, it might be a mistake to assume that these formulae represent all the children know about the way these verbs work with NFCCs. Sarah's use of *see* + NFCC, for example, is not limited to one routine, and she does occasionally vary both the arguments and the form of the verb (e.g. *see the birdie drinking* 3;10; *let me see you open it* 4;1). And even when Adam's use of *want*+NFCC is completely rigid, it's not clear that he can't understand the verb in other uses as well. What is clear is that these sorts of precompiled formulae make it easier for children to compose increasingly complex utterances on line for an increasingly wide variety of situations.

In this light, children's formulaic language need not be merely a sign of grammatical naiveté (though it may be that as well)—it is also a basic tool for the development of more sophisticated grammatical abilities. Narrow scope formulae allow children to break into complex syntax without having to compose complex sentences from scratch. Once they master a fixed formula, they can gradually learn what sorts of substitutions it affords, what sorts of constituents it contains, and what sorts of grammatical relations hold between them. In this sense, rigid formulae provide the foundation on which abstract and flexible constructions are built.

This process is evident in the way particular predicates come to combine with increasingly complex constituents over the course of development. For example, the *make*+NFCC constructions from Nina in (6) and Adam in (7) reveal a clear progression from simple, linear slot-and-frame patterns to complex, hierarchically organized structures. Early uses, as in (6a, 7a), have no inflection or expressed subject, and feature a stripped down complement with a pronoun and an intransitive verb. Examples like those in (6b, 7b) feature more complex constituents, with overt subjects and auxiliary constructions in the matrix clauses, and postverbal adjuncts and complements in the NFCC itself. The crowning achievements come in (6c, 7c), where the compound complements of the matrix verb *make* suggest some real understanding of the constituency of the NFCC as a whole, and of the NFCC as a constituent.

(6)	a.	make her stand up .	(Nina)	2;0.24
		make a sit down .		2;0.24
	b.	let's make him fly on this house .		2;9.21
		did the band+aid make it feel better ?		2;9.21
		you can't make # make these wheels move .		2;9.21
	c.	let's make them sit down and talk # Mommy .		2;11.6
		let's make the little doggy stand up and the mother stand up .		3;0.3

(7)	a.	make it walks .	(Adam)	2;11.28
		make him run		3;0.11
	b.	I gon make you drive on it .		3;1.9
		dis sometimes # makes me cry .		3;1.26
	c.	dat makes it stick out and stay up .		4;3.13

By the time Nina and Adam can manipulate these complex constituents, they have had months of practice with simpler substitutions in the same structures.

All seven children use the verb *help* with an NFCC at least once, and all of them use it specifically to request (or demand) assistance. Abe is particularly prolific with this use: out of 149 utterances with the verb *help,* Abe produces 43 distinct tokens with an NFCC, and of these 31 (72%) are (in)direct requests or imperatives, as in (8).

(8) Abe's [...(you)...help me VP] Construction

I can't find it, Dad, you help me find it, ok ?	2;10.15
help me take my sock off .	2;10.27
can you help me put it back in ?	3;1.1
will you help me find a hockey stick ?	3;1.26
are you gonna help me put em back in ?	3;3.15
you need to help me put em up I'm gonna do just two of em .	3;8.2
how'd you like to help me do wings for that [/] for that arrow ?	3;9.12
know what you could help me do ?	3;9.27

This data clearly shows that Abe has learned to use a specific construction with the verb *help* for a specific type of speech act. There is little evidence here that Abe has an adult-like syntactic representation for these sentences, or even that he understands the basic pragmatic principles governing his "indirect" speech acts. Rather Abe has a pivot-like *help me X* construction, which he productively combines with bare VP complements for particular pragmatic purposes. On the other hand, Abe's use of *help*+NFCC is not entirely rigid either: his earliest such utterance features a 1st person subject—*I help my mommy cooking* (2;6.4)—and later uses even include inanimate 3rd person subjects—*the candy we took to the movie helps you grow*, (3;8.17). Indeed, the significance of the narrow scope [*help me X*] formula is not so much that it reveals Abe's grammatical limitations, but rather that it provides a handy framework for further syntactic development. By relying on a well-entrenched, precompiled formula, Abe is able to experiment with a range of syntactically complex embedding constructions—*can you X, will you X, are you gonna X, you need to X, how'd you like to X,* etc.—all of which require a complement of the same syntactic type and all of which fulfill analogous pragmatic functions. In this sense, the very simplicity of Abe's formulaic usage actually lays the foundation for his mastery of more complex and abstract structures.

4. Early Agrammaticality and Groping

Children's earliest combinations are occasionally among their most creative, or at least their most anomalous. The examples in (9) illustrate some of the unusual liberties children take with basic word order in NFCC constructions.

(9)	I sock put on .	1;7	Eve 4
	mine take out .	2;3.18	Adam 2
	outside put book .	2;4.3	Adam 3
	take it Nina away .	2;0.10	Nina 5
	cheek put it on .	2;3.28	Nina 19
	put in my hair my barrette .	2;2.14	Peter 9
	we put in the glasses in the milk .	2;8.23	Naomi 70
	it throw away ?	3;3.20	Sarah 054

Braine (1976) refers to such apparently free word orders as groping patterns. If grammar is learned from experience, this sort of freedom makes sense early in grammatical development. Early on, when children have only a small and heterogeneous set of concrete constructions at their disposal, they will lack reliable patterns on which to model new utterances. And although local consistency requires that children should conform to the usages they know, global consistency predicts that children's utterances may be less constrained when they know fewer usages. The smaller the repertoire of constructions, the harder it is to recognize general patterns uniting them, so when children do come up with novel combinations, they will be free to combine them in any way that suits their pragmatic purposes. In this sense, groping constructions are not really *un*grammatical, but rather *a*grammatical.

Such uses need not be entirely random. In general, groping may occur whenever a child has not yet mastered the combinatorial niceties of a given construction. The examples in (10) reveal a pattern of innovations where children have begun to use verb+particle constructions, but have not yet learned that light, pronominal objects obligatorily precede the particle.

(10)	I need to plug in it .	3;1.9	Adam 22
	I hang up this .	2;1.1	Naomi 44
	get away this .	2;2.6	Nina 13
	you screw off it .	2;3.28	Nina 19
	put on it .	1;11.7	Peter 04

Presumably, the various verb+particle combinations here have been learned as units to which the children simply add a pronominal object. It is worth noting that most of these children also used the [V NP Particle] order at the time of these utterances, though usually with different verbs and particles: what these children apparently lack is a general schema for postposed particles, and a general understanding of when this schema is obligatory.

5. Interference and Mixed Constructions

Global consistency predicts that linguistic representations should maximize similarities across linguistic units: whether such similarities reflect deep structural relations or fortuitous surface resemblances, both will be grist for the mill of schema abstraction. False analogies and overgeneralizations are thus to be expected. Interference effects arise as a child's linguistic repertoire expands and constructions with similar formal and semantic properties compete to license usage events. The basic principle seems to be that utterance types which share some features are likely to share more, and the clearest manifestations of this are mixed constructions—nonce uses combining properties of two or more distinct constructions. Typical examples involve a verb or other lexical head appearing in a semantic frame or grammatical structure associated with some closely related expression. Thus, in (11), Peter incorrectly uses a *to*-infinitive with *make* after correctly using it in a similar causative construction with *get*; in (12) he makes the opposite mistake, erroneously using the same bare stem infinitive with *get* that he correctly uses with *make* just moments later.

(11)	made him stand up .	2;8.14	Peter 17
	I'm gonna get the cow to drink some milk .	2;8.14	Peter 17
	*make a boy to ride on here .	2;8.14	Peter 17
(12)	*let's just get it stand .	3;1.21	Peter 20
	I can make it stand .	3;1.21	Peter 20

Similarly, Nina in (13) and Adam in (14) produce NFCCs with peculiar gerundive complements in contexts where the gerundive complement has been primed by other, well-formed NFCC constructions.

(13)	I wanna see the lady dancing .	2;5.28	Nina 31
	I don't want the water falling .	2;5.28	Nina 31
	*<I want a> [//] I want my doll's waking up .	2;5.28	Nina 31
(14)	I saw a duck swimming in (th)e water .	3;0.25	Adam 21
	*why you goin(g) put truck parking ?	3;0.25	Adam 21

These sorts of errors, though very common, are rarely very productive. But while they may reflect mere momentary confusions, they also clearly depend on some implicit recognition of the relations among NFCC constructions.

Interestingly, certain types of confusion are quite common across subjects. For instance, children regularly seem to overextend causative predicates in NFCC constructions. Thus in (15) *make,* which normally requires a verbal or stative complement, shows up with a locative complement. Similarly, in (16), children use *put* (='cause X to be located at Y') with stative complements, and in (17) *give* (='cause Y to have X') with locatives.

(15)	I # I make cream on dolly's hair .	2;3.28	Nina 19
	make the duck off .	3;5.20	Sarah 63
(16)	I'm gonna put my suitcase full of stuff .	3;3.28	Abe 089
	I tryin(g) put the sink off .	3;7.23	Sarah 070
(17)	I am I'm gonna give it up there at the ceiling, see?	4;6.19	Abe 185
	I give milk in .	2;3.28	Nina 19
	I # feed xxx (s)paghetti # on my leg .	2;10.2	Adam 15

Since many matrix predicates allow two or more different types of X-comp (e.g. *get* allows stative, locative and verbal X-comps) this would appear to be a well-motivated type of overextension. Indeed, it is striking that most of the seven children come up with the same (or very similar) overextensions.

6. Persistence and Constructional Blending

While mixed constructions are fairly common, they also tend to be somewhat fleeting in their overall effect. Where a schema is supported by a relatively small number of instances, its potential to license new utterances will be relatively weak; but where schemas capture robust similarities across a large set of utterance types they should be productive and resilient. Consistency predicts that the productivity of any schema should be a function of its global motivation in a range of exemplars and its overall compatibility with local consistency.

It is interesting in this light to consider Sarah's *(e)rase X off* and Nina's *wear X on* constructions in (18-19), both of which are produced over a period of eight months or more, and both of which are motivated by a family of well-entrenched and closely related constructional routines.

(18) Sarah: *(e)rase X off*: 7 tokens; 4;3 - 4;11

can you erase dis off ?	4;3.13	Sarah 100
&c # can I (e)rase the red off ?	4;3.13	Sarah 100
(e)rase it off .	4;4.11	Sarah 104
I got ta (e)rase some off .	4;4.25	Sarah 106
for a minute # to (e)rase that off .	4;7.0	Sarah 116
mmhm # I have to erase that off .	4;11.13	Sarah 131

(19) Nina: *wear X on,* 11 tokens; 2;6-3;3

you wear gloves on in Mantha's house .	2;5.26	Nina 29
bears don't wear clothes on .	2;10.13	Nina 36
you can wear this blanket on tonight .	2;11.6	Nina 39
no # with my undershirt I wanna wear that on .	3;2.12	Nina 51
I want the kitties to wear clothes on .	3;2.16	Nina52
these kind of dresses that I'm wearing on .	3;3.1	Nina54

The examples in (20) show some precursors for Sarah's use of *erase X off*. By the time Sarah starts using this construction at 4:3.13, she has (a) already used the verb *erase* correctly in simple transitives, (b) used the particle *off* with at least 10 other transitive predicates (among others, *take, wipe, cut, get, pull, peel, pick)*, and (c) created at least three other novel *V X off* combinations.

(20)	a.	just erase it .	4;3.7	Sar 99
	b.	I take em off .	3;0.27	Sar 42
		wipe my boo+boo off .	3;1.3	Sar 43
		I wan(t) (t)a cut the corners off .	4;1.28	Sar 93
		how do you get the head off ?	3;10.30	Sar 83
		I tryin(g) pick it off .	3;7.16	Sar 69
	c.	you can work this off huh # Mommy ?	4;0.14	Sar 88
		I got ta write some off .	4;2.28	Sar 98

Similarly, by the age of 2;6, when Nina begins her *wear X on* usage, she has been using *wear* in simple transitives for at least six months (21a), and has used the particle *on* with at least 6 distinct transitive verbs: 9 times with *have* (21b), 81 times with *put* (21c), and once each with *leave, need, make* and *keep* (21d). And in most of these uses, *on* has the same idiomatic semantics it takes with *wear,* specifically denoting a state of being dressed.

(21)	a.	I want # I want to wear it .	2;3.18	Nina 18
		what's he wearing ?	2;3.28	Nina 19
	b.	she have jamas@f on .	2;2.6	Nina 13
		you have a blanket on and we go in a carriage .	2;2.28	Nina 15

c.	put dress on .	2;1.6	Nina 09
	let me have # my put my pants on .	2;1.29	Nina 12
d.	leave it on .:	2;1.22	Nina 11
	I made a basket on .	2;2.6	Nina 13
	her need this seat on # ok ?	2;3.14	Nina 17
	no # I keep it on .	2;4.26	Nina 23

Sarah and Nina's innovations here are, in fact, so well-motivated that it is hard to see why they are not part of the adult language, and it is easy to see why the girls might be reluctant to give up such seemingly natural usages.

7. Conclusions

I take my observations here to support three basic conclusions. First, early child grammar involves both rote-learning and complex creativity, and these are complementary rather than antithetical processes. Second, children need not depend on a conservative learning strategy: they will over-extend their grammars where the principles of consistency give them good reason to do so, and recovery from such overextensions may take a very long time. Finally, there is more to creativity than first meets the eye. In this paper I have identified four distinct types of linguistic creativity, all more or less directly observable in children's spontaneous performance data. Further study is of course required. The important point is that one may learn a great deal about children's grammatical competence just from a careful analysis of the ways children really do say the darnedest things.

Acknowledgments

Much of the research for this paper was conducted during my time at the Max Planck Institute for Evolutionary Anthropology. Earlier versions were presented at the Amsterdam Center for Language and Communication (October 2000) and at the ICLC in Santa Barbara (July 2001). I am grateful to members of those audiences, to the audience at BLS, and to Kirsten Abbot-Smith, Heike Behrens, Nancy Chang, George Lakoff, Elena Lieven, and Michael Tomasello for their comments, as well as to Patricia Brooks and Tess Wood for comments on an earlier draft. What imperfections remain are entirely my own responsibility.

References

Barlow, Michael & Suzanne Kemmer (eds.). 2000. *Usage based models of language.* Stanford: CSLI Publications.

Bowerman, Melissa. 1988. The 'no negative evidence' problem: Ho do children avoid constructing an overly general grammar? In J. A. Hawkins (ed.) *Explaining Language Universals* , 73-101. Oxford: Basil Blackwell.

Bowerman, Melissa. 1996. Argument structure and learnability: Is a solution in sight? *Proceedings of the Berkeley Linguistics Society* **22**: 454-468.

Braine, Martin. 1971. On two types of models of the internalization of grammars. In D. I. Slobin (ed.) *The Ontogenesis of Grammar: A theoretical symposium*, 153-86. New York: Academic Press.

Braine, Martin. 1976. *Children's first word combinations.* Monographs of the Society for Research in Child Development. Chicago: University of Chicago Press.

Braine, Martin & Patricia Brooks. 1995. Verb argument structure and the problem of avoiding an overgeneral grammar. In M. Tomasello & W. Merriman (eds.) *Beyond names for things: Young children's acquisition of verbs*, 353-376. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bybee, Joan & Paul Hopper (eds.). 2001. *Frequency and the emergence of linguistic structure.* Amsterdam/Philadelphia: John Benjamins.

Fisher, Cynthia. 2002. The role of abstract syntactic knowledge in language acquisition: A reply to Tomasello (2000). *Cognition* **82**: 259-78.

Israel, Michael, Patricia Brooks, & Michael Tomasello. In progress. *The ReVerb code book.* Leipzig, Germany: Max Planck Institute for Evolutionary Anthropology.

Johnson, Christopher. 1999. *Constructional grounding: The role of interpretational overlap in lexical and constructional acquisition.* PhD dissertation, UC Berkeley.

Langacker, Ronald W. 1987, 1991. *Foundations of cognitive grammar*, vols. 1 and 2. Stanford, California: Stanford University Press.

Langacker, Ronald W. 1999. *Grammar and conceptualization.* Cognitive Linguistics Research **14**. Berlin & New York: Mouton.

MacWhinney, Brian. 1995. *The CHILDES Project: Tools for analyzing talk.* Hillsdale, NJ: Erlbaum.

Pine, Julian, Elena Lieven, & Caroline Rowland. 1998. Comparing different models of the development of the English verb category. *Linguistics* **36**: 807-30.

Pinker, Stephen. 1989. *Learnability and cognition: The acquisition of argument structure.* Cambridge, Massachusetts: MIT Press/Bradford Books.

Tomasello, Michael. 2000a. Do young children have adult syntactic competence? *Cognition* **74**: 209-253.

Tomasello, Michael. 2000b. The item-based nature of children's early syntactic development. *Trends in Cognitive Science* **4**.

3101 Susquehanna Hall
University of Maryland
College Park, MD 20742

israel@wam.umd.edu

Constellations, Polysemy, and Hindi –ko

SHRAVAN VASISHTH
University of Saarland

BRIAN JOSEPH
Ohio State University

1. Background: General Issues

Much of linguistic analysis rests on a single key question: given entities X and Y as objects for analysis, are they the same or different? This issue pervades all components of grammar: phonetics, phonology, morphology, syntax, etc. Moreover, in addressing this issue, one often needs to recognize the relevance of different levels of analysis, especially underlying versus surface, since underlying sameness can be surface difference, and vice versa. For example, phonemic analysis takes phones with decidedly different phonetic realizations (e.g. aspirated vs. unaspirated stops in English) and treats them as the same at the phonemic level if their distribution does not overlap. But at the same time, segments that seem to be the same phonetically on the surface and even phonemically as well, e.g. the [*d*] of *recede* and the [*d*] of *invade*, might need to be treated as different from a morphophonemic standpoint, since, in this example, the former alternates with [*š*] in the related noun *recession* whereas the latter alternates with [*ž*] in *invasion*, both nominal formations having ostensibly the same suffix. In syntax, too, patterns that are alike on the surface, such as control constructions (e.g., *Skippy tried not to mind*) and raising constructions (e.g., *Skippy seemed not to mind*), can show some unlike properties that lead, in most current theoretical frameworks at least, to structural differentiation in some way, e.g., in underlying structure (cf., **It tried not to be raining* vs. *It seemed not to be raining*). As these examples indicate, an answer to the above key question regarding sameness often involves a recognition of differences too. Thus, the issue becomes one of measuring similarities and differences against one another and weighing the relative importance of one or the other, as well as deciding how to represent the sameness or difference that one ends up positing.

2. Same vs. Different in Morphology

The few examples in section 1 involve phonology and syntax, but, as noted at the outset, the same-vs.-different question pertains to morphology also, where the issue is rather: When are two morphs to be considered related to one another?

While there is the purely conditioned sort of difference that can be appealed to, as with phonemics (i.e., where there are allomorphs of a single morpheme),at the level of morphology, meaning is also involved; therefore one has to balance purely formal similarities and differences with semantic/functional similarities and differences. The traditional approach to this issue in morphology has been essentially all or nothing relatedness: one looks for recurring elements of form matching with recurring elements of meaning and posits a single underlying form that unites the variant realizations. For example, *repose* [ripoz] and *repository* [ripazɪtori] share [ripVz-] in form and 'put to rest' (or the like) in meaning, and this sameness can be expressed through an underlying root form having (roughly) the shape /ripoz-/.

But this discovery procedure is not without some difficulties. One the formal side, suppletion, as in *go/went*, presents a problem as there is usually little or no formal sameness to draw on (e.g., is the [+back] specification of [*g*] and [*w*] enough to allow for *go* and *went* to be connected formally?). On the semantic side, a problem arises with words that are etymologically, and still formally, related but have come to show different meanings, e.g., *suppose/suppository* (which as a pair is formally similar to, but semantically different from, *repose/repository*). More generally, then, any sort of drift, especially in form but also in meaning or even both, poses problems. For example, in *two/twelve* there is (some) formal separation (#t- vs. #tw-, and different vowels) but still a plausible semantic connection, while in *two/twine* there is both formal and semantic separation; other parallel cases can be easily found.

3. Other Solutions

In a sense, the null case from the formal standpoint comes when the forms one is considering relating are identical, differing only in meaning. Such a situation typically lends itself to a different type of solution. In particular, when the items under consideration show no formal differences, generally one can invoke polysemy as the means of representing the connection, and see the differences in meaning as a series of extensions, possibly metaphorical in nature, along one of the dimensions of the meaning of a given form. An example is the classic analysis by Brugman (1988) of *over* in English (where the link has to do with mapping the relation of a "trajector" to a "landmark", i.e. ABOVE x ACROSS) or with a putative connection linking English prepositional *to* with infinitival *to*, with directional *to*, and with indirect object marking *to* (in terms of movement towards some goal), etc.

However, when the forms themselves are not identical, in order to capture unity-in-diversity as well as diversity-in-unity, a different approach has been argued for, the CONSTELLATIONAL APPROACH of Janda and Joseph (1986). This approach involves two constructs, defined in (1):

(1) a. THE CONSTELLATION: A group of elements which share at least one characteristic property of form but are distinguished by individual idiosyncrasies of both form and function that prevent their being collapsed with one another.

b. Meta-template: Ameta-level redundancy statement, which ranges over all relevant candidates and equates instances of a particular formal configuration that meet certain criteria of uniting properties.

Thus, diversity is characterized by the Constellation, where differences in the elements are recognized, and unity by the Meta-template, which "parses", and thus identifies, all relevant equitable instances of a determined form. Moreover, the Constellation and Meta-template are related, in that, e.g., morphological constellations are ensembles of word-formational elements (e.g., morphemes) united by meta-templates which express the formal and functional identities shared by a set of distinct morphemes or, alternatively, uncollapsible morphological rules or constraints. The Constellation and the Meta-template together provide a mechanism that allows a realistic, non-procrustean approach to sameness in linguistic analysis – a recognition of how elements can simultaneously be same but also different (uncollapsible). Examples in the literature include Sanskrit reduplication (Janda & Joseph 1999), Arapesh plurals (Dobrin 2001), and Finnish definitives (Välimaa-Blum 1989).

4. Our Goal: A Constellational Account of *-ko*

We argue here that invoking constellations is appropriate even when identical forms are involved, and use constellations to clarify the relationship among several interconnected elementsin Hindi, all with the shape *ko*, that have preciously been misanalyzed as merely polysemous. Moreover, we extend the range of evidence available to confirm constellational status, going beyond distributional facts and syntactic behavior, and bring in relevant experimental results from language comprehension studies that bear on the similarities and differences among these elements *ko*.

5. The element(s) *–ko* in Hindi

The relevant elements are the Hindi postposition markers with the shape *ko*. Most relevant analyses (e.g., McGregor 1995, Wunderlich 2000) either conflate them or claim several distinct uses for this single element:

(2) a. Riṭa Sita-**ko** akhbaar de-gii
Rita Sita-ko newspaper give-fut
'Rita will give (a/the) newspaper to Sita.'

b. Ram-**ko** bhuukh lagii hai
Ram-ko hunger feel is
'Ram is feeling hungry.'

c. Kavita kitaab(**-ko**) paṛh-rahii hai
Kavita book(-ko) read-cont is
'Kavita is reading a (the) book.'

d. Kalika-ne Seema-ko kitaab khariidne-**ko** bolaa
Kalika-erg Seema-dat book buy-inf told
'Kalika told Seema to buy a book.'

e. Hari shaam-**ko** Ravi-ke ghar gayaa
Hari evening-ko Ravi-gen house went
'Hari went to Ravi's house in the evening.'

These *–ko*'s appear to be a single polysemous morpheme, with a common form *-ko*, and identical positioning vis-à-vis topicalizer *–to* and focalizer-*tak*, occurring only to the left of the marker with *–to*, and only to the right with *–tak*.

(3) a. Riṭa Sita-**ko-to**/*to-ko akhbaar de-gii
Rita Sita-ko-top newspaper give-fut
'As for Sita, Rita will give her (a/the) newspaper.'

b. Ram-**ko-to**/*-to-ko bhuukh lagii hai
Ram-ko-top hunger feel is
'Ram is feeling hungry.'

c. Kavita kitaab-**ko-to**/*-to-ko paṛh-rahii hai
Kavita book-ko-top read-cont is
'As for the book, Kavita is reading it.'

d. Kalika-ne Seema-ko kitaab khariidne-**ko-to**/*-to-ko bolaa
Kalika-erg Seema-dat book buy-inf-top told
'Kalika did tell Seema to BUY a book.'

e. Hari dopehar-**ko-to**/*-to-ko jaaye-gaa
Hari afternoon-ko-top go-fut
'Hari will go in the AFTERNOON (but maybe not any other time).'

(4) a. Riṭa Sita-**tak-ko**/??ko-tak akhbaar de-gii
Rita Sita newspaper give-fut
'Rita will give even Sita (a/the) newspaper.'

b. Ram-**tak-ko**/??ko-tak bhuukhq lagii hai
Ram hunger feel is
'Even Ram is feeling hungry.'

c. Kavita kitaab-**tak-ko**/??ko-tak paṛh-rahii hai
Kavita book read-cont is
'Kavita is reading even the book.'

d. Kalika-ne Seema-ko kitaab khariidne-**tak-ko**/??ko-tak bolaa
Kalika-erg Seema-dat book buy told
'Kalika told Seema to even buy a/the book.'

e. Hari dopehar-**tak-ko**/??ko-tak jaaye-gaa
Hari afternoon-ko-even go-fut
'Hari will go even in the afternoon.'

Even though these *ko*'s all behave alike here, we claim that they do not represent mere polysemy. In particular, a polysemy account cannot explain the semantic spread for *ko*: even if some functions are relatable, e.g., indirect objects and infinitivals as goal-oriented, no semantic principles or pathways can derive definiteness from any of the other functions. Moreover, these *ko*'s always show differentiated behavior.

The constellational approach predicts the occurrence of non-semantic differentiation. This is found among the various *ko*'s. Consider the interaction of the *ko*'s with *bhii* 'also, even'[1] and *hii* 'only': only the Indirect Object and Subject *ko*'s can be freely ordered; the other *ko*'s must occur to the left of these particles. (5) shows the cooccurrence patterns with *bhii*; those with *hii* are identical.

(5) a. Rita Sita-**ko-bhii/bhii-ko** akhbaar de-gii
Rita Sita-ko-even newspaper give-fut
'Rita will give even Sita (a/the) newspaper.'

b. Ram-**ko-bhii/bhii-ko** bhuuk lagii hai
Ram-ko-even hunger feel is
'Even Ram is feeling hungry.'

c. Kavita kitaab-**ko-bhii**/??**bhii-ko** par̥h-rahii hai
Kavita book-ko-even read-cont is
'Kavita is reading even the book.'

d. Kalika-ne Seema-ko kitaab khariidne-**ko-bhii**/??**bhii-ko** bolaa
Kalika-erg Seema-dat book buy-inf-even told
'Kalika told Seema to even buy a/the book.'

e. Hari dopehar-**ko-bhii**/??**bhii-ko** jaaye-gaa
Hari afternoon-ko-even go-fut
'Hari will go even in the afternoon.'

6. Evidence from language comprehension studies

The constellational approach assumes that there is a purely formal and grammatical basis for differentiating among the various *ko*'s, and that these *ko*'s really are different entities (though united by a meta-level redundancy statement).

We present evidence from language processing that further supports the constellational view by demonstrating the existence of undifferentiated as well as differentiated treatment in processing of the case-marker versus infinitival *–ko*.

Davison (1991) and Butt (1993) have argued that *ko*-marked infinitivals are simply nominals with ordinary case-marking: "...the constituent headed by the infinitive not only has the distribution of an NP, it can take case markers and

[1] See Schwenter & Vasishth 2000 for the distinction between *–tak* and *–bhii*.

undergo some further morphological processes that only apply to NPs. The entire infinitival 'clause' must therefore be analyzed as an NP" (Butt 1993:52).

They provide linguistic evidence that this is the case: (1) The inflections on the infinitivals are similar to nominal inflections; (2) In an infinitival, *-ko* can be replaced by clear postpositions, like *keliye* 'for'; and (3) Coordination of two NPs versus two infinitivals behaves identically. By contrast, Mohanan (1994:13-14), Bickel & Yadava (2000), and others suggest that infinitivals are verbal.

If infinitivals are merely NPs with *–ko* case marking, then the expectation is that these will behave like NPs during real-time sentence comprehension (assuming that there is a close connection between linguistic constructs and their mental representation). In particular, they should be stored in short-term or working memory as NPs. Center embeddings are a good way to test these hypotheses. Consider the double self-center embedding below:

(6)	Siitaa-ne	Hari-ko	Ravi-ko	kitaab	khariid-neko	bol-neko	kahaa
	Sita-erg	Hari-dat	Ravi-dat	book	buy-inf	tell-inf	told

'Sita told Hari to tell Ravi to buy a/the book.'

Pre-theoretically, parsing such a sentence in real time involves (a) **storing** each NP as it is encountered; and (b) **integrating** the NPs with verbs as the verbs are encountered. However, when NPs are stored in memory, they are encoded in some way rather than being stored as-is, and there is much evidence that NPs in the context of a sentence generate predictions (Lewis & Nakayama 2001, Gibson 2000, Vasishth 2002). These predictions are expectations of verbs and of sentence structures. In self-paced reading experiments involving Dutch (Dickey & Vonk 1997, Kaan & Vasić 2000), we find an invariant pattern: arriving at a verb after seeing an array of NPs results in faster reading time – the integration of NPs with a verb reduces local processing load, perhaps because the NPs and verb are now stored in working memory as one unit, a "chunk" (Miller 1956, Lewis 1996).

If infinitivals are really NPs, they should not be involved in any integration-related speedups, which (we are assuming) is a property of verbs. Thus, there should be no speedup at the innermost infinitival if it is an NP. Psycholinguistic experiments show that the infinitival behaves just like a verb – there is a significant fall in reading time (RT) at the infinitival. This is in contrast to the monotonically nondecreasing RT observed with the successive appearance of NPs (a reflection of storage costs).

The two experiments presented here are noncumulative moving window self-paced reading tasks (Just, Carpenter, & Woolley 1982). The procedure is as follows. First, a set of blank lines appear on the computer screen, each corresponding to a word in the sentence to be read. Then, with each press of the space bar, the subject sees each phrase successively, and the previous phrase disappears. The dependent measure is RT for each phrase. After the sentence ends, a yes/no comprehension question is presented to ensure that subjects are attending to the sentence.

The first experiment uses single center embeddings containing ditransitive (lexical causative) embedded verbs; although the critical comparison here is between presence versus absence of specificity/definiteness marking on direct object (1 x 2 design), in the present discussion we are interested in the difference between RTs at NPs versus verbs. The sentence types of interest are shown in (7); each phrase separated from others by white spaces is a separate region.

(7) a. Siitaa-ne Hari-se Ravi-se kitaab le-neko kahaa
Sita-erg Hari-abl Ravi-abl book take-inf told
'Sita told Hari to take a book from Ravi.'

b. Siitaa-ne Hari-se Ravi-se kitaab-ko le-neko kahaa
Sita-erg Hari-abl Ravi-abl book-acc take-inf told
'Sita told Hari to take the book from Ravi.'

As Figure 1 shows, there is a significant speedup at the infinitival element. This is consistent with integrative ("verb"-like) processes occurring at the infinitival, but not with storage ("noun"-like) processes. Therefore, the results support the assumption in Mohanan 1994 and Bickel & Yadava 2000 that infinitivals are verbs.

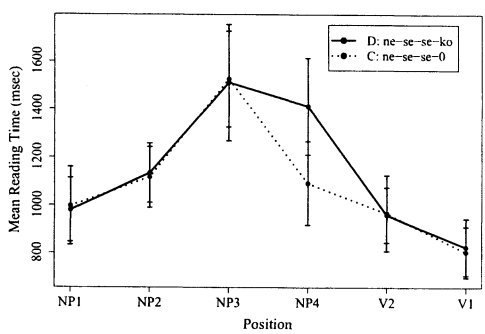


Figure 1: Results for Experiment 1 (95% confidence intervals)

A second experiment provides independent support for the conclusion that in real-time processing infinitivals behave like verbs.

Research on working memory suggests that human parsing involves a predictive component: possible sentence completions are anticipated as a sentence is being processed. Coupled with the fact (Mahajan 1990:87-88) that an adverb in Hindi must attach to a verb-projection, we can manipulate the degree of confidence in the prediction that the next word is a verb. A higher degree of confidence would mean a faster RT at the verb. By contrast, if the infinitival is an NP, then there is no reason to expect a speedup at the infinitival as a consequence of adverb-insertion; if anything, there should be a slowdown, since the infinitical NP would then have to be "reconstituted" as a verb for parsing purposes.

In order to test these hypotheses, a self-paced reading study was conducted with a 1 x 2 design. The factors were presence or absence of adverb between final NP and first verb.

(8)	a.	Siitaa-ne	Hari-ko	[Ravi-ko	[kitaab-ko
		Sita-erg	Hari-dat	Ravi-dat	book-ko
		khariid-neko]	bol-neko]	kahaa	
		buy-inf	tell-inf	told	
		'Sita told Hari to tell Ravi to buy the book.'			
	b.	Siitaa-ne	Hari-ko	[Ravi-ko	[kitaab-ko
		Sita-erg	Hari-dat	Ravi-dat	book-acc
		jitnii-jaldi-ho-sake	khariid-neko]	bol-neko]	kahaa
		as-soon-as-possible	buy-inf	tell-inf	told
		'Sita told Hari to tell Ravi to buy the book as soon as possible.'			

As Figure 2 shows, RT is significantly faster at the infinitival when an adverb is present. This is consistent with the assumption that the infinitival is a verb and not with it being an NP.

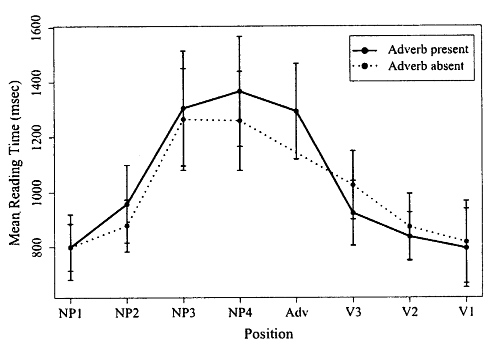


Figure 2: Double embeddings in Experiment 2 (95% confidence intervals)

In sum, the all-or-nothing assumption leads to the infinitival-as-NP debate, and raises two questions; (1) Is the *ko* marking on the infinitival ordinary case marking or not? (2) How do we resolve the apparent contradiction that these infinitival embedded clauses behave like NPs (purely linguistic evidence) as well as verbs (reading comprehension studies)? By contrast, the Constellational approach suggests that the infinitival-as-NP debate is asking the wrong questions. Under this view, it is completely acceptable if the *ko* of *...neko* is not segmentable and categorizable in exactly the same way as in *Sita-ko*—the meta-level redundancy statement will "pull out" a *ko* and equate the *ko* of *...neko* with the *ko* of *Sita-ko*, due, e.g., to their parallel behavior vis-à-vis *to* and *tak* (as a feature that unites them). Moreover, under the Constellational view, these infinitival embedded clauses are *expected* to be similar to NPs in some respects (linguistic tests) and different in other respects (their behavior in real-time sentence comprehension).

7. Conclusion

The Hindi *ko*'s, therefore, show that constellations achieve finer granularity in empirical coverage than pure polysemy accounts, thus allowing the analyst to have his/her cake and eat it too: elements can be same and different, but in a principled manner.

Moreover, the experimental evidence independently motivates the unity in diversity that is absent from the strictly unifying polysemy approach. The experimental results also provide a good example of the empirical consequences of adopting the Constellational approach; without it, debates like (Butt, 1993) *versus* Mohanan (1994) ("are infinitivals NPs or not?") are bound to arise, but these debates (although useful) are raising the wrong question.

References

Bickel, B. & Y.P. Yadava. 2000. A fresh look at grammatical relations in Indo-Aryan. *Lingua* **110**: 343-373.

Brugman, C. 1988. *The Story of* Over*: Polysemy, Semantics, and the Structure of the Lexicon.* New York: Garland Publishing.

Butt, M. 1993. Hindi-Urdu infinitives as NPs. *South Asian Languages Review* **3**(1): 73-84.

Davison, A. 1991. Finiteness and case in Hindi-Urdu complements. *Proceedings of the Thirteenth South Asian Languages Roundtable.*

Dickey, M.E. & W. Vonk. 1997. Center-Embedded Structures in Dutch: An On-Line Study. Ms.

Dobrin, L. 2001. The morphological status of Arapesh plurals. Talk presented at the 2001 Linguistic Society of America Annual Meeting, San Francisco, CA.

Gibson, E. 2000. Dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, & W. O'Neil (eds.) *Image, Language, Brain: Papers from the First Mind Articulation Project Symposium.* Cambridge, MA: MIT Press.

Janda, R. & B. Joseph. 1986. One rule or many? Sanskrit reduplication as fragmented affixation. In B. Joseph (ed.) *OSUWPL* **34**: 84-107.

Janda, R. & B. Joseph. 1999. The modern Greek negator *mi(n)(-)* as a morphological constellation. In G. Babiniotis (ed.) *Greek Linguistics: Proceedings of the 3rd International Conference on Greek Linguistics*, 341-351. Athens: Elinika Gramata Press.

Just, M.A., P.A. Carpenter, J.D. Woolley. 1982. Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General* **111**(2): 228-238.

Kaan, E. & N. Vasić. 2000. Cross-serial dependencies in Dutch: Evidence for a discourse-based metric? Poster presented at CUNY Sentence Processing Conference, University of California, San Diego.

Lewis, R.L. 1996. Interference in short-term memory: The magical number two (or three) in sentence processing. *Journal of Psycholinguistic Research* **25**(1): 93-115.

Lewis, R.L. & M. Nakayama. 2001. Syntactic and positional similarity effects in the processing of Japanese embeddings. In M. Nakayama (ed.) *Sentence Processing in East Asian Languages*, 85-113. Stanford, CA.

Mahajan, A.K. 1990. The A/A-bar Distinction and Movement Theory. Unpublished doctoral dissertation, MIT, Cambridge, MA.

McGregor, R.S. 1995. *Outline of Hindi Grammar.* New York: OUP.

Miller, G.A. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* **63**: 81-97.

Mohanan, T. 1994. *Argument Structure in Hindi.* Stanford, CA: CSLI.

Schwenter, S.A. & S. Vasishth. 2000. Absolute and relative scalar particles in Spanish and Hindi. *Proceedings of the 26th Annual Meeting of the Berkeley Linguistics Society.* Berkeley, CA.

Väkunaa-Blum, R. 1989. Finnish word order as a set of syntactic constructions. In J. Powers & K. de Jong (eds.) *Proceedings of ESCOL 1988 (Eastern States Conference on Linguistics),* 500-511. Columbus, OH.

Vasishth, S. 2002. An Abductive Inference Based Model of Human Parsing. Unpublished Master's thesis, Ohio State University.

Wunderlich, D. 2000. Optimal case in Hindi. Ms.

Department of Linguistics
Ohio State University
222 Oxley Hall
Columbus, OH 43210

vasishth@ling.osu.edu; bjoseph@ling.osu.edu

Unidirectionality in Grammaticalization and Lexical Shift: The Case of English *Rather*

MATTHEW L. JUGE
Southwest Texas State University

1. Traditional models of grammaticalization

Traditional models of grammaticalization posit a process (or a set of processes) whereby words belonging to one lexical class shift to another lexical class considered to be more grammatical, as illustrated schematically in Figure 1, adapted from Hopper & Traugott 1993:108).

Figure 1—Typical one-dimensional model of verb grammaticalization
full verb > auxiliary > clitic > affix

In this paper I argue that the seemingly unrelated shift of English *rather* from comparative adverb to verb shows that such a model for grammaticalization is untenable.

2. Subprocesses of grammaticalization

Newmeyer argues (1998:252-259) that many of the processes that comprise grammaticalization are not unique to grammaticalization. In addition to the examples he provides, consider the following examples:

• phonological reduction	English *says* [sɛz] (cf. *pays* [pejz])
• inferential change	English *want* 'lack' > 'desire'
• metaphor	Gothic *wait* 'I have seen/I saw' > 'I know'
• harmony	English *picker-upper, rathering interesting*

With the integrity of grammaticalization as a coherent process in jeopardy, it now becomes even more important than ever to clearly identify what the components of grammaticalization are and how they relate to other kinds of linguistic changes. First, I will explore some earlier treatments of unidirectionality.

3. Previous arguments against unidirectionality

Other authors have argued that grammaticalization is not a unidirectional process, among them Newmeyer (1998), Beths (1999), and Janda (2001), while Ramat (1992) provides a number of counterexamples to the unidirectionality hypothesis.

Newmeyer argues quite strongly against the unidirectionality hypothesis. Part of his claim depends on the assertion—supported with extensive argumentation and examples (1998:232-59)—that grammaticalization is not a distinct process of its own. He further argues that, if it were a process in any meaningful sense of the term, it would, like aging, erosion, or evolution be uninterestingly unidirectional.

A key consideration that Newmeyer raises concerns the distinction between those grammaticalization researchers who treat unidirectionality essentially as part of the definition of grammaticalization and those who treat it as an empirical hypothesis or claim, in which case, he writes, '[T]he claim is false' (1998:261). Newmeyer suggests that for those who treat unidirectionality as a part of the definition of grammaticalization, there is no arguing, since definitions cannot be right or wrong.

Beths (1999) develops an argument against unidirectionality in the context of an in-depth analysis of a specific verb, claiming,

> The historical development of *dare*, from a semilexical verb in O[ld]E[nglish], to a modal auxiliary, a lexical verb, and a semimodal verb in Mod[ern]E[nglish] shows that grammaticalization can no longer be considered to be a unidirectional diachronic process of language change. (1999:1105)

Finally, Janda focuses on the discontinuous nature of language transmission and the fact that scholars have generally ignored this fact, paying attention instead to linguistic forms themselves.

In short, it now seems clear that, if there is not yet consensus, there is certainly a growing body of literature expressing the idea that the claim of unidirectionality is simply false.

But I wish to argue that all these researchers are missing an important point. In discussing unidirectionality and reversability, these scholars reveal that they are working with a one-dimensional model of lexical categories. The evidence shows that a multi-dimensional model is required to accommodate the various types of changes we see in the data, some of which are illustrated by the history of English *rather*, to which I now turn.

4. Lexical shift

In this section I would like to put aside questions of unidirectionality momentarily and address one of the similarities between grammaticalization and another kind of language change, namely **lexical shift**, called **transcategorization** by Ramat (2001). This is the shift of a linguistic unit, usually a word, from one lexical class to another. For example, in English, the participle *during* has shifted to the class of prepositions.

I want to explore these two (putative) phenomena together because grammaticalization frequently involves lexical shift. Thus, if we are to understand grammaticalization—and any unidirectionality it may show—then we must first understand lexical shift. In order to do this, I will perform a case study of the English word *rather*.

5. *Rather*

In modern English, the word *rather* has as one of its uses a verbal usage whereby it marks preference, as in (1).

(1) She hates listening about castles,/She rathers the battles. (http://homepage.tinet.ie/~portlawns/Pages/poems.htm)

This non-standard usage is further exemplified in the Appendix.

5.1 English

Historically *rather* is a comparative adverb meaning 'sooner, more quickly', as shown in the following example from around 1475 (Simpson & Weiner 1989):

(2) This worle..goth vnto decline, Rather or later to an endly fine.

Such a meaning is suitable in the right circumstances for expressing preference, as we see in (3):

(3) I would **sooner** cut off my right arm than be converted to right-handedness.... (from http://www.emf.net/~estephen/facts/lefthand.html)

In this situation, the expression of preference depends on two main components: (a) the real-world knowledge that cutting off one's right arm is a dispreferred action (even for a southpaw!) and (b) the tendency to, when possible, do what we like doing before doing things we do not like doing. We have here a case of pragmatic implicature (sometimes called **inferential change** in the grammaticalization literature) that allows the addressee to realize that the writer is so fond of left-handedness that, given the choice of losing it or their right arm, they would give up their right arm.

With *rather* the common pragmatic implicature of preference ultimately came to outshine the temporal meaning. This process was aided by the loss of the positive and superlative forms *rathe* and *rathest*, respectively, around the sixteenth century (contrast *sooner*, which the speaker can easily relate to the positive and superlative *soon* and *soonest*).

It is important at this point to note that because English lexical classes do not have characteristic morphological structures, *rather* does not advertise its status as an adverb, having essentially the same structure as the noun *lather* and the verb *gather*. This morphological ambiguity makes *rather* well-suited for reinterpretation as a word of virtually any lexical class, but I will focus exclusively on the verbal use and ignore other uses, including the adverbial degree usage (*It's rather cold in here*) and the compound prepositional usage (*I had a salad rather than a burger*).

Let us now consider the placement of *rather* relative to the verbal elements with which it frequently appears, namely *would* and *have*. Examples (4-10) show that English is flexible not just with *rather* but with the placement of adverbials in general.

(4) I **gladly** would've paid the extra $20....
(http://www.mexconnect.com/mex_/jrrimmig.html)

(5) I would **gladly** have paid twice the amount....
(http://www.pcesoft.com/)

(6) I would have **gladly** ransomed him....
(http://www.audiencemag.com/archives.html)

(7) We would have **rather** had a trial by ourselves.
(http://www.cnn.com/2001/LAW/01/29/combs.trial.02/index.html)

(8) ...I would have **rathered** a contractor do this.
(http://www.easton.ma.us/Directory/selectmen/minutes/min9-5-00.htm)

(9) I would **rather** a person (including entered contenstants) didn't win mr.net.art (http://bak.spc.org/iod/MisterNet.Art.html)

(10) I would **rather** a year of suffering than a week of living the mundane (http://www.pagans.org/~fyrecat/fyrecat/lyght.html)

This flexibility is not without pattern, however. Table 1, the results of a search of three patterns on the World Wide Web, shows that in cases in which *rather* appears with *would* and *have*, there is an overwhelming tendency for it to come between them.

Table 1—Frequency of patterns of rather *and the verbal complex*

I rather would have	80
I would rather have	19097
I would have rather	2990

This positioning of *rather* gives it the appearance of a second auxiliary marking preference.

The next piece of the conspiracy to make *rather* look like a verb concerns the behavior of *would.* Historically it is the past tense of *will*, which as a main verb used to mean 'to want'. It also was the past subjunctive of *will*, with the meaning 'would like', but this usage has faded into disuse:

(11) *I would one of those chocolate chip cookies.

The loss of this main verb use of *would* leaves the phrase *would rather* (11) looking like a phrase consisting of an auxiliary (*would*) and a main verb (*rather*).

(12) "...or would you rather coffee?"
(http://www.lisacerasoli.com/fanfic/authors/rebecca/memories/memories8.html)

Once *rather* was (re)interpreted as a verb, the door was opened to the analogical formation of additional verb forms, including past tense *rathered*, participial *rathered*, and third singular non-past *rathers*.

The exact chronology of some of these changes is difficult if not impossible to ascertain since some of the uses are non-standard and thus unlikely to be recorded, except on the internet (see §7 for further discussion of this point).

So, difficulties in dating certain changes notwithstanding, we may summarize the development of *rather* from adverb to main verb as follows:

chronology of *rather*

1. positive *rathe* and superlative *rathest* lost in modern English
2. loss of earlier temporal meaning of *rather*
3. un(der)specified morphological structure: cf. noun *lather,* verb *gather*
4. *rather* between *would* and a main verb—appearance of a second auxiliary
5. loss of *would* as subjunctive equivalent to *would like*
6. reinterpretation of *rather* as a verb
7. analogical past tense, past participle *rathered*, third singular non-past *rathers*

5.2 Brief comparison with Italian *piuttosto*

It may be useful at this point to explore why cases of the kind of shift shown by *rather* are not more common. To do this, I examine the potentially equivalent term in Italian, *piuttosto*, exemplified in (13).

(13) piuttosto andrei a chiedere l'elemosina '...I would rather go ask for alms.'
(http://www.theo.it/R&L/inmorte.html)

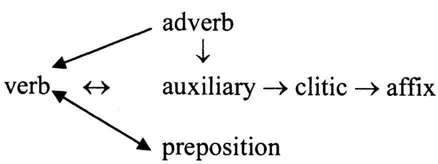
Like *rather*, *piuttosto* has the appearance of a comparative form (>*più* 'more' + *tosto* 'soon'), although since *tosto* is now literary, there is a certain opacity to the form. Again like *rather*, there is a separate root that now means 'soon', *presto*. Quite unlike rather, however, *piuttosto* does not resemble a verb (except that it ends in *-o* like first person singular indicative forms).

In Italian, the chance that an adverb will essentially accidentally resemble a verb, inflected or otherwise, is slim because verbs have a characteristic set of shapes, while in English this is not the case. In the end, a number of factors have to come together for the kind of shift shown by *rather* to take place. Thus we do not see such changes often. Since this type of change is one of the major kinds that contradict the appearance of unidirectionality and it is relatively infrequent, researchers have been led to believe that (apparent) unidirectionality is a significant phenomenon.

6. A unified model of lexical category change

The case of *rather*, then, suggests a model of language change that does not rely on a simple, one-dimensional view of language, but rather a conceptualization of lexical space with at least two dimensions, if not more (see Figure 2).

Figure 2—A two-dimensional model for category changes



On such a view of the relationships among lexical categories, the debate over unidirectionality takes on a less compelling character. At the same time, we are more likely to recognize patterns previously overlooked because researchers were only looking in a restricted set of places for their data.

7. Coda: Lowered standards and improved research

In the course of conducting the research for this paper, I was reminded of certain difficulties associated with researching non-standard language forms. Beyond the well-known obersver's paradox is the fact that in certain kinds of contexts, the very fact that these forms are non-standard excludes them from appearing. In this section I argue that the relatively untamed nature of the World Wide Web is a boon to the linguistic researcher working on some kinds of non-standard patterns.

While some might think of the internet as a place where important standards are being violated daily, in my research I found that not only is there a great degree of systematicity and order—how could it be otherwise if the medium is to be used for communication?—but that there is also the added benefit of increased likelihood of appearance of certain linguistic forms. The case of *rather* shows this well, for when we compare a corpus of materials that adhere fairly well to these standards, namely the SARA corpus (http://sara.natcorp.ox.ac.uk/lookup.html), we find no examples of verbal *rather*, while on the Internet, we find many examples.

7.1 Complications associated with 'lowered' standards

Along with the increased research possibilities of the unregulated nature of the internet come certain complications. Foremost among these is that when using the internet as a corpus, one must keep in mind that some tasks will require extra effort. One example is performing word frequency counts or distribution analyses; a search for *receive*, for instance, would be more successful if one also searched for *recieve* (and possibly other variants). Of course there is a good chance that additional variants may occur that the researcher will not be able to anticipate, but the same is true of research with an unfamiliar spoken language.

7.2 Directions for future research on the Internet

My experience in researching the distribution of verbal *rather* strongly suggests that use of the internet as a new kind of corpus holds a great deal of promise so long as researchers keep in mind, as they should in all research situations, the importance of context to data they obtain.

Appendix: Examples of verbal *rather*

Zionism was born in the late 1800s, the Jews did not want to leave their countires, no need to mention on how the Zionists were fought & attack, both politically & physically, by the Jewish peoples across europe, but thanks to the Holocaust that left the Jews to no choice but to immigrate to Russia, US, and England, you know like every one esle know that the bulk of the Jews **rathered** to go to russia than going to the "Promised land".	www.salam-shalom.net/ salam-shalom/arcjan27.htm
LBJ much **rathered** to fight a war against poverty and make his vision of a Great Society a reality, but performing in the shadow of an assassinated president did not afford him that luxury.	www.coe.uh.edu/ hypergroups/courses/ hist3322/0043.html
Webster noted for 2.2m I would have **rathered** a contractor do this.	www.easton.ma.us/Directory/ selectmen/minutes/min9-5-00.h
But I would have **rathered** a slightly better pc for slightly more...	www.linuxsa.org.au/mailing-lis 2001-01/429.html
"I think the players would have **rathered** to play on when the break came, but we have the chance when we go back to show what we can do again."	www.celticfc.co.uk/newsdesk/ latest_news/01-01/am_21_2009.htm
Erris themselves would have **rathered** to take the full Complement of points on offer but at least a point keeps them within touching distance of the leaders.	www.mayo-ireland.ie/WPeople WP9809/WP980909/Sport.htm

References

Aaron, Jane. 2001. *The Little, Brown Compact Handbook*. New York: Longman.

Beths, Frank. The history of *dare* and the status of unidirectionality. *Linguistics* **37-6**: 1069-1110.

Hopper, Paul, and Elizabeth Traugott. 1993. *Grammaticalization*. Cambridge: Cambridge Univerisity Press.

Janda, Richard D. 2001. Beyond "pathways" and "unidirectionality": On the discontinuity of language transmission and the counterability of grammaticalization. *Language Sciences* **23**: 265-340.

Ramat, Paolo. 1992. Thoughts on degrammaticalization. *Linguistics* **30**: 549-60.

______. 2001. Degrammaticalization or transcategorization? In Schaner-Wolles, Torino, Rosenberg, & Sellier (eds.) *Naturally! Linguistic studies in honour of W.U. Dressler presented on the occasion of his 60th birthday*, 393-401.

Newmeyer, Frederick J. 1998. *Lanuage Form and Language Function*. Cambridge: Massachusetts Institute of Technology.

Simpson, J. A. & E. S. C. Weiner, (eds.). 1989. *The Oxford English Dictionary*. Oxford: Oxford University Press.

Department of Modern Languages
214 Centennial Hall
601 University Dr.
Southwest Texas State University
San Marcos, TX 78666-4616

juge@modlang.swt.edu

Keeping and Losing Contrasts[1]

JOHN KINGSTON
University of Massachusetts, Amherst

0. Introduction

In this paper, I will first show that contrasts aren't licensed in particular contexts because they are necessarily realized with more salient cues in those contexts than in contexts where they aren't licensed. Next, I will argue that the releases of consonants, particularly non-continuants, are not packages of particularly salient acoustic information about the consonant's identity but instead only one of many kinds of acoustic information about how the utterance is segmented prosodically. Prosodic segmentation helps the listener find words in the stream of speech. Finally, I will show that segments are usually perceived to be different from their neighbors, i.e. to contrast with them, except when the target sound is C_1 in a VC_1C_2V string, which is instead often perceived to be the same as C_2, i.e. to assimilate to C_2. Both the general and specific effects of neighboring sounds phonetically explain where contrasts are kept or lost quite differently than the licensing by cue account.

1. Licensing by cue

Licensing by cue explains where phonological contrasts are kept or lost in terms of where they are or are not expressed by perceptually salient cues (Steriade 1995, 2000, 2001; Wilson 2001). The explanation rests on two incorrect assumptions. First, that contrasts are pronounced in the same way, and thus would be recognized by the same potential cues, in all contexts, but that these cues differ in salience or even audibility between contexts. And second, that the cues themselves differ in their inherent perceptual salience. The first assumption is incorrect because contrasts are typically pronounced differently in different contexts, and these differences may be designed to optimize conveying them in each context. The second assumption is incorrect because languages differ in

[1] I am grateful to the audiences at the 28th Meeting of the Berkeley Linguistic Society, the University of Massachusetts, Amherst, and the University of Pennsylvania for comments and questions that have substantially improved this paper. Maciej Baranowski, Gene Buckley, Ed Flemming, Mark Liberman, John McCarthy, Pawel Nowak, and Bob Rothstein deserve special thanks. All errors are mine.

which cues they select to convey the same contrast, and which cues they transform or transfer the contrast to in the course of sound change.

1.1. Salience differences between contexts and languages?

Kingston & Diehl (1994; also Keating (1984), Maddieson (this volume)) have shown that speakers pronounce the members of contrasts differently in different contexts, languages, and styles, and that as a result, the acoustic differences between these contrasting sounds aren't always the same. A familiar example of such variation across contexts is that of the [voice] contrast in English stops (Lisker 1986, Kingston & Diehl 1994):

(1) English stops contrasting for [voice] differ:

(a) Word-initially and pretonically in:
  (i) Aspiration

(b) Intervocalically and not pretonically in:

(i)	Closure duration:	/b,d,g/ < /p,t,k/
(ii)	Preceding vowel duration:	/b,d,g/ > /p,t,k/
(iii)	Closure voicing:	/b,d,g/ > /p,t,k/
(iv)	F_1 at vowel edge:	/b,d,g/ < /p,t,k/
(v)	F_0 at vowel edge:	/b,d,g/ < /p,t,k/

(c) Word- and syllable-finally in:
  (i) Preceding vowel duration
  (ii) Glottalization

(More detail is given for intervocalic stops because the perceptual value of these differences is discussed below.)

Laeufer (1992) documents particularly interesting differences and similarities across languages in the realization of the word-final [voice] contrast in English and French. Since Chen's 1970 survey, received wisdom had been that vowels lasted 50% or more longer before [+voice] than [-voice] stops in English but only 10-20% longer in French and other languages (also Mack (1982)). Laeufer showed instead that the vowel duration differences were much more similar to one another in these two languages so long as the stops are syllabified in the same way, which is difficult to achieve given their very different prosodies. When prosody re-syllabifies the word-final consonant into the following syllable's onset in either language, preceding vowels differ much less in duration than when the prosody leaves the stop in the preceding syllable's coda. Moreover, in French, final [+voice] and [-voice] obstruents are very often released into voiced and voiceless vocoids, respectively, and these vocoids behave like genuine vowels in permitting the final consonants to re-syllabify. Because these distinctive releases are so frequent, there are many more re-syllabified tokens in French than in English, and consequently preceding vowel durations differ less over all tokens in

French than in English. In other words, French speakers rely more on the distinctive consonant releases to convey a final [voice] contrast than English speakers do, but when they don't pronounce the stops with distinctive releases, differences in preceding vowel duration appear that are comparable in size to those observed in English. (Davis & Summers (1989) also show that vowel duration differences depend on syllabification in English, being much larger post-tonically, in words like *rabid* vs. *rapid*, than pre-tonically, in words like *rebel* (verb) vs. *repel*.)

Kingston & Diehl also argued that in some cases speakers choose different pronunciations to optimize conveying the contrast in each context that it occurs. The argument has three parts. First, all contrasts are conveyed by many, covarying acoustic differences. Second, many of these acoustic differences are produced by independently controlled articulations, not by some articulations occurring mechanically as a consequence of another controlled articulation. (As Kingston & Diehl defend this claim extensively, it won't be discussed further here.) Third, covarying acoustic differences often integrate perceptually so as to enhance the contrast.

Perceptual integration has been studied most thoroughly for the [voice] contrast in intervocalic stops, where it produces three perceptual properties: the low (acoustic) frequency property, the closure duration property, and the consonant:vowel duration ratio.

Low Frequency. Kingston & Diehl (1995) and Kingston, *et al.* (submitted) found that two stops were easy to discriminate when one had low F_1 or F_0 at the edges of flanking vowels and voicing lasted well into the stop closure and the other had high F_1 or F_0 at flanking vowel edges and voicing stopped early in the stop closure. Stops with opposite combinations of these two properties, low F_1 or F_0 at vowel edge and short-lasting voicing vs. high F_1 or F_0 at vowel edge and long-lasting voicing, were much harder to discriminate, even though the physical size of the differences in both properties was the same for both combinations of values. In the more discriminable pair of stops, energy is either concentrated at low frequencies in and near the stop closure or it isn't (see also Stevens & Blumstein, 1981). The combinations of properties in the less discriminable pair cancel out each other's effects on the concentration of energy at low frequencies.

Closure Duration. Parker, *et al.* (1986), Kingston, *et al.* (1990), and Kingston & Diehl (1995) showed that when the stop closure contains voicing, i.e. low frequency periodic energy, listeners are more likely to label an intervocalic stop with a particular closure duration as [+voice] or short (recall that [+voice] stops have shorter closure durations) than when it contains no voicing. However, they only do so as long as F_1 is low at the flanking vowels' edges (low F_0 at the vowels' edges doesn't have the same effect). Voicing and low F_1 together make the closure more continuous spectrally with the flanking vowels at low frequencies and thereby reduce the perceived interruption of the vowels by the closure.

Consonant:Vowel Duration Ratio. Kluender, *et al.* (1988; see also Kohler (1979), Port & Dalby (1982), and Fowler (1992)) showed that listeners judge a particular closure duration to be [+voice] or short more often if the preceding vowel is long than if it's short. That listeners judge closure duration in contrast to the perceived duration of the preceding vowel suggests that this property is the ratio of these two durations: small values of this ratio induce [+voice] judgments, large values [-voice] judgments.

In each of these cases, the combinations of properties that enhance contrasts or increase [+voice] over [-voice] judgments are those observed in naturally produced stops contrasting for this feature, as laid out in (1). However, listeners' behavior cannot be due entirely their having (over-)learned the combinations of acoustic properties that do occur because all the same perceptual interactions have also been obtained with non-speech analogues of vowel-stop-vowel stimuli, which would not evoke listeners' experience with speech sounds. These interactions or perceptual properties arise from audition not experience.

Moreover, integration's pervasiveness shows that listeners often don't hear the cues as individual properties of speech sounds. Instead, they hear perceptual properties like the low frequency property that arise from the integration of these cues. Picking out one or another cue as essential to conveying a contrast misrepresents the listener's perceptual experience as well as the nature of the perceptual information they find in the signal.

So far, I have shown that the first assumption of the licensing by cue proposal is wrong: that contrasts are pronounced in the same way in all contexts or languages. Rather than being the same, the pronunciations of contrasts differ systematically between contexts and languages. Some of these differences, furthermore, produce combinations of acoustic properties that enhance contrasts. Finally, if speakers can choose different pronunciations in different contexts to improve the distinctiveness of a contrast there, then the contexts themselves don't differ intrinsically in how well contrasts can be realized in them.

1.2. Salience differences between cues?

Do the cues themselves differ in intrinsic perceptual salience? Sound changes in which a contrast is transferred to one of its cues naturally test the relative salience of cues as one would expect the contrast to be transferred to its most salient cue. However, the dis-integration of final [voice] contrasts shows instead that the contrast can be transferred to any of its cues.

The Friulian forms in (2) show the transfer of the [voice] contrast to a vowel duration difference in preceding vowels (Baroni & Vanelli 2000). The stems in (a-d) alternate between a final [+voice] obstruent when followed by a vowel and a [-voice] consonant when word-final. Vowel lengthening accompanies devoicing. The non-alternating stems in (e-h) show that nothing happens when the stem ends in [-voice] obstruent, i.e. when its pronunciation is the same finally as before a vowel. The lengthened vowel before a devoiced obstruent in a form such as *ˈla:t* 'gone (m.)' is more than twice as long as the vowel in *ˈlat* 'milk', while one before a [+voice] consonant in a form such as *ˈlade* 'gone (f.)' is only half again longer. The lengthened vowels are also more peripheral, their F_0 peaks occur earlier and F_0 falls more across the vowel, and the following stop closure is shorter than in a comparable non-alternating form. The [voice] contrast has not just been transferred to vowel length but the longer vowel duration that otherwise occurs before [+voice] obstruents in this language has been both greatly exaggerated and augmented by other phonetic differences that enhance it.

(2) [voice] and vowel length alternations in Friulian

a. 'gone (m.)'	ˈla:t	ˈlade	'gone (f.)'
b. 'fire'	ˈfu:k	fogoˈlar	'fire place'
c. 'weight'	ˈpe:s	peˈza	'to weigh'
d. 'snow'	ˈne:f	neveˈa	'to snow'
e. 'milk'	ˈlat	laˈta	'to breast-feed'
f. 'piece'	ˈtɔk	tuˈkut	'little piece'
g. 'pass'	ˈpas	paˈsa	'to pass'
h. 'slap'	paˈtaf	pataˈfa	'to slap'

In (3) we can see that in the same context where Friulian transferred the [voice] contrast to vowel length, Polish has instead transferred it to vowel height (Steele 1973, Stieber 1973, Gussman 1980, Carlton 1990, Buckley 2001). The stems in (a-d) end in a [+voice] consonant before a vowel, which devoices finally; devoicing is accompanied by raising of the mid vowel [o] to high [u]. The non-alternating stems in (e-h) have a final [-voice] consonant before a vowel as well as finally, and their [o] remains unraised. The transfer in this case reinterprets the low F_1 that would ordinarily precede a [+voice] consonant as a high vowel, which has a lower F_1 value than a mid vowel.

(3) [voice] and [high] alternations in Polish

	Nom. Sg.	Nom. Pl.
a. 'ice'	l[ut]	l[od]y
b. 'corner'	r[uk]	r[og]i
c. 'beans'	b[up]	b[ob]y
d. 'knife'	n[uʃ]	n[oʒ]e
e. 'flight'	l[ot]	l[ot]y
f. 'juice'	s[ok]	s[ok]i
g. 'peasant'	chł[op]	chł[opʲ]y
h. 'basket'	k[oʃ]	k[oʃ]e

This case is complicated in two ways, however. First, vowel height also alternates in stems that end in non-nasal sonorants (4a-e), which don't devoice finally, although not in those that end in nasal sonorants (5a-b), e.g.:

(4) [high] without [voice] alternations in non-nasal sonorant-final Polish stems

a. 'peace'	pok[u]j (gen. sg.)	pok[o]ju (nom. sg.)
b. 'cow'	kr[u]w (gen. pl.)	kr[o]w (nom. sg.)
c. 'moth'	m[u]l (gen. sg.)	m[o]la (nom. sg.)
d. 'barn'	stod[u]ł (gen. pl.)	stod[o]ł (nom. sg.)
e. 'time'	p[u]r (gen. pl.)	p[o]ra (nom. sg.)

(5) Neither [high] nor [voice] alternations in nasal sonorant-final Polish stems

a. 'home'	d[o]m (nom. sg.)	d[o]mu (gen. sg.)
b. 'side'	str[o]n (gen. pl.)	str[o]na (nom. sg.)

Raising before non-nasal sonorants is unsurprising as F_1 is low at the edges of vowels before any voiced consonant, whether it's a sonorant or an obstruent. The failure to raise before nasal sonorants is equally unsurprising as F_1 is instead raised by nasalization, and high nasalized vowels often lower to mid (Maeda, 1993; Hajek, 1997).

The second complication is at first glance more serious, however. (The following discussion is based equally on Stieber (1971) and Buckley (2001).) The alternants in (3-5) where the stem-final consonant is now word-final originally ended in weak yers. These yers were lost with compensatory lengthening of the stem vowel before final voiced consonants (sonorants as well as obstruents) around 1000 CE, e.g. 'ice' *lod-ɨ̆* (nom. sg.) > *lo:d* vs *lod-y* (nom. pl.). (Present-day forms are used for illustration.) Compensatory lengthening was blocked by a final voiceless consonant, e.g. 'flight' *lot-ɨ̆* (nom. sg.) > *lot* vs. *lot-y* (nom. pl.). By the late 14th century, the lengthened mid vowels were higher and more peripheral than the unlengthened ones: *lọ:d* vs *lɔd-y* and 'flight' *lɔt* and *lɔt-y*. Final obstruents devoiced about the same time or perhaps somewhat earlier in the 14th century: *lọ:d* > *lọ:t* vs *lɔd-y* and 'flight' *lɔt* and *lɔt-y*. Vowel quantity differences only began to disappear more than a century later, after 1450 CE: *lọ:t* > *lọt* vs *lɔd-y* and 'flight' *lɔt* and *lɔt-y*, and later still, certainly after 1600 CE, /ọ/ and /u/ merged to /u/ and /ɔ/ raised to /o/: *lọt* > *lut* vs *lɔd-y* > *lod-y* and 'flight' *lɔt* > *lot* and *lɔt -y* > *lot-y*. Like Friulian, Polish first develops a longer vowel before a [+voice] consonant, albeit because a [+voice] consonant is compatible with compensatory lengthening, not via transfer of the [voice] contrast. Moreover, raising appears to accompany length.

Nonetheless, the higher vowels emerged before consonants whose voicing would lower F_1 in the preceding vowel, and raising coincided with devoicing in the 14th century. Compensatory lengthening may have been the first step in the development of the present-day alternation, but the following consonant's voicing remains the most likely cause of raising. By itself, lengthening would instead have lowered vowels, as lower vowels are inherently longer than higher ones. Because word-final obstruents devoiced at the same time, raising immediately came to share the burden of the [voice] contrast with vowel length. Once quantity differences were lost, only vowel height differences remained to convey the contrast. Later mergers led to the present-day high vowel in the bare stem alternants in (3-5).

The last example of dis-integration of a final [voice] contrast underlies a sound change taking place in southern American English. Thomas (2000) and Moreton (2002) show that the offglide in the diphthongs /aɪ, aʊ, ɔɪ, ɛɪ/ before [-voice] obstruents has lower F_1 values and more extreme F_2 values, higher in the fronting diphthongs / aɪ, ɔɪ, ɛɪ/ and lower in the backing diphthong /aʊ/, than in other contexts. [Summers (1987) observed more extreme F_1 values and larger jaw movements in /a, æ/ before [-voice] than [+voice] obstruents; both F_1 value and jaw height also changed faster before [-voice] than [+voice] obstruents.] Moreton (2002) also found that listeners identify final stops as [-voice] more often when F_1 is low and F_2 is high in the offglide of /aɪ/. Moreton's speakers and listeners came from northern as well as southern parts of the United States, so more extreme pronunciations of these diphthongs before [-voice] obstruents than elsewhere aren't just found in the south. However, in much of the south, this difference has been phonologized to the extent that /aɪ/ is still pronounced as a diphthong before [-voice] obstruents but has become a low fronted monophthong [a̟] elsewhere. Although following obstruents contrasting for [voice] usually remain phonetically different, before flaps the difference between a diphthong and monophthong alone conveys the [voice] contrast.

In each of these three examples, what was only one of many phonetic cues to the [voice] contrast, vowel duration, F_1, or diphthongization, is now the only means of distinguishing morphemes that once ended in [+voice] vs. [-voice] consonants. This diversity suggests that none of cues is any more salient than the others, contrary to the second assumption of the licensing by cue proposal that cues differ intrinsically in salience. In short, no cue is privileged by greater perceptual salience.

1.3. Optimal pronunciations

Before ending this critique of licensing by cue, it is useful to take up a case of dis-integration that does not involve the [voice] contrast. Besides showing the generality of dis-integration, this case also shows that speakers exert themselves to convey contrasts in ways that are entirely unexpected if they couldn't optimize their pronunciations to ensure that contrasts are conveyed.

In languages where nasalization is not contrastive in vowels, soft palate height covaries directly with tongue height, causing lower vowels to be more nasalized than higher vowels (see Kingston (1991) for a review of the evidence). English is such a language, and for English listeners nasalization separates perceptually from height in higher vowels, but integrates with height in lower vowels: greater acoustic nasalization makes a lower vowel sound lower rather than more nasalized (Kingston & Macmillan 1995, Macmillan *et al.* 1999). More nasal coupling is needed to get a lower vowel to sound nasalized (Maeda, 1993).

The perceptual integration of nasalization into the percept of vowel lowness and the need for more nasal coupling are unexpected given the ubiquity of low vowels in nasal vowel inventories. In the expanded version of UPSID with 451 languages (Maddieson & Precoda 1992) are 101 that contrast nasal with oral vowels. 49 of these languages have the same number of nasal as oral vowels, but in the remaining 52, one or more nasal vowels corresponding to oral vowels are missing. The examples in (6) show the patterns and frequencies of gaps in the nasal vowel inventories of these languages:

(6) Typology of gaps in inventories of nasal vowels corresponding to oral vowels

a. *Headless*, missing high nasal vowels (21 languages), e.g., Amuzgo:

Oral			Nasal		
i		u			
e		o	ẽ		õ
æ	a	ɒ	æ̃	ã	ɒ̃

b. *Gutless*, missing mid nasal vowels (39 languages), e.g., Senadi:

Oral			Nasal		
i		u	ĩ		ũ
ẹ		ọ			
ɛ		ɔ	ɛ̃		ɔ̃
	a			ã	

c. *Footless*, missing low nasal vowels (6 languages), e.g., Chatino:

Oral			Nasal		
i		u	ĩ		ũ
e		o	ẽ		õ
	a				

The gutless type (6b), where mid nasal vowels are missing, outnumbers the headless type (6a), where high nasal vowels are missing, by nearly 2:1. The footless type (6c) is comparatively very rare. How can nasal vowel inventories nearly always include a low nasal vowel if nasalization is integrated into the percept of vowel lowness in lower vowels?

Speakers must lower the soft palate more in lower than higher vowels to ensure that they're perceived as nasal as well as low. Fortuitously, the vowel also sounds lower because some of that nasalization is still integrated into the percept of vowel lowness. Low vowels' inherently greater duration may also help make the spectral modifications caused by nasalization easier to detect (Hajek 1997; see also Krakow, *et al.* (1988) and Whalen & Beddor (1989)).

Otherwise, gutless nasal vowel inventories are most common because mid nasal vowels either lower, if nasalization is perceived as lowness, or raise, to improve the perceptual separability of nasalization. Headless nasal vowel inventories are next most common because high vowels are inherently short and may thus sometimes be too brief for nasalization to be perceived reliably. Nasalization either doesn't become contrastive in high vowels or integration lowers them to mid.

In this case, the most common instance of a type, a low nasal vowel, is the one that requires the most extreme articulation, the greatest soft palate lowering, to ensure that its acoustic effect is dis-integrated perceptually from other properties of that sound. Speakers must be altruists.

2. What are releases good for?

2.1. Do releases package acoustic information about segment identity?

Much of the concrete discussion of licensing by cue has turned on the presence and perceptual value of an audible release in consonants. The release's acoustics are thought to bear much of the burden of conveying the consonants' phonation, place, and manner contrasts. I bear no little responsibility for this emphasis, having argued in Kingston (1985, 1990) that releases had these virtues. Here, I first review the evidence and arguments that led me and others to emphasize the importance of releases for segment identification and then turn to the perceptual evidence which shows that listeners don't after all rely on the acoustics of consonant releases for this purpose. Finally, I review evidence which suggests that releases, along with much other allophonic variation, are probably perceptually valuable because they aid listeners in segmenting the utterance prosodically, which in turn helps them find words in the stream of speech.

My original argument was that the release of a consonant constriction "packages" the consonant's values for phonation, manner, and place of articulation into a brief, salient acoustic event (see also Stevens & Keyser (1989) and Liu (1996)). This event might be salient because neurons in the VIIIth (auditory) nerve fire robustly for a brief time when signal energy rises abruptly; subsequently, their firing rate quickly drops off (Delgutte & Kiang 1984a,b; Delgutte 1996; Silverman 1997; Wright 1999, 2001). The representation of spectral energy distributions in the peripheral auditory system as well as higher up should therefore be best at the release, particularly of a non-continuant consonant.

Stevens & Keyser (1989) showed how the most common contrasts for manner and place of articulation in the world's languages correspond to large differences in how spectral energy distributions change at consonant releases. Earlier, Stevens & Blumstein (1978; Blumstein & Stevens 1979, 1980) had shown that stops could be automatically classified for place of articulation using just gross features of spectra calculated across a short interval (25.6 ms) beginning with the release, and Kewley-Port, *et al.* (1983) and Lahiri, *et al.* (1984) had been able to classify stops' place of articulation automatically using changes in the spectral distribution of energy across a brief interval between the release and a point early in the following vowel. Released final stops are also identified much better than unreleased ones (Malécot, 1958), and in a cluster of stops between vowels, only the released one is heard if the closure duration is short enough that the listener thinks only a single stop has been pronounced (Fujimura, *et al.* 1978; Repp 1978, 1983; Ohala 1990).

Nonetheless, releases don't convey the most salient information for either place or manner perception. When release spectra are pitted against formant transitions in place perception, formant transitions nearly always determine what place of articulation is perceived (Walley & Carrell 1983, Smits *et al.* 1996a,b). The rate at which energy rises at signal onset, a.k.a. "rise time", should directly affect VIIIth nerve firing rate, yet rise time contributes little or nothing to perceiving the [continuant] contrast between affricates vs. fricatives—listeners rely on noise duration instead (Kluender & Walsh 1992)—nor to the combined [continuant, sonorant] contrast between stops vs. glides—listeners rely on transition duration instead (Diehl & Walsh 1989, Walsh 1991).

The original idea that releases packaged segmental information assumed, like the licensing by cue proposal it inspired, that it wasn't sounds' pronunciations that varied systematically with context but instead only the audibility of cues to their identity. Consonants are always released, but the release is only audible in some contexts, specifically when the following sound has a much more open articulation than the consonant, ideally a vowel or a non-nasal sonorant. Before such a more open articulation, energy would be audible across the spectrum, and any effects of glottal or oral articulations on its distribution could be detected. As consonants preceding a vowel or non-nasal sonorant are most likely to be syllabified into the onset with that sound, releases are much more likely to be audible and informative about segment identity in syllable onsets than codas (Kingston 1985; Lombardi 1991, 1995).

However, speakers actually articulate consonants differently in onsets than they do in codas. Soft palate and lip movements are closely synchronized in [m] in onsets, but soft palate movement begins long before lip movement in [m] in codas (Krakow 1989, 1993). Similarly, tongue dorsum and tip movements are closely synchronized in [l]s in onsets, but tongue dorsum movement begins long before tip movement in [l]s in codas (Sproat & Fujimura 1983). Kelso, *et al.* (1986) had speakers repeat either [ip] or [pi] and then speed up repeatedly. In repetitions of [pi], peak glottal opening was delayed relative to lip opening onset by a constant proportion at all speaking rates, but in repetitions of [ip], peak opening abruptly shifted from coinciding with lip opening onset to a delay equal to that observed in [pi] when the speaking rate exceeded about 4 syllables/second. Although this shift was apparently involuntary, it shows that glottal opening is timed differently relative to lip opening for a voiceless stop in a coda from one in an onset. Tuller & Kelso (1991) show that these differences affect listeners' percept of the stop's syllabification. Finally, Smith (2002) shows that the pronunciation of segments in onsets and other *prominent* positions may even neutralize contrasts in order to increase the position's salience. For example, in Chamicuro, the glottal consonants /ʔ, h/ are prohibited from occurring in syllable onsets; that is, the contrast between consonants with and without an oral constriction is neutralized in favor of those with an oral constriction in this position. These results all show that speakers strive to pronounce onsets differently from codas; consonants don't merely sound different in the two syllable positions.

2.2. Do releases signal prosodic boundaries instead?

Even if listeners don't rely on the acoustic information packaged in releases to identify segments, do they use releases for any other purpose? Because a consonant is usually only released audibly when it precedes a segment with a more open articulation and in that case is usually syllabified with that segment, the release may convey that syllabification to the listener. The absence of an audible release would instead convey that the consonant is not syllabified with the following segment. If this hypothesis is correct, then releases are simply one of a large number of allophonic properties that convey the grouping of segments in syllables, as well into higher-level constituents in the prosodic hierarchy.

There is also considerable evidence that the allophones which occur at the beginning of prosodic constituents interrupt the signal more than those which occur inside prosodic constituents (English: Pierrehumbert & Talkin 1992, Turk 1993, Dilley, *et al.* 1996; French: Fougeron & Keating 1997, Fougeron 2000; and Korean: Silva 1992, Jun 1993, Cho & Keating 2000). A greater interruption separates the next prosodic constituent more from the preceding one. An audible release may not only link a consonant prosodically to the following segment, but also increase the interruption of the signal and thereby sharpen the separation of prosodic constituents.

Although the listener would benefit from the marking of the beginning of any prosodic constituent by releases and other allophones, these phonetic events could be particularly useful for finding the beginnings of phonological words in the stream of speech. Initial allophones which aren't word-initial inhibit word spotting in longer non-word strings in Dutch (McQueen 1998) and English (Kirk 2000, 2001). For example, *wine* is harder to spot following the aspirated [kʰ] in [vukʰwaɪn] than the unaspirated [k] in [vukwaɪn] and *rock* is harder to spot following the retroflexed, affricated [ɖʐ̩] in [vuɖʐ̩ɹak] than the alveolar stop [d] in [vudɹak]. Kirk argued that the initial allophones [kʰ] and [ɖʐ̩] lead listeners to syllabify the end of the preceding residue with the target word, and the resulting mismatch between syllable and word boundaries makes the word harder to spot than when the residue ends instead with a non-initial allophone. Both allophonic variations are furthermore differences in the presence vs absence of a particular kind of release of the final consonant of the residue.

2.3. Summary

In this section of the paper, I have tried to show that releases aren't packages of acoustic information about segment identity, at least not ones that listeners rely on much. Instead, releases are one of a large number of allophonic variants that signal the beginning of prosodic constituents and thus aid the listener in parsing the stream of speech into words and higher level prosodic constituents. Word-spotting experiments show that listeners do indeed use releases for this purpose.

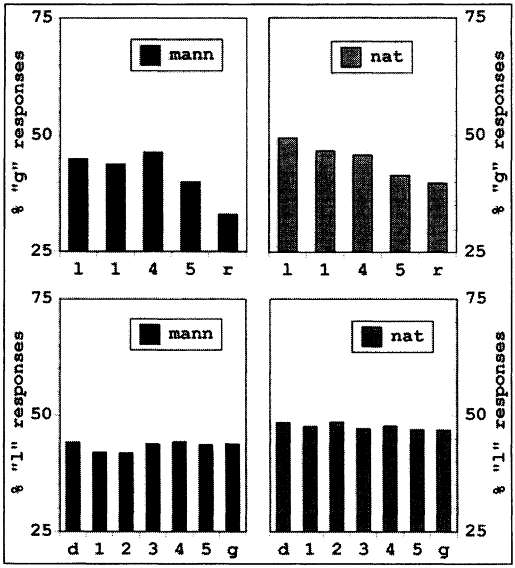
3. Listening in context

3.1. Perceptual contrast between adjacent segments

It is a truism that in everyday life, speech sounds are pronounced and must be recognized next to other speech sounds. Both a very general and a very specific characteristic of neighboring sounds' effects on the perception of the current sound phonetically explain the distribution of contrast and neutralization and thus replace the licensing by cue account of this distribution that I argued against in §1 above.

Nearey (1990, 1995, 1997, in press), Smits (2001a,b), and Benkí (this volume) have shown that listeners extract segments from the speech stream even though perceptually relevant information often comes from acoustic intervals corresponding to adjacent segments and neighboring segments mutually influence one another's recognition. The effect of a neighboring segment is typically contrastive: next to a segment which is at one extreme of some acoustic dimension, a segment which is intermediate along that dimension will sound like it's at the opposite extreme. For example, an intermediate member of a [d-g] continuum synthesized by incrementally varying F_3 onset frequency is more likely to be labeled heard as the continuum's low endpoint "g" after [l], whose F_3 is high, than after [ɹ], whose F_3 is low (Mann 1980). Lotto & Kluender (1998) obtained the same "g" bias when the liquid was replaced by a pure tone whose frequency corresponded to the liquid's F_3, which shows the contrast arises from audition rather than compensation for coarticulation (cf. Fowler, *et al.* 1990, 1999). Gouskova & Kingston (2002, unpublished data) recently replicated Mann's results using two liquid continua as contexts, one synthesized with Mann's parameter values (*mann* stimuli) and the other with values based on Stevens's 1998 description and modeling of naturally produced liquids' acoustics (*nat* stimuli). Listeners identified the liquid as well as the stop. The top two panels in Figure (7) show that "g" responses drop by about 10% as the liquid goes from [l] to [ɹ], and the bottom two panels show that the percentage of "l" responses changes little as the stop goes from [d] to [g]. The effects and their sizes don't differ between the *mann* and the *nat* stimuli. The progressive effect of the liquid on stop judgments is contrastive, but there is no regressive effect of the stop on liquid judgments.

(7) Frequency of "g" responses as a function of the preceding liquid (top) and "l" responses as a function of the following stop (bottom), for *mann* (left) vs *nat* (right) liquids.



Other context effects are also contrastive. An intermediate member of a [t-k] continuum, also synthesized by varying F_3 onset frequency, is more likely to be labeled the high endpoint "t", after [ʃ], whose noise is concentrated at relatively low frequencies, than after [s], whose noise is concentrated at relatively high frequencies (Mann & Repp 1981; Repp & Mann 1981, 1982). Consonants' place also contrastively shifts the labeling of neighboring vowels' backness. Listeners label intermediate members of an [ɪ-ʊ] continuum more often as the high F_2 endpoint "ɪ" between labiovelar glides [w_w], whose F_2 values are low, than between palatal glides [j_j], whose F_2 values are high. Similarly, listeners label intermediate members of a steady-state [i-u] continuum more often as the high F_2 endpoint "i" between the labials [f_p], whose energy is concentrated at relatively low frequencies, than between the coronals [s_t], whose energy is instead concentrated at high frequencies (Ohala *et al.* 1978). Listeners also label a steady-state [i-u] continuum's intermediate members as "i" more often before labial [b] than coronal [d], whether they actually hear the stop or restore it (Ohala & Feder 1995; Bradlow & Kingston 1990). Finally, Holt *et al.* (2000) report that listeners label intermediate members of an [ɛ-ʌ] continuum more often as "ɛ" between labials [b_b] than coronals [d_d], as well as FM glides mimicking the F_2 and F_3 transitions to and from [b] vs. [d] and pure tones matching the F_2 and F_3 onset frequencies next to [b] vs. [d]. In all these cases, the listeners is more likely to respond with the category corresponding to the continuum's high frequency endpoint when the neighboring segment concentrates energy at relatively low frequencies than when it instead concentrates it at relatively high frequencies.

3.2. Perceptual assimilation as well as contrast

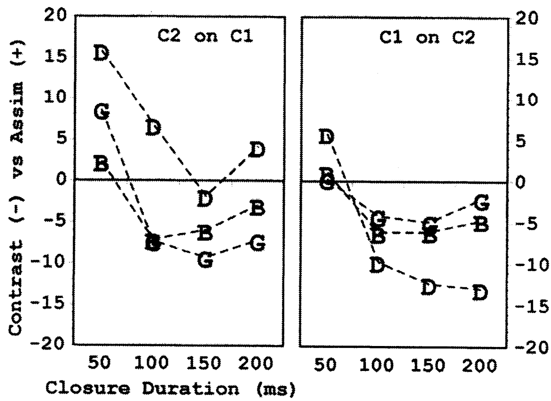
Although context effects are nearly always contrastive, like these examples, in one circumstance they are consistently assimilatory. In stop-stop clusters between vowels (VC_1C_2V; Repp 1978, 1983; Ohala 1990), $C_1$is perceived as having the same place of articulation as C_2 when the closure duration is short enough (relative to the range of closure durations heard) to be just one consonant rather than two. Alternatively, the listener hears only C_2 when the closure duration is short. Even if C_1 isn't heard as a separate segment at these short closure durations, its presence and distinct place of articulation nonetheless contrastively shifts C_2 labeling. At closure durations long enough to be two consonants rather than just one, labeling is mutually contrastive, although the contrastive effect of C_1 on C_2 remains stronger than the reverse. Shinya & Kingston (2002, unpublished data) replicated Repp (1983) using a 7-step [b-d-g] continuum synthesized in both VC_1 and C_2V by varying only F_2 and F_3 offset and onset frequencies. VC_1 and C_2V were mirror images in that no burst preceded formant onset in C_2V. VC_1 and C_2V were separated by silent gaps lasting 50, 100, 150, or 200 ms. Listeners labeled the stops as "b", "d", or "g" if they heard just one, and labeled both C_1 and C_2 as one of these three alternatives if they instead heard two. Assimilation and contrast were quantified using the formulae like those in (8) (following Repp 1983):

(8) a. Assimilation: [% "b" / V_.bV] - [% "b" / V_.{d,g}V] > 0
b. Contrast: [% "b" / V_.bV] - [% "b" / V_.{d,g}V] < 0

These two formulae assess the regressive effect of C_2 on the entire range of C_1 stimuli: if more "b" responses are given before [b] than before [d] or [g], then C_1 has assimilated to C_2, but if "b" responses before [d] or [g] instead outnumber those before [b], then C_1 instead contrasts with C_2. Scores were calculated for "b" responses before the best [b] vs. the best [d] and [g], and "b" responses before the best [b] stimulus vs. all other stimuli. The resulting scores differed very little, so only those calculated the second way are presented here. Entirely parallel calculations were carried out for assessing the progressive effect of C_1 on C_2 and for the other places of articulation. These scores are displayed as a function of closure duration for the regressive effects of C_2 on C_1 in the left panel of Figure (9), and for the progressive effects of C_1 on C_2 in the right panel. At the shortest closure duration, a listener is more than 15% more likely to label C_1 as "d" when C_2 is the best [d] exemplar than when C_2 is any other stimulus, and a bit less than 10% more likely to label C_1 as "g" when C_2 is the best [g] exemplar than when C_2 is any other stimulus. "b" responses to C_1 are at most just slightly more likely before the best [b] exemplar than other stimuli at this closure duration. At all longer closure durations, both "b" and "g" responses to C_1 are instead more likely

(5-10%) when C_2 is any other stimulus than the best [b] and [g] exemplars, respectively. "d" responses, on the other hand, remain more likely before the best [d] exemplar than other stimuli for all closure durations except 150 ms, where "d" responses are equally likely before all C_2s.

(9) Regressive (left) and progressive (right) assimilation (+ %) vs. contrast (- %) in VC1C2V for "b", "d", and "g" responses as a function of closure duration.



Listeners are somewhat more likely (5%) to label C_2 as "d" following the best [d] exemplar at the shortest closure duration, but at all other closure durations, they are much more likely (up to 15%) to label C_2 as "d" following any other stimulus. Neither "b" nor "g" responses differ as a function of C_1 at the shortest closure duration and both become somewhat more likely (5%) following any stimulus other than the best exemplars of [b] and [g], respectively, at longer closure durations.

Coronal C_1 percepts are assimilative or neutral across closure durations, but non-coronal C_1 percepts are instead contrastive, except for dorsal C_1 percepts at the shortest duration. C_2 both weakens the percept that C_1 has occurred and that C_1 is different from C_2 when C_1 is perceived to have occurred. Both effects are stronger for coronals than non-coronals. The phonological consequences of this weakening are that C_1 is likely to delete, assimilate to C_2, or neutralize for place of articulation, especially if C_1 is coronal. On the other hand, coronal C_2 percepts are neutral or strongly contrastive, and non-coronal C_2 percepts are neutral or weakly contrastive. C_1 strengthens the percept that C_2 has occurred and that it's different from C_1. Again, both effects are stronger for coronals than non-coronals. C_2 is therefore unlikely to delete, assimilate to C_1, or neutralize for place, particularly if it's a coronal. These perceptual asymmetries explain both the much greater susceptibility of coda than onset consonants to loss, assimilation, and neutralization and the particular susceptibility of coronal consonants in codas to these losses of contrast.

another and no burst preceded C_2V. It is only the position of the consonants in the coda that puts them at perceptual disadvantage, not their lack of any place information or acoustic substance that the consonants in the onset possess.

These preliminary results are very encouraging but they leave a number of questions unanswered. First and most basic, why are progressive effects contrastive while regressive effects are assimilative? The onset consonant may dominate simply because it's more recent, but this hypothesis hasn't been tested yet. Second, would similar perceptual asymmetries be observed between onsets and codas for manner and phonation contrasts, which are also lost in codas but kept in onsets? The necessary experiments simply haven't been done. Third, how can this explanation of the distribution of contrast and neutralization be generalized to word-final position, where contrasts are also often lost? In texts from 20 disparate languages, Janda (1979) found that a word is far more likely to be followed by a word beginning with a consonant than one beginning with a vowel. He argued that a following word boundary and a following consonant are so often common environments for phonological constraints or processes because # is in fact usually C. The conditions for perceptual assimilation and contrast between coda and onset consonants therefore usually exist at the ends of words as well as inside them.

4. Summary

In this paper, I have tried to show three things. First, contrasts aren't licensed by cue but instead by the adjustments and effort speakers make to ensure their differences are perceived. Second, listeners don't rely on releases as packages of salient acoustic information for identifying segments but instead for segmenting the signal prosodically, particularly into prosodic words. Third, a sound is most often heard as different from its neighbors, except when it's a consonant before another consonant, in which case it's instead likely to be heard as the same as the following consonant, while the following consonant is likely to be heard as different from the preceding one. This perceptual asymmetry may explain why contrasts are typically lost in codas but kept in onsets. As nearly all the neutralizations that licensing by cue purports to explain occur in codas, this alternative explanation covers the ground equally well. Furthermore, it depends on a quite general and well-established perceptual interaction between neighboring segments.

References

Baroni, M., & Vanelli, L. 2000. The relationship between vowel length and consonantal voicing in Friulian, In L. Repetti (ed.), *Phonological Theory and the Dialects of Italy*, 13-44. Amsterdam, Philadelphia: John Benjamins Pub. Co.

Benkí, J. R. this volume. Effects of signal-independent factors in speech perception, *Proc. 28th Ann. Mtg. Berkeley Ling. Soc.*

Blumstein, S. E., & Stevens, K. N. 1979. Acoustic invariance in speech production: Evidence from the measurement of the spectral characteristics of stop consonants, *J. Acoust. Soc. Am.*, **66**: 1001-1017.

Blumstein, S. E., & Stevens, K. N. 1980. Perceptual invariance and onset spectra for stop consonants in different vowel environments, *J. Acoust. Soc. Am.*, **67**: 648-662.

Bradlow, A. R., & Kingston, J. 1990. Cognitive processing in the perception of the speech. J. *Acoust. Soc. Am.*, **88**: S56. (Abstract).

Buckley, E. 2001. Polish o-raising and phonological explanation, Ann. Mtg. Ling. Soc. Am., Washington, DC.

Carlton, T. R. 1990. *Introduction to the Phonological History of the Slavic Languages*, Columbus: Slavica.

Chen, M. 1970. Vowel length variation as a function of the voicing of the consonant environment, *Phonetica*, **22**: 125-159.

Cho, T., & Keating, P. 2000. Articulatory and acoustic studies on domain-initial strengthening in Korean, *J. Phon.*, **29**: 155-190.

Davis, S., & Summers, W. V. 1989. Vowel length and closure duration in word-medial VC sequences, *J. Phon.*, **17**: 339-353.

Delgutte, B., & Kiang, N. Y. S. 1984a. Speech coding in the auditory nerve III: Voiceless fricative consonants, *J. Acoust. Soc. Am.*, **75**: 887-897.

Delgutte, B., & Kiang, N. Y. S. 1984b. Speech coding in the auditory nerve IV: Sounds with consonant-like dynamic characteristics, *J. Acoust. Soc. Am.*, **75**: 897-907.

Delgutte, B. 1996. Auditory neural processing of speech. In W. J. Hardcastle and J. Laver (eds.), Handbook of Phonetic Sciences, 507-538. Oxford: Blackwell Publishing.

Diehl, R. L., & Walsh, M. A. 1989. An auditory basis for the stimulus-length effect in the perception of stops and glides, *J. Acoust. Soc. Am.*, **85**: 2154-2164.

Dilley, L., Shattuck-Hufnagel, S., & Ostendorf, M. 1996. Glottalization of word-initial vowels as a function of prosodic structure, *J. Phon.*, **24**: 423-444.

Fougeron, C. 2000. Articulatory properties of initial segments in prosodic constituents in French, *J. Phon.*, **29**: 109-135.

Fougeron, C., & Keating, P. A. 1997. Articulatory strengthening at edges of prosodic domains, J. Acoust. Soc. Am., **106**: 3728-3740.

Fowler, C. A. 1992. Vowel duration and closure duration in voiced and unvoiced stops: There are no contrast effects here, *J. Phon.*, **20**: 143-166.

Fowler, C. A., Best, C. T., & McRoberts, G. W. 1990. Young infants' perception of liquid coarticulatory influences on following stop consonants, *Percept. & Psychophys.*, **48**: 559-570.

Fowler, C. A., Brown, J. M., & Mann, V. A. 1999. Compensation for coarticulation in audiovisual speech perception, *Proc. XIVth Intl. Cong. Phon. Sci.*, San Francisco, 639-642.

Fujimura, O., Macchi, M. J., & Streeter, L. A. 1978. Perception of stop consonants with conflicting transitional cues: A cross-linguistic study, *Lg. & Sp.*, **21**: 337-346.

Gussman, E. 1980. *Studies in Abstract Phonology*, Cambridge, MA: MIT Press.

Hajek, J. 1997. *Universals of Sound Change in Nasalization*, Publications of the Philological Society, 31, Oxford, UK: Blackwell Publishers.

Holt, L. R., Lotto, A. J., & Kluender, K. R. 2000. Neighboring spectral context influences vowel identification, *J. Acoust. Soc. Am.*, **108**: 710-722.

Janda, R. 1979. Double-cross in phonology: Why word boundary (often) acts like a consonant, *Proc. 5th Ann. Mtg. Berkeley Ling. Soc.*, 397-412.

Jun, S.-A. 1993. *The Phonetics and Phonology of Korean Prosody*, Ph.D. dissertation, Ohio State University.

Keating, P. A. 1984. Phonetic and phonological representation of stop consonant voicing, *Lg.*, **60**: 286-319.

Kelso, J. A. S., Saltzman, E. L., & Tuller, B. 1986. The dynamical perspective on speech production: Data and theory, *J. Phon.*, **14**: 29-59.

Kewley-Port, D., Pisoni, D. B., & Studdert-Kennedy, M. 1983. Perception of static and dynamic acoustic cues to place of articulation in initial stop consonants, *J. Acoust. Soc. Am.*, **73**: 1779-1793.

Kingston, J. 1985. *The Phonetics and Phonology of the Timing of Oral and Glottal Events*, Ph.D. dissertation, University of California, Berkeley.

Kingston, J. 1990. Articulatory binding, In J. Kingston & M. E. Beckman (eds.), *Papers in Laboratory Phonology I*, 406-434. Cambridge, UK: Cambridge University Press.

Kingston, J. 1991. Integrating articulations in the perception of vowel height, *Phonetica*, **47**: 149-179.

Kingston, J., & Diehl, R. L. 1994. Phonetic knowledge, *Lg.*, **70**: 419-454.

Kingston, J., & Diehl, R. L. 1995. Intermediate properties in the perception of distinctive feature values, In A. Arvaniti & B. Connell (eds.), *Papers in Laboratory Phonology IV*, 7-27. Cambridge, UK: Cambridge University Press.

Kingston, J., Diehl, R. L., Kirk, C. J., & Castleman, W. A. submitted. On the internal perceptual structure of phonological features: the [voice] distinction, Ms. University of Massachusetts, Amherst and University of Texas, Austin.

Kingston, J., & Macmillan, N. A. 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. Acoust. Soc. Am.*, **97**: 1261-1285.

Kirk, C. J. 2000. Syllabic cues to word segmentation, *Proc. Workshop Spoken Word Access Processes*, 131-134. Nijmegen: Max Planck Inst. Psycholing.

Kirk, C. J. 2001. *Phonological Constraints on the Segmentation of Continuous Speech*, Ph.D. dissertation, U. Massachusetts, Amherst.

Kohler, K. J. 1979. Dimensions in the perception of fortis and lenis plosives, *Phonetica*, **36**: 332-343.

Kluender, K. R., Diehl, R. L., & Wright, B. A. 1988. Vowel-length differences before voiced and voiceless consonants: An auditory explanation, *J. Phon.*, **16**: 153-169.

Kluender, K. R., & Walsh, M. A. 1992. Amplitude rise time and the perception of the voiceless affricate-fricative distinction, *Percept. & Psychophys.*, **51**: 328-333.

Krakow, R. A. 1989. *The Articulatory Organization of Syllables: A Kinematic Analysis of Labial and Velar Gestures*, Ph.D. dissertation, Yale University.

Krakow, R. A. 1993. Nonsegmental influences on velum movement pattersn: Syllables, sentences, stress, and speaking rate, In M. K. Huffman & R. A. Krakow (eds.), *Nasals, Nasalization, and the Velum, Phonetics and Phonology, 5*, 87-116. San Diego: Academic Press.

Krakow, R. A., Beddor, P. S., Goldstein, L. M., & Fowler, C. A. 1988. Coarticulatory influences on the perceived height of nasal vowels, *J. Acoust. Soc. Am.*, **83**: 1146-1158.

Laeufer, C. 1992. Patterns of voicing-conditioned vowel duration in French and English, *J. Phon.*, **20**: 411-440.

Lahiri, A., Gewirth, L., & Blumstein, S. E., 1984. A reconsideration of acoustic invariance for place of articulation in diffuse stop consonants: Evidence from a cross-language study, *J. Acoust. Soc. Am.*, **76**: 391-404.

Lindblom, B., & Studdert-Kennedy, M. 1967. On the role of formant transitions in vowel recognition, *J. Acoust. Soc. Am.*, **42**: 830-843.

Lisker, L. 1986. "Voicing" in English: A catalogue of acoustic features signaling /b/ vs /p/ in trochees. *Lg. & Sp.*, **29**: 3-11.

Liu, S. A. 1996. Landmark detection for distinctive feature-based speech recognition, *J. Acoust. Soc. Am.*, **100**: 3417-3430.

Lombardi, L. 1991. *Laryngeal Features and Laryngeal Neutralization*, Ph.D. dissertation, University of Massachusetts, Amherst.

Lombardi, L. 1995. Laryngeal neutralization and syllable wellformedness, *Nat. Lg. Ling. Theory*, **13**: 39-74.

Lotto, A. J., & Kluender, K. R. 1998. General contrast effects in speech perception: Effect of preceding liquid on stop consonant identification, *Percept. & Psychophys.*, **60**: 602-619.

Mack, M. 1982. Voicing-dependent vowel duration in English and French: Monolingual and bilingual production, *J. Acoust. Soc. Am.*, **71**: 173-178.

Macmillan, N. A., Kingston, J., Thorburn, R., Walsh Dickey, L., & Bartels, C. 1999. Integrality of nasalization and F_1. II. Basic sensitivity and phonetic labeling measure distinct sensory and decision-rule interactions, *J. Acoust. Soc. Am.*, **106**: 2913-2932.

Maddieson, I. this volume. Phonetics in the field, Proc. 28th Ann. Mtg. Berkeley Ling. Soc.

Maddieson, I., & Precoda, K. 1992. *UPSID and Phoneme, User Manual*, Los Angeles, University of California.

Maeda, S. 1993. Acoustics of vowel nasalization and articulatory shifts in French nasal vowels, In M. K. Huffman & R. A. Krakow (eds.), *Nasals, Nasalization, and the Velum, Phonetics and Phonology 5*, 147-167. San Diego: Academic Press.

Malécot, A. 1958. The role of releases in the identification of released final stops, *Lg.*, **34**: 370-380.

Mann, V. A. 1980. Influence of preceding liquid on stop-consonant perception, *Percept. & Psychophys.*, **28**: 407-412.

Mann, V. A., & Repp, B. H. 1981. Influence of preceding fricative on stop consonant perception, *J. Acoust. Soc.Am.*, **69**, 548-558.

McQueen, J. M. 1998. Segmentation of continuous speech using phonotactics, *J. Mem. Lg.*, **39**: 21-46.

Moreton, E. 2002. Realization of the English postvocalic [voice] contrast in F_1 and F_2, Ms., Johns Hopkins University.

Nearey, T. M. 1990. The segment as a unit of speech perception, *J. Phon.*, 347-373.

Nearey, T. M. 1995. A double-weak view of trading relations: Comments on Kingston and Diehl, In A. Arvaniti & B. Connell (eds.), *Papers in Laboratory Phonology IV*, 28-40. Cambridge, UK: Cambridge University Press.

Nearey, T. M. 1997. Speech recognition as pattern recognition, *J. Acoust. Soc. Am.*, 101, 3241-3253.

Nearey, T. M. in press. The factorability of phonological units in speech perception: Simulating the results of speech reception in noise, In *Papers in Laboratory Phonology VI*, Cambridge, UK: Cambridge University Press.

Ohala, J. J. 1990. The phonetics and phonology of aspects of assimilation, In J. Kingston & M. E. Beckman (eds.), *Papers in Laboratory Phonology I*, 258-275. Cambridge, UK: Cambridge University Press.

Ohala, J. J., & Feder, D. 1995. Listeners' identification of speech sounds is influenced by adjacent 'restored' phonemes, *Phonetica*, **51**: 111-118, 1994.

Ohala, J. J., Riordan, C. J., & Kawasaki, H. 1978. The influence of consonant environment upon identification of transitionless vowels, *J. Acoust. Soc. Am.*, **64**: S18. (Abstract)

Parker, E. M., Diehl, R. L., Kluender, K. R. 1986. Trading relations in speech and non-speech, *Percept. & Psychophys.*, **39**: 129-142.

Pierrehumbert, J. B. & Talkin, D. 1992. Lenition of /h/ and glottal stop, In G. R. Docherty & D. R. Ladd (eds.), *Papers in Laboratory Phonology II*, 90-116. Cambridge, UK: Cambridge University Press.

Port, R., & Dalby, J. 1982. Consonant/vowel ratio as a cue for voicing in English, *Percept. & Psychophys.*, **32**: 141-152.

Repp, B. H. 1978. Perceptual integration and differentiation of spectral information of spectral cues for intervocalic stop consonants, *Percept. & Psychophys.*, **24**: 471-485.

Repp, B. H. 1983. Bidirectional context effects in the perception of VC-CV sequences, *Percept. & Psychophys.*, **33**: 147-155.

Repp, B. H., & Mann, V. A. 1981. Perceptual assessment of fricative-stop coarticulation, *J. Acoust. Soc. Am.*, **69**: 1154-1163.

Repp, B. H., & Mann, V. A. 1982. Fricative-stop coarticulation: Acoustic and perceptual evidence, *J. Acoust. Soc. Am.*, **71**: 1562-1567,

Silva, D. J. 1992. *The Phonetics and Phonology of Stop Lenition in Korean*, Ph.D. dissertation, Cornell University.

Silverman, D. 1995. *Phasing and Recoverability*, Ph.D. dissertation, U. California, Los Angeles.

Smith, J. 2002. *Phonological Augmentation in Prominent Positions*, Ph.D. dissertation, U. Massachusetts, Amherst.

Smits, R. 2001a. Evidence for hierarchical categorization of coarticulated phonemes, *J. Exptl. Psych.: Human Percept. Perform.*, **27**: 1145-1162.

Smits, R. 2001b. Hierarchical categorization of coarticulated phonemes: A theoretical analysis, *Percept. & Psychophys.*, **63**: 1109-1139.

Smits, R., ten Bosch, L., & Collier, R. 1996a. Evaluation of various sets of acoustic cues for the perception of prevocalic stop consonants. I. Perception experiment, *J. Acoust. Soc. Am.*, **100**: 3852-3864.

Smits, R., ten Bosch, L., & Collier, R. 1996b. Evaluation of various sets of acoustic cues for the perception of prevocalic stop consonants. II. Modeling and evaluation, *J. Acoust. Soc. Am.*, **100**: 3865-3881.

Sproat, R., & Fujimura, O. 1993. Allophonic variation in English /l/ and its implications for phonetic implementation, *J. Phon.*, **21**: 291-311.

Steele, R. D. 1973. *The Segmental Phonology of Contemporary Standard Polish*, Ph.D. dissertation, Harvard University.

Steriade, D. 1995. Positional neutralization, Ms. University of California, Los Angeles.

Steriade, D. 2000. Directional asymmetries in place perception, In E. Hume & K. Johnson (eds.), *The Role of Speech Perception in Phonology*, 219-250. San Diego: Academic Press.

Steriade, D. 2001. The phonology of perceptibility effects: The P-Map and its consequences for constraint organization, Ms., University of California, Los Angeles.

Stevens, K. N. 1998. *Acoustic Phonetics*, Cambridge, MA: MIT Press.

Stevens, K. N., & Blumstein, S. E. 1978. Invariant cues for place of articulation in stop consonants, *J. Acoust. Soc. Am.*, **64**: 1358-1368.

Stevens, K. N., & Blumstein, S. E. 1981. The search for invariant acoustic correlates of phonetic features, In P. D. Eimas & J. L. Miller (eds.), *Perspectives on the Study of Speech*, 1-38. Hillsdale, NJ: Lawrence Erlbaum Associates.

Stevens, K. N., & Keyser, S. J. 1989. Primary features and their enhancement in consonants, *Lg.*, **65**: 81-106.

Stieber, Z. 1973. *A Historical Phonology of the Polish Language*, Heidelberg: Carl Winter Universitätsverlag.

Summers, W. V. 1987. Effect of stress and final-consonant voicing on vowel production: Articulatory and acoustic analyses, *J. Acoust. Soc. Am.*, **82**: 847-863.

Thomas, E. R. 2000. Spectral differences in /ai/ offsets conditioned by the voicing of the following consonant, *J. Phon.*, **28**: 1-25.

Tuller, B., & Kelso, J. A. S. 1991. Production and perception of syllable structure, *J. Sp. Hear. Res.*, **34**: 501-508.

Turk, A. E. 1993. Articulatory phonetic clues to syllable affiliation: Gestural characteristics of bilabial stops, In P. A. Keating (ed.), *Papers in Laboratory Phonology III*, 107-135. Cambridge, UK: Cambridge UP.

Walley, A. C., & Carrell, T. D. 1983. Onset spectra and formant transitions in adult's and child's perception of place of articulation, *J. Acoust. Soc. Am.*, **73**: 1011-1022.

Walsh, M. A. 1991. Formant transition duration and amplitude rise time as cues to the stop-glide distinction, *Quat. J. Exptl. Psych., Hum. Exptl. Psych.*, **43A**: 603-620.

Whalen, D. S., & Beddor, P. S. 1989. Connections between nasality and vowel duration and height: Elucidation of the Eastern Algonquian intrusive nasal, *Lg.*, **65**: 457-486.

Williams, D. R. 1986. *Role of Dynamic Information in the Perception of Coarticulated Vowels*, Ph.D. dissertation, University of Connecticut.

Wilson, C. 2001. Consonant cluster neutralization and targeted constraints, *Phonology*, **18**: 147-197.

Wright, R. (1999). Tsou consonant clusters and auditory cue preservation, In E. Zeitoun & P. Li (eds.), *Selected Papers from the Eighth International Conference on Austronesian Linguistics*, v. 1, 227-312. Taipei: Academia Sinica, Institute of Lingiustics.

Wright, R. (2001). Perceptual cues in contrast maintenance, In E. Hume & K. Johnson (eds.) *The Role of Speech Perception in Phonology*, 251-277. San Diego: Academic Press.

Linguistics Department
University of Massachusetts
Amherst, MA 01003

jkingston@linguist.umass.edu

Anaphoric R-Expressions as Bound Variables[1]

FELICIA LEE
University of British Columbia

0. The Problem

San Lucas Quiaviní Zapotec (SLQZ), an Otomanguean language of southern Mexico, regularly allows apparent Principle B and C violations.[2] R-expressions may bind identical R-expressions (1-2) and pronouns may locally bind identical pronouns (3):

(1) R-yu'lààa'z Gye'eihlly Gye'eihlly
hab-like Mike Mike
"Mike likes himself."

(2) R-cààa'z bxuuhahz ch-iia bxuuhahz
hab-want priest irr-go priest
"The priest wants to go."

(3) R-yu'lààa'z-ëng la'anng
hab-like-3s.prox 3s.prox
"He/she likes himself/herself."

Thai also allows apparent Principle C violations:

(4) John koonnuat John
John shaved John
"John shaved himself." [Thai]

[1] I am grateful to Rodrigo Garcia and Sugunya Ruangjaroon for providing the SLQZ and Thai data and judgments in this paper. I am also grateful to Irene Heim, Jim Huang, Pamela Munro, Tim Stowell, and audiences at UBC, University of Canterbury, and NELS 32 for their suggestions and questions on earlier stages of this project. Any remaining errors are my own.
[2] This pattern was described in detail in Munro 1994. Some of the grammaticality judgments reported in this earlier work differ from those found here.

(5) Aajarn kit waa puak rau chççp aajarn
teacher think that all we like teacher
"The teacher$_i$ thinks we like him$_i$." [Thai]

1. Proposal

This paper will show that Principles B and C do indeed hold in SLQZ and Thai. I will argue that putatively "bound" R-expressions are not, in fact, true R-expressions, but bound variables spelled out as copies of their antecedents. Likewise, "locally bound" pronouns are bound copies of their antecedents.

This is consistent with the view that reflexive predicates represent functions mapping a single argument to both argument positions (Reinhart and Reuland 1993):

(6) λx (P..x…x..)

Under Reinhart and Reuland's assumptions, reflexivity is realized in one of two ways: either a predicate is lexically specified as reflexive, or it needs to be 'reflexive-marked' by a reflexive morpheme. SLQZ lacks an independent series of reflexive pronouns; thus it uses bound copies to reflexive-mark predicates.

2. An Earlier Proposal

The Thai pattern was noted by Lasnik (1986). He concluded that Principle C is subject to parametric variation: Principle C holds in languages such as English, but not in languages such as Thai. This theory has been invoked to account for the binding facts of Quiegolani Zapotec (related to, but mutually unintelligible from, SLQZ) (Black 1994).

3. Testing the Hypothesis

The idea that Principle C is absent in languages such as Thai and SLQZ forces several predictions about the behavior of these languages. The following sections examine these predictions, and show that Principle C does indeed hold in Thai and SLQZ. Thus, an alternate account must be made for their binding patterns.

3.1. Prediction 1: All R-expressions Should Be Bindable

If Principle C did not hold in some languages, then these languages should allow R-expressions to be bound in any context. However, the contexts in which R-expressions may appear to be bound in Thai and SLQZ are extremely limited.

Lasnik notes that in Thai, R-expressions cannot be bound by pronouns (7). The same constraint holds in Quiegolani Zapotec (8) (Black 1994) and SLQZ (9):

(7) *Khaw chççp John
he likes John
"He$_i$ likes John$_i$." [Thai: Lasnik 1986, p.154]

(8) *Per n-an-t men pa go r-zak Merse
but s-know-neg 3s what thing h-have Mercedes
"But she_i didn't know what $Mercedes_i$ had."
[Quiegolani Zapotec: Black 1994, p. 98]

(9) *B-gwi'ih-ëng lohoh Gye'eihlly
perf-look-3s.prox at Mike
"He_i looked at $Mike_i$."

Lasnik attributes this constraint to a referential hierarchy on binding: less referential elements may not bind more referential ones. Thus, R-expressions may be bound by R-expressions, but not by pronouns.

However, constraints on bound R-expressions are stricter than Lasnik suggests. In Thai and SLQZ, R-expressions cannot be bound by different (equally referential) R-expressions (10-11); pronouns can be locally bound neither by R-expressions (12) nor pronouns differing in formality or proximity features (13). The same constraints hold for Thai (14):

(10) R-yu'lààa'z Gye'eihlly me's
hab-like Mike teacher
"$Mike_i$ likes the teacher $_{j/*i}$."

(11) R-yu'lààa'z me's Gye'eihlly
hab-like teacher Mike
"The $teacher_i$ likes Mike $_{j/*i}$."

(12) R-yu'lààa'z Gye'eihlly la'anng
hab-like Mike 3s.prox
"Mike likes him/*himself."

(13) R-yu'lààa'z -ih la'anng
Hab-like-3s.prox 3s.dist.
"He/she_i likes him/her $_{j/*i}$."

(14) *John koonnuat aajarn
John shave teacher
"$John_i$ shaved the $teacher_{*i}$." [Thai]

Thus, apparent binding of R-expressions and local binding of pronouns appear to be subject to the following constraint:

- *The Identical Antecedent Requirement*: The only R-expressions that can be bound, and pronouns that can be locally bound, are exact copies of their antecedents.

This requirement calls into question the idea that Principle C may be freely disobeyed in languages such as Thai and SLQZ.

3.2. Prediction 2: Thai and SLQZ Should Not Show Crossover Effects

If Principle C were absent in Thai and SLQZ, then wh-traces are predicted to be subject to A-binding, and these languages should not show crossover effects.

Both languages, however, show both strong and weak crossover effects. In (15) and (16), the wh-trace cannot be coindexed with any c-commanding arguments:

(15) Q: Tu r-ralloh la'anng r-yu'làà'z (t) Li'eb (t)
Who hab-think 3s.prox hab-like Felipe
"Who does he think Felipe likes? /Who does he think likes Felipe?"

A: Lia Paamm-zhi'
Ms. Pam-maybe
"Maybe Pam."

(16) *Khray$_i$ thii khaw$_i$ khit t$_i$ waa Nit rak t$_i$
who COMP he think COMP Nit love
"Who$_i$ does he$_i$ think Nit loves?" [Thai: Ruangjaroon 2001]

Because SLQZ has basic VSO word order and no subject agreement on verbs with non-pronominal subjects, argument wh-questions are often ambiguous between subject and object readings. (This is seen in the two possible interpretations of (15)). (17), however, shows a weak crossover effect: the possessed nominal *x:nnàaanni'*, 'his/her mother', can only be interpreted as the object (not the subject) of the sentence:

(17) Tu r-yu'làà'z t x:-nnàaan-ni' *t
who hab-like gen-mother-refl.poss
"Who like his/her own mother/*Who does his/her own mother like?"

Thai also shows weak crossover effects. In (18), 'who' must be disjunct from 'his':

(18) MQ khong kao chççp kray
Mother his$_i$ like who$_j$"
"Who$_j$ does his$_i$ mother like?" [Thai]

These data show that wh-traces cannot be A-bound in Thai and SLQZ, which is unexpected if these languages lacked Principle C.

3.3. Prediction 3: Bound DPs Should Be Fully Referential

If Principle C were absent in Thai and SLQZ, then fully referential DPs should be able to be bound, and should have the same referential force as other R-expressions.

This, however, proves not to be the case. "Bound" R-expressions in Thai and SLQZ do not receive strict readings in VP deletion contexts, as do typical R-expressions. Rather, they allow only sloppy (bound variable) readings:

(19) B-gwi'ih Gye'eihlly lohoh Gye'eihlly zë'cy cahgza' Li'eb
perf-look Mike at Mike likewise Felipe
"Mike looked at himself, and Felipe did too."
(*Felipe looked at Mike/ Felipe looked at himself)

(20) John koonnuat khong John lae Peter ko muankan
John shave of John and Peter the same
"John shaved himself, and Peter did too."
(*Peter shaved John/ Peter also shaved himself) [Thai]

Thus, bound copies in SLQZ and Thai do not have the referential force of normal R-expressions. Rather, they behave like bound variables.

4. Bound Copies as Bound Variables: The Solution, Revisited

The presence of Principle C effects and the bound variable readings of apparently bound R-expressions support the proposal that they are not truly referential, but are bound variables spelled out as copies of their antecedents. Likewise, "locally bound" pronouns are also bound copies. This accounts for the general constraint against bound R-expressions and locally bound pronouns, as well as the presence of crossover effects. The copy status of these apparently bound expressions also accounts for the Identical Antecedent Requirement.

5. Semantic Consequences

This proposal also predicts another binding constraint in languages such as SLQZ and Thai. If bound copies are bound variables, then they should be, in type-theoretical terms, elements of type *e*. Thus, only DPs of type *e* should be able to appear as bound copies.

This prediction is borne out. Referential DPs and pronouns, which are elements of type *e*, may appear as bound copies, but quantified phrases, elements of type <<e,t>,t>, cannot be bound copies[3]:

(21) *B-guhty cho'nn ra bxuuhahz cho'nn ra bxuuhahz
perf-kill three pl. priest three pl. priest
"Three priests killed themselves."

[3] I assume that in these contexts, both names and definite DPs are elements of type *e*, consistent with Partee's (1986) claim that natural languages allow definite descriptions to be either type *e* or type <<e,t>, t>.

(21) *B-guhty cho'nn ra bxuuhahz cho'nn ra bxuuhahz
perf-kill three pl. priest three pl. priest
"Three priests killed themselves."

(22) ??R-a txup tson wnaa r-ka txup tson wnaa gyus
Hab-go two three woman hab-buy two three woman pot
"A few women went to buy a pot."
[Quiegolani Zapotec: Black 1994, p. 103]

(23) *Thuk khon konnuad thuk khon
every one shave every one
"Everyone shaved himself." [Thai]

5.1. QPs and Bound Copies

Bound QP copies cannot appear—at least not with a reflexive reading—because they would cause a semantic type clash. Recall that reflexive predicates are assumed to be functions mapping a single argument to both argument positions:

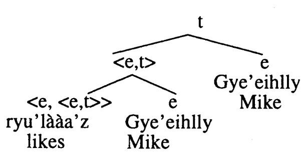
(24) $\lambda x\ (P..x...x..)$

Thus, a reflexive predicate such as "kill oneself" is a function that takes an entity to form a second function, which takes the same entity to form a proposition:

(25) [[kill oneself]]: $=[\lambda x\in D_e.\ \lambda x\in D_e.\ x \text{ kill } x]$

According to this representation, then, bound copies must be of type *e*. Thus, a simple reflexive expression such as (1), gets the representation in (26):

(26) [[Mike likes himself]]:$=[\ \lambda x\in D_e.\ \lambda x\in D_e.\ x \text{ likes } x]$ (Mike) = 1

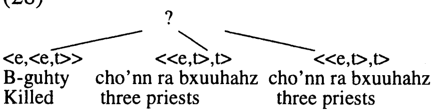


If the reflexive argument is a QP, however, a type clash occurs. Consider the ungrammatical example (21), repeated below:

(27) *B-guhty cho'nn ra bxuuhahz cho'nn ra bxuuhahz
perf-kill three pl. priest three pl. priest
"Three priests killed themselves."

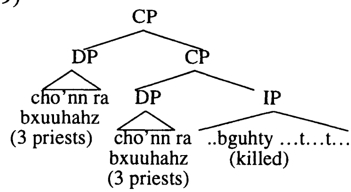
From (28), we see that the QPs cannot combine compositionally with the predicate, which only selects entities as arguments:

(28)

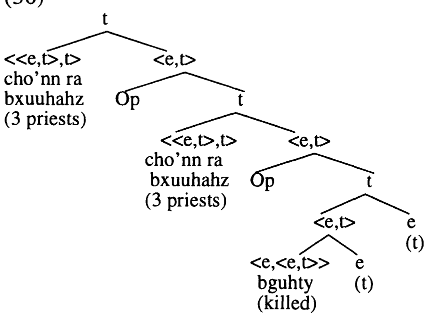


One means of making this structure licit is to type-shift the copies by raising them at LF, leaving traces of type *e* to combine with the predicate. The following syntactic and semantic structures will result:

(29)



(30)



The LF structure in (30), however, does not denote the intended reflexive reading: since each of the traces is bound by a different QP, the sentence can only mean that three priests killed three other priests.

5.2. Reflexivization of QPs

SLQZ uses a different pattern to reflexivize QPs: the QP is base-generated as a preverbal topic, and the actual subject of the reflexive predicate is realized as a distal pronoun[4]. The reflexive object is a bound copy of the subject pronoun:

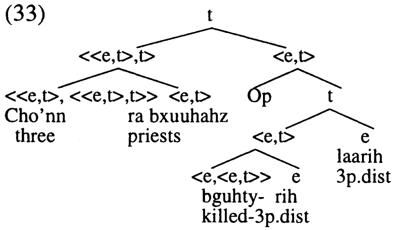
(31) Cho'nn ra bxuuhahz b-guhty-rih la'arih
Three pl priest perf-kill-3p.dist 3p. dist
"Three priests killed themselves."

[4] See Munro (to appear) for the uses of proximate and distal forms in narrative.

(32) Yra'ta' ra bxuuhahz b-guhty-rih la'arih
Every pl priest perf-kill-3p.dist 3p. dist
"Every priest killed himself."

In this construction, reflexive arguments are realized as pronominal variables (thus, elements of type *e*), with the QP base-generated in an A' (operator-like) position. (33) shows the semantic representation of (31):

(33)



Here, a single QP takes scope over both pronominal, giving a reflexive reading.

To sum up, the failure of quantified arguments to appear as bound copies is due to the bound variable status of bound copies: variables are elements of type *e*, and only elements of type *e* may appear as bound copies.[5]

6. SLQZ Bound Copies as Long-Distance Anaphora

SLQZ also allows non-locally-bound copies, which also behave as bound variables. This section will propose that non-locally-bound copies are long-distance anaphora, and share syntactic and semantic features attested in long-distance anaphora crosslinguistically.

6.1. Bound Copies as Nominative Anaphors

Bound copies can be subjects of embedded clauses. (SLQZ has no infinitival clauses; all verbs are marked with tense/aspect markers.) Embedded subject copies are also interpreted as bound variables:

(34) R-cààa'z Gye'eihlly g-ahcnèe Gye'eihlly Lia Paamm zë'cy cahgza' Li'eb
hab-want Mike irr-help Mike fem. Pam likewise Felipe
"Mike wants to help Pam, and so does Felipe."
(Felipe also wants to help Pam /*Felipe also wants Mike to help Pam)

6.2. Embedded Object Copies

Bound copies can also be objects of embedded clauses:

[5] Bare nouns in SLQZ may be interpreted as singular or plural, definite or indefinite, depending on context. I will assume that bare nouns are DPs with silent heads, and these DPs are treated as entities.

(35) R-ralloh Gye'eihlly r-yu'làà a'z Lia Paamm Gye'eihlly
Hab-think Mike hab-like fem. Pam Mike
"Mike$_i$ thinks Pam likes him$_i$."

Object bound copies in embedded clauses can get apparently referential readings:

(36) R-ralloh Gye'eihlly r-yu'l àà a'z-ënn Gye'eihlly
Hab-think Mike hab-like-1p Mike
"Mike$_i$ thinks we likes him$_i$

chiru' zë'cy cahgza' Li'eb
also likewise Felipe
and so does Felipe."
(Felipe thinks we like Mike/Felipe thinks we like him (Felipe))

This, however, is not necessarily evidence against the anaphoric status of the bound copy. Thráinsson (1993) notes that the Icelandic long-distance anaphor *sig* allows only a sloppy reading in VP-deletion contexts when bound locally, but allows both strict and sloppy readings when its antecedent binds it across a clause:

(37) Jón rakaDi sig og Péter gerDi DaD líka
John shaved self and Peter did so too
(Peter shaved himself (Peter)/ *Peter shaved John)

(38) Jón sagDi [aD Dú hefDir svikiD sig] og Péter gerDi DaD líka
John$_i$ said that you had betrayed self$_i$ and Peter did so too
(Peter said that you betrayed John / Peter$_i$ said that you betrayed him$_i$)

This shows that SLQZ bound copies show the same interpretive behavior as local and long-distance anaphors crosslinguistically.

6.3. Bound Copies in Adjunct Clauses

Bound copies may also appear in adjunct clauses, where they also get bound variable readings:

(39) Zi'cygàa' nih cay-uhny Gye'eihlly zèèiny b-ìi'lly-ga' Gye'eihlly
While that prog-do Mike work perf-sing-also Mike
"While Mike worked, he sang."

This is unexpected given that no c-command relation holds between the copy and its antecedent. However, the relation between the copy and its antecedent is not unattested: Huang and Tang (1993) note that Chinese long-distance anaphors may appear in adjunct clauses with antecedents in main clauses:

(40) Ta zhidao [[suiran Lisi piping-le ziji]
He know though Lisi criticise-ASP self
"He$_i$ knows that although Lisi$_j$ criticized self$_{i/j}$

dajia haishi hen xihuan ta
all still very like him
we still like him." [Chinese: Huang and Tang 1993, p. 279]

Bound copies in adjunct clauses, like other copies, get bound variable readings and are subject to the Identical Antecedent requirement:

(41) Zi'cygàa' nih cay-uhny Gye'eihlly zèèiny b-ìi'lly-ga' Gye'eihlly
While that prog-do Mike work perf-sing-also Mike
"While Mike was working, he sang

zë'cy cahgza' Li'eb
likewise Felipe
and so did Felipe."
(*Felipe sang while Mike worked/Felipe sang while he (Felipe) worked)

(42) Zi'cygàa' nih cay-uhny Gye'eihlly zèèiny b-ìi'lly-ga'-ng
While that prog-do Mike work perf-sing-also-3s.prox
"While Mike was working, he/she (someone else) sang."

This supports the idea that adjunct bound copies are semantically dependent on their antecedents, and are not merely instances of accidental coreference.

6.4. Non-Local Copies Are Not Logophors

It has been claimed that the non-bound anaphora cross-linguistically are logophors: pronominal elements representing one whose speech, thoughts, or perceptions are being reported. An example of this is the use of *myself* in (43):

(43) As for myself, I like sugar in my coffee.

Long-distance anaphora have also been claimed to be logophors (Reinhart and Reuland 1993, Huang and Liu 2001). Under these accounts, the distribution of long-distance anaphora is constrained primarily by semantic and pragmatic factors, rather than purely structural ones.

However, non-locally bound copies in SLQZ can appear in contexts in which their use is not logophoric. Such contexts are noted by Dubinsky and Hamilton (1998), who argue that epithets can only be c-commanded by their antecedents in anti-logophoric environments:

(44) *Felipe$_i$ is afraid his teacher failed the poor guy$_i$.

(45) Felipe$_i$ ran over the man who tried to help the poor guy$_i$.

Example (44) is ungrammatical because *the poor guy* is c-commanded in a logophoric context: Felipe's feelings are being described. (45), on the other hand, is grammatical because the context in which the epithet appears is anti-logophoric: it says nothing about Felipe's state of mind during the action.

If non-locally-bound copies in SLQZ were logophors, they would be ruled out in contexts similar to that in (45). However, this prediction is not borne out: (46) shows that non-locally-bound copies can occur in anti-logophoric contexts:

(46) B-tàa'az Gye'eihlly bèe'cw nih b-da'uhgya'ah Gye'eihlly
perf-hit Mike dog REL perf- bite Mike
"Mike hit the dog that bit him."

Here, the copy is licit even though (46), like (45), does not describe the perspective of the subject or its copy. Hence, it cannot be the case that bound copies are logophors.

7. Conclusion

This paper has shown that the apparently bound R-expressions and locally bound pronouns that appear in SLQZ and Thai do not represent Principles B or C violations. Rather, these elements are semantically and syntactically bound variables that are spelled out as copies of their antecedents. In SLQZ, which lacks a morphologically distinct series of anaphors, bound copies serve as both local and long-distance anaphora.

References

Black, Cheryl. 1994. Quiegolani Zapotec Syntax. PhD dissertation, UC-Santa Cruz.

Dubinsky, Stanley & R. Hamilton. 1998. Epithets as Antilogophoric Pronouns. *Linguistic Inquiry* 29.4: 685-673.

Hornstein, Norbert. 2001. *Move! A Minimalist Theory of Construal.* Oxford: Blackwell.

Huang, C.T. James. 1982. Logical Relations in Chinese and the Theory of Grammar. PhD dissertation, MIT.

Huang, C.T. James & C.-S. Luther Liu. 2001. Logophoricity, Attitudes, and *Ziji* at the Interface. In P. Cole et al (eds.) *Long-Distance Reflexives, Syntax and Semantics*. San Diego: Academic Press.

Huang , C.T. James & C.-C. Jane Tang. 1993. The Local Nature of Long-Distance Anaphora in Chinese. In J. Koster & E. Reuland (eds.) *Long Distance Anaphora*. Cambridge: Cambridge University Press.

Koopman, Hilda & D. Sportiche. 1989. Pronouns, Logical Variables, and Logophoricity in Abe. *Linguistic Inquiry* 20.4: 555-888.

Lasnik, Howard. 1991. On the Necessity of Binding Conditions. In R. Frieden, (ed.) *Principles and Parameters in Comparative Grammar*. Cambridge: MIT Press.

Lee, Felicia. 1998. Evidence for Tense in a 'Tenseless' Language. *Proceedings of NELS 29.*

Lee, Felicia. 1999. Antisymmetry and the Syntax of San Lucas Quiaviní Zapotec. PhD dissertation, UCLA.

Munro, Pamela. 1994. Coreference Phenomena in San Lucas Quiaviní Zapotec. Ms., UCLA.

Munro, Pamela. To appear. Hierarchical Pronouns in Discourse: Third-Person Pronouns in San Lucas Quiaviní Zapotec Narratives. *SWJL.*

Munro, Pamela & F.H. Lopez (with R. Garcia and O. V. Méndez). 1999. *Dicyonaary X:te:e'n Di:i'zh Sah Sann Luu'c: San Lucas Quiaviní Zapotec Dictionary. Diccionario Zapoteco de San Lucas Quiaviní.* Los Angeles: UCLA Chicano Studies Research Center Publications.

Partee, Barbara. 1986. Noun Phrase Interpretation and Type-Shifting Principles. In Jeroen Groenendijk et al (eds.) *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*. Dordrecht: Foris.

Reinhart, Tanya & E. Reuland. 1993. Anaphors and Logophors: An Argument Structure Perspective. In J. Koster & E. Reuland (eds.) *Long Distance Anaphora*. Cambridge: Cambridge University Press.

Ruangjaroon, Sugunya. 2001. Thai Wh-Movement as Focus. Ms., UBC.

Thráinsson, Höskuldur. 1993. Long-Distance Reflexives and the Typology of NPs. In J. Koster & E. Reuland (eds.) *Long Distance Anaphora*. Cambridge: Cambridge University Press.

Department of Linguistics
University of British Columbia
E270-1866 Main Mall
Vancouver, B.C. V6T 1Z1 Canada

leefa@interchange.ubc.ca

Japanese Listeners' Use of Duration Cues in the Identification of English High Front Vowels[1]

GEOFFREY STEWART MORRISON
Simon Fraser University

0. Introduction

General Canadian English (along with several other English dialects) has two high front vowel phonemes differing in both spectral and duration properties. In identical contexts, the tense /i/ is both more peripheral in the vowel space and longer than the lax /ɪ/ (Nearey & Assmann 1985, Morrison 2002b). Japanese also has two high front vowel phonemes: short /i/ and long /iː/. Japanese long and short vowels are traditionally described as differing only in terms of duration with identical spectral properties (Akamatsu1997).

Ingram & Park (1997) found that Japanese listeners identified Australian English /i/ and /ɪ/ at near perfect levels (an error rate of only 1%). In an L1 assimilation test, Japanese listeners classified English /i/ as Japanese /iː/ at a rate of 100% and classified English /ɪ/ as Japanese /i/ at a rate of 90% (10% as /iː/). The stimuli had been produced by two native English speakers, one of whom spoke faster than the other resulting in different vowel durations: one speaker's /ɪ/ and /i/ vowels were approximately 75 and 255ms long respectively, and the other speaker's vowels were 165 and 390ms. English vowels were consistently classified as long or short Japanese categories irrespective of the inter-speaker duration differences. Ingram & Park hypothesised that the Japanese listeners were able to normalise for speaking rate differences because of their experience with the phonemic vowel duration contrast in Japanese. Korean listeners tested on the same stimuli were not able to normalise for speaking rate differences (modern Korean lacks a vowel duration contrast). The implication is that the Japanese listeners identified English /ɪ/ and /i/ using the same duration criteria that they used to identify Japanese /i/ and /iː/, and normalised for speaking rate differences as they would when identifying Japanese long and short vowels. Presumably the Japanese participants also made some use of spectral cues, since English /ɪ/ was more sometimes correctly identified as /ɪ/.

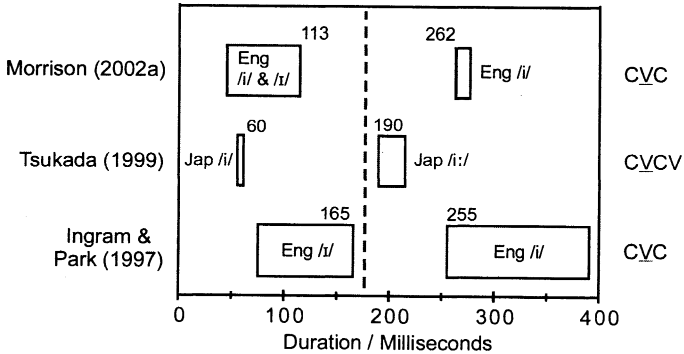
[1]This paper is based on part of my forthcoming MA thesis. My thanks to Dr. Murray J. Munro for comments on an earlier version of the paper, and to the anonymous participants and model speakers.

The results of several studies, however, cast doubt on the ability of Japanese listeners to normalise for speaking rate. Strange et al. (1998) found that in US English /h_bɑ/ words, long English vowels /i e æ ɑ ɔ o u/ and short English vowels /ɪ ɛ ʌ ʊ/ were assimilated to Japanese long and short vowels at rates of 39% and 92% respectively. When the final /ɑ/ in the /h_bɑ/ words was truncated to the same duration as /ʌ/ in order to simulate a faster speaking rate, there was little difference in the duration assimilation pattern: Long and short English vowels were assimilated to Japanese long and short vowels at rates of 43% and 94% respectively. Likewise, Guion et al. (2002) found no significant difference in the assimilation of English vowels to Japanese long and short categories in isolated /b_bo/ words in which /bo/ had two durations. Toda (1999) measure native Japanese listeners' long-short boundary in the final vowel of (C)VCV - (C)VCVV continua. Altering the duration of the first vowel so that it was 100% longer in one series of continua than the other resulted in only a 10% increase in the location of the boundary. Given the small degree of normalisation for simulated speaking rate in these studies, it seems unlikely that normalisation could account for the ability of the Japanese listeners in Ingram & Park (1997) to correctly identify English /i/ and /ɪ/ irrespective of large speaking rate differences. It seems especially unlikely given that the vowels were presented in isolated /h_d/ words in which only duration differences in the two consonants would be available as speaking rate cues.

Morrison (2002a) had Japanese listeners identify Scottish / Northern English vowels in isolated CVC words in which both consonants were either voiced orvoiceless. The correct-identification rate for English /ɪ/ was 93%, and for English /i/ in the voiced consonant condition it was 95%. In contrast, English /i/ in the voiceless consonant condition had correct-identification rate of only 27%; it was identified as /ɪ/ at a rate of 72%. Given that English vowels are typically longer before voiced than before voiceless consonants (Chen 1970), Morrison (2002a) proposed that: English /i/ is assimilated to Japanese /iː/ before a voiced consonant but assimilated to Japanese /i/ before a voiceless consonant; that English /ɪ/ is assimilated to Japanese /i/ before both a voiced and voiceless consonant; that English vowels assimilated to Japanese /iː/ are identified as English /i/, and that English vowels assimilated to Japanese /i/ are identified as English /ɪ/. This requires that the Japanese long-short vowel boundary fall between the longest duration for an English /i/ stimulus in the voiced consonant condition or for an English /ɪ/ stimulus in either condition (whichever be the greater), and the shortest duration for an English /i/ stimulus in the voiceless consonant condition. Comparison of Morrison's (2002b) stimulus durations with Japanese vowel production data from Tsukada (1999) and the English stimulus durations from Ingram & Park (1997) (Figure1) suggested that the Japanese long-short vowel duration boundary does indeed fall within this range.

Under this hypothesis, Morrison (2002a) proposed that the near perfect correct identification rates for English /i/ and /ɪ/ in Ingram & Park (1997) were not due to

(Figure 1) Duration ranges for English /i/ before voiced consonants and English /ɪ/ before voiced and voiceless consonants (left), and for English /i/ before voiced consonants (right) from Morrison (2002a). Duration ranges for Japanese /i/ (left) and /iː/ (right) from Tsukada (1999); and for English /ɪ/ (left) and /i/ (right) from Ingram & Park (1997). The dashed line represents a categorical boundary that would account for the identification patterns in Ingram & Park (1997) and Morrison (2002a).



the Japanese listeners normalising for speaking rate. Rather, the results may be due to the English /ɪ/ stimuli in Ingram & Park (1997) being short enough to be assimilated to Japanese /i/, and the English /i/ stimuli being long enough to be assimilated to Japanese /iː/. Although there was a large difference in vowel durations, the duration ranges for vowels within each English category did not straddle the Japanese long-short categorical boundary due to the fact that Ingram & Park had used context in which the vowels occurred before a voiced consonant.

The present paper will test Morrison's (2002a) hypothesis by determining whether Japanese listeners identify English /i/ and /ɪ/ using the same duration criteria that they use to identify Japanese /iː/ and /i/.

1. Method

1.1. Participants

Canadian English participants (3 women, 4 men) had lived until age 16, and immediately prior to data collection, in an Anglophone region of Canada west of Quebec. The group had a mean age of 30 (range 25-38).

Japanese participants (4 women, 3 men) were undergraduate students attending an academic exchange in Vancouver. Their programme included content-based language instruction but no courses in English pronunciation. They had lived in Vancouver for a period of less than 2 months prior to data collection. They had lived until age 16 in Japan, and as children, had not been immersed in a language other than Japanese. They had never lived outside Japan for a continuous period of more than three months, and had not lived in an English speaking country during the year prior to their arrival in Canada. The group had a mean age of 20 (range 19-22). They had studied English in school for 7-9 years, starting at age 13 or 14.

1.2. Stimuli

1.2.1 Target segments and carrier sentences

The English target words ("bit, beat, bid, bead" /bɪt bit bɪd bid/) were chosen to exemplify the English vowels /i/ and /ɪ/ in a stressed position followed by both a voiced and voiceless plosive.

Since responses to English and Japanese stimuli were to be compared, the segment immediately following the English target word had to match the parallel segment in the Japanese stimuli. The segment selected was /s/, resulting in voiceless affricated release for English /t/ and /d/, corresponding to the Japanese affricate [t͡s]. It was not possible in Japanese to exactly match the English voiced alveolar plosive with voiceless affricated release; in Japanese both parts of the affricate are voiced, i.e. [d͡z]. The Japanese target words read by the model speaker were therefore "bitsu, biitsu, bittsu, bidzu, biidzu" exemplifying possible combinations of short and long vowels with voiced and voiceless singleton and geminate consonants. Although phonotactically rare, the words "biddzu, biittsu, biiddzu" were added as possible response categories in the perception test since some of the stimuli had properties that might suggest these responses.

Matching English and Japanese carrier sentences were "What they're wearing are <u>beat</u> suits," and "*motteiru mono wa <u>bitsu</u>utsu*" ('What I have is a [nonsense word]-suit'). Strange et al. (1998) found that the use of a carrier sentence, as compared to the presentation of target words in isolation, can have a significant impact on perception, such that long English vowels were assimilated to Japanese long vowels more often in a sentence condition than in an isolated word condition (85 versus 42%). They hypothesised that the participants had used the rhythmic pattern of the sentence to calibrate their duration-based responses.

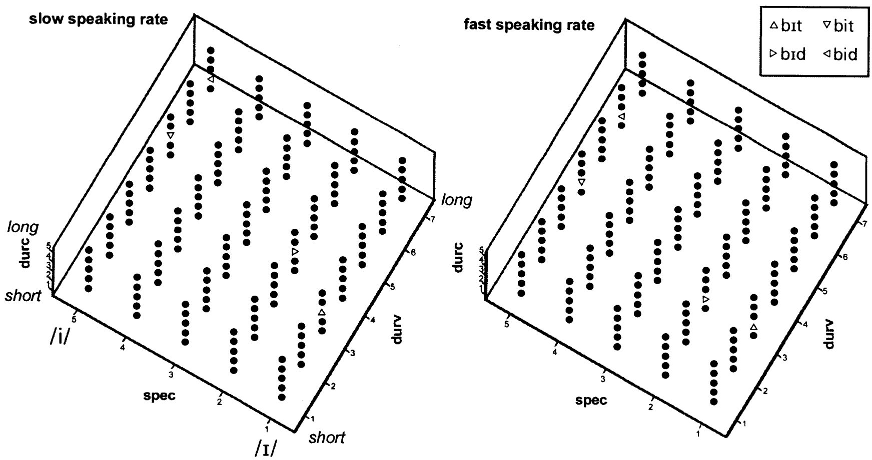
1.2.2 Acoustic properties of stimuli

Perceptual stimuli were constructed using edited natural speech based on the productions of two model speakers, a 34 year-old male monolingual English speaker who had lived most of his life in Vancouver, and a 38 year-old male native speaker of Japanese who had lived most of his life in Fukuoka. Recordings were made in a soundproofed room using a Sony MZS-R5ST Mini Disc recorder and a Sony ECM-MS907 microphone. The speakers read randomised lists of stimulus sentences at normal, slow, and fast speaking rates. The recordings were digitally transferred to computer, were analysed, and were used to create the stimuli using Praat computer software. Details of the procedure are given in Morrison (2002b).

A multidimensional continuum was created in which acoustic properties varied along the following dimensions: vowel spectra (5 points), vowel duration (7 points), plosive closure duration (5 points), and speaking rate (2 points). The dimensions can be visualised as the cuboids shown in Figure 2.

The F1, F2, and F3 values at the midpoints of the vowels in the English stimuli

(Figure 2) Cuboids representing the dimensions along which the English stimuli vary. Vowel duration (durv) from front to back: 1 is shortest and 7 longest. Vowel spectral properties (spec) from right to left: 1 is most /ɪ/-like and 5 most /i/-like. Consonant closure duration (durc) from bottom to top: 1 is shortest and 5 longest. Speaking rate: left cube is slow and right cube is fast. The triangles represent the position of typical "bit, beat, bid, bead" produced by the model speaker.



are given in Table 1. The spectral properties of spectral points1 and 5 were the mean spectral values from the model English speaker's /i/ and /ɪ/. The five points were evenly spaced on the mel scale.

The vowel and consonant durations of the stimuli are given in Table 2. The range of durations for the English stimuli was expanded to cover the same range as the Japanese durations. The properties of the duration endpoints were based on the mean values from contexts resulting in the longest and shortest vowels and consonants produced by the Japanese model speaker. The duration points were evenly spaced on a logarithmic scale. Based on mean values from the English and Japanese model speakers, sentence durations for the slow and fast speaking rates were 1.541 and 1.265 seconds respectively (excluding the duration of the target segments).

The Japanese continua had the same duration properties as the English continuum, but had no spectral dimension. The spectral values for the Japanese stimuli (F1, F2, F3 = 318, 2126, 3115Hz) were those of a model vowel which had spectral values close to the mean values produced by the model speaker. Since Japanese phonemic voicing is always cued by phonetic voicing, two continua were created, one with, and one without phonetic voicing. The dimensions can be visualised as the rectangular matrices shown in Figure 3.

Stimulus vowels had durations and spectral properties within 2.5% of the values given in Tables 2 and 3. Consonant closure durations for English consonants and Japanese voiceless consonants consisted of silence.

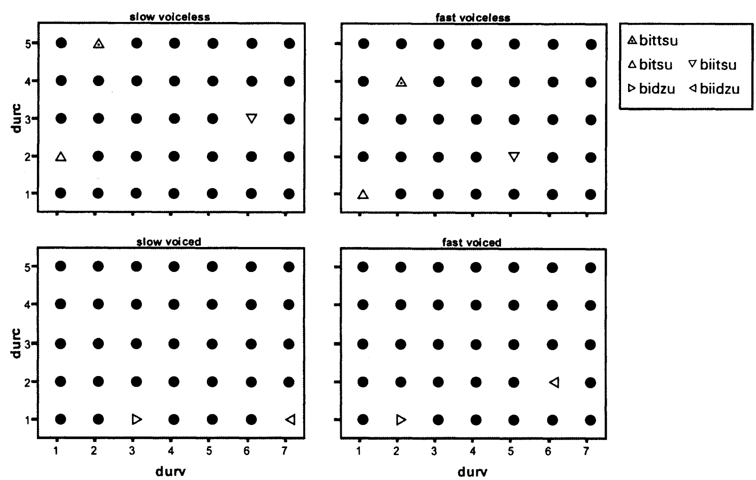
(Table 1) Spectral values at the midpoints of the vowels in the English stimuli.

Endpoint vowel	Dimension point	F1		F2		F3	
		Hz	mel	Hz	mel	Hz	mel
ɪ	1	410	496	1700	1433	2465	1793
	2	370	454	1834	1503	2569	1836
	3	330	412	1974	1572	2676	1878
	4	292	370	2121	1642	2786	1921
i	5	255	328	2275	1712	2900	1964

(Table 2) Durations of the vowels and consonants in the English and Japanese stimuli.

Dimension point	1	2	3	4	5	6	7
Vowel duration (ms)	60	72	86	104	124	149	179
Consonant closure duration (ms)	40	58	84	122	177		

(Figure 3) Rectangular matrices representing the dimensions along which the Japanese stimuli vary. Vowel duration (durv) from left to right: 1 is shortest and 7 longest. Consonant closure duration (durc) from bottom to top: 1 is shortest and 5 longest. Speaking rate: left rectangles are slow right rectangles fast. Consonant voicing: top rectangles are voiceless and bottom rectangles voiced. The triangles represent the position of typical “bittsu, bitsu, biitsu, bidzu, biidzu” produced by the model speaker.



1.3 Data Collection

Participants were tested individually in a soundproofed room. They listened to the stimuli presented in random order via MEDS computer software over Optimus HP340 headphones, and responded to each English stimulus by clicking on one of five pictures on the computer screen. Four of the pictures represented the words "bit, beat, bid, bead." The fifth picture, an "X," was a null response which the participants were instructed to use if they heard a word other than one of the four target words. Pictures were used in order to avoid potential confusion from orthographic representations of the target words. Participants were trained to interpret the pictures before the perception test. The computer played a stimulus once and did not proceed until the participant had given a response. The inter-trial interval was 500ms. The order of the pictures was assigned randomly for each trial. A subset of 24 stimuli was used as a warm-up before the full set of 350 stimuli, each of which was identified once.

The Japanese perception test was conducted approximately one month after the English test using the same procedures as the English test, except that the eight response options "bittsu, biitsu, bitsu, biittsu, biddzu, bidzu, biidzu, biiddzu" were written in *katakana* (a Japanese orthographic system). The full set of 140 stimuli were presented in two randomisations.

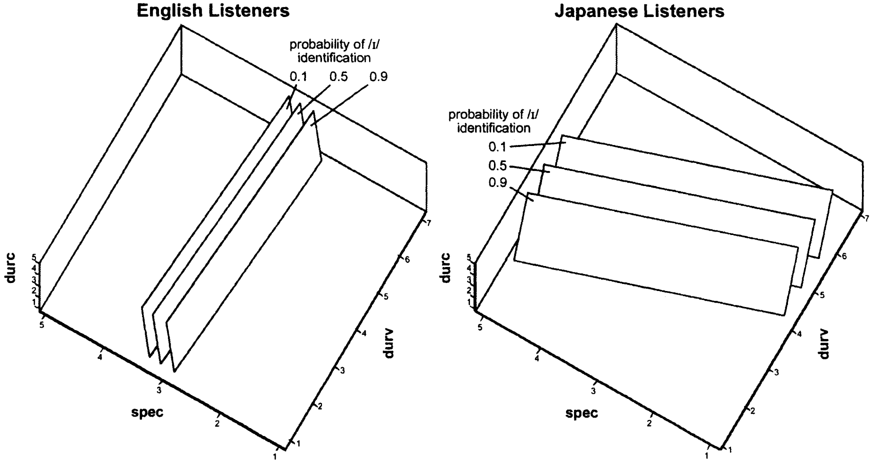
2. Results and Discussion

Categorical boundaries in the listeners' responses were modelled using discriminant analyses. Vowel spectral properties, vowel duration, consonant duration, and speaking rate were candidate variables entered stepwise into the model if they produced a between-group F ratio significance of .05 or less. Discriminant function coefficients and group centroid values were used to calculate 0.1, 0.5, and 0.9 probability of /ɪ/ identification points which were then graphed in terms of the original variables. The 0.5 probability points indicate the position of the categorical boundary. A small distance between the 0.1 and 0.9 probability points indicates a sharp boundary, and a large distance indicates a fuzzy boundary. Due to space limitations, only probability point values, variable weightings expressed as percentages, and cross-validated correct classification rates are reported here. Details of the analysis procedure and full statistical results are given in Morrison (2002a).

2.1. Canadian English Listeners

The discriminant analysis was highly successful at categorising the English listeners' responses (the cross-validated correct classification rate was 95.5%). All variables except speaking rate were entered into the model. This suggests that speaking rate did not affect the English listeners' vowel perception, and that they identified the vowels primarily according to their spectral properties. The relative weighting given to the variables were 84% for vowel spectral properties, 11% for vowel duration, and

(Figure 4) Planes describing the boundary between English /i/ and /i:/ derived form the discriminant analyses carried out on the identification responses of native English and Japanese listeners.



5% for consonant duration. The small effect for vowel duration is predictable since tense /i/ is longer than lax /ɪ/ all else being equal. The slight effect for consonant closure duration is also predictable since a long consonant closure will create the perceptual contrast effect of a shorter vowel (Kluender, Diehl, & Wright, 1988).

The categorical boundary is graphed in Figure 4. The planes are almost perpendicular to the *spec* axis, indicating that the native English participants based their vowel identification primarily on spectral properties. The mean spectral value of the boundary between /i/ and /ɪ/, was 3.07 (F1, F2, F3 = 415, 1568, 1875mel; 333, 1964, 2668Hz). The spectral distance between 0.1 and 0.9 probability of /ɪ/ identification was 0.44 (ΔF1, ΔF2, ΔF3 = 18, -30, -19mel; 17, -62, -47Hz, Hertz values calculated at the centre of the planes) indicating a sharp boundary.

2.2. Japanese Listeners

2.2.1. English Test

The discriminant analysis was moderately successful at categorising the Japanese listeners' responses to the English stimuli (the cross-validated correct classification rate was 85.9%). All variables except speaking rate were entered into the model. This suggests that speaking rate did not affect the Japanese listeners' vowel perception, and is in accordance with the results of other studies (Strange et al., 1998; Guion et al., 2002, Ingram & Park, 1997) that found little effect for speaking rate on Japanese listeners' identification of English lax versus tense vowels. The Japanese listeners identified the vowels primarily according to vowel duration but also made considerable use of vowel spectral properties. The relative weightings were 63% for vowel duration, 28% for vowel spectral properties, and 9% for consonant duration. The predominantly duration-based identification pattern was radically different to the almost exclusively spectral identification pattern used by the native English speakers.

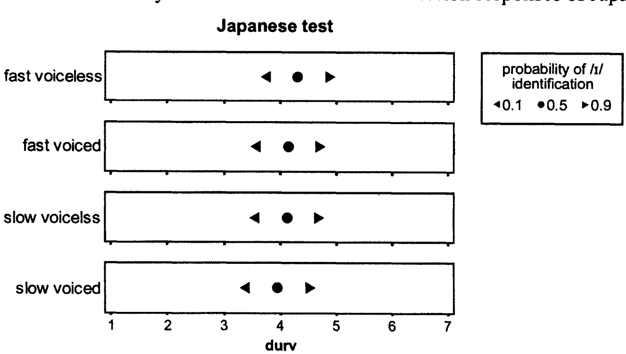
The Japanese listeners' English /i/-/ɪ/ categorical boundary is graphed in Figure 4. The orientation of the planes indicates that the Japanese listeners based their identification primarily on vowel duration. The mean duration of the boundary was 4.22 (108ms). The vowel duration distance between 0.1 and 0.9 probability of /ɪ/ identification was 1.49 (29ms at the centre of the planes) indicating a relatively sharp boundary. The slopes due to the use of secondary cues resulted in a considerable duration range for the boundary: The duration range for the 0.5 plane was 2.98 to 5.46 (86 to 135 ms), and for the 0.1 and 0.9 planes 2.24 to 6.20 (75 to 155ms).

The Japanese participants' use of spectral properties may be a result of learning English. Although still relying primarily on transfer of L1 duration criteria they may have realised that English /i/ and /ɪ/ differ spectrally and shifted their cue weighting towards greater use of spectral properties. L1 experience, however, may also have caused sensitivity to spectral properties: Fitzgerald (1996) found that Japanese speakers produced long vowels that were more peripheral in the vowel space than short vowels. He found that, with only F1 and F2 as independent variables, a discriminant analysis on all vowels correctly classified 77% of /i:/ and 69% of /i/ vowels. The existence of spectral differences leads to the possibility that spectral properties may in fact be a secondary cue differentiating long-short vowel pairs in Japanese.

2.2.2. Japanese Test

The discriminant analysis was moderately successful at categorising the Japanese listeners' responses to the Japanese stimuli (the cross-validated correct classification rate was 88.6%). The Japanese listeners based their identification of /i/ and /i:/

(Figure 5) Points describing the categorical boundary between Japanese /i/ and /i:/ derived from the discriminant analysis carried out on the identification responses of Japanese participants.



primarily on vowel duration. The relative weightings given to the variables were 81% for vowel duration, 13% for consonant voicing, and 5% for speaking rate. The small effect due to consonant voicing is consistent with the finding of Tsukada (1999) that the consonant voicing effect on vowel duration in Japanese was small compared to the size of the effect in English. The small effect for speaking rate is consistent with the results of Toda (1999) who found only a small effect for speaking rate on Japanese listeners' identification of Japanese long versus short vowels.

The categorical boundary points are graphed in Figure 5. The mean duration of the boundary between /i/ and /iː/ was 4.07 (105ms). The vowel duration distance between 0.1 and 0.9 probability of /i/ identification was 1.16 (22ms centred on the mean 0.5 probability) indicating a relatively sharp boundary. Due to the effects from consonant voicing and speaking rate, the duration range for the 0.5 planes was 3.89 to 4.25 (102 to 108 ms), and for the 0.1 and 0.9 planes 3.31 to 4.83 (91 to 121ms).

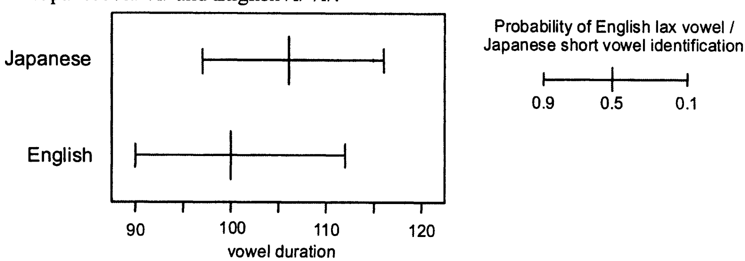
2.2.3. Comparison of English and Japanese Boundaries

Since the English stimuli all had silent consonant closures, comparison was made between the English boundary and the Japanese /iː/-/i/ boundary based only on the responses to silent-consonant-closure stimuli. Likewise, the Japanese boundary was compared with the English /i/-/ɪ/ boundary based on stimuli with the same spectral properties as the Japanese stimuli. The formant frequencies of the Japanese stimuli were closest in terms of F1 and F2 to the English stimuli with spectral value 4.

A discriminant-analysis model based on responses to Japanese silent-consonant-closure stimuli resulted in a cross-validated correct classification rate of 89.9%, and a cue weighting of 92% for vowel duration and 8% for consonant duration. The mean Japanese /i/-/iː/ categorical boundary was 4.13 (106ms). The distance between 0.1 and 0.9 probability of /i/ identification was 0.99 (19ms centred on the mean 0.5 probability) indicating a relatively sharp boundary. Due to the effects from consonant closure, the duration range for the 0.5 line was 3.98 to 4.29 (103 to 109ms), and for the 0.1 and 0.9 lines 3.31 to 4.83 (94 to 120ms).

A discriminant-analysis model based on the Japanese listeners' responses to English /ɪ/ and /i/ stimuli with spectral value 4 resulted in a cross-validated correct

(Figure 6) Comparison of the Japanese listeners' mean categorical duration boundary (in milliseconds) for Japanese /iː/-/i/ and English /ɪ/-/i/.



classification rate of 87.9%, and a cue weighting of 84% for vowel duration and 16% for consonant duration. The mean English /ɪ/-/i/ categorical boundary was 3.82 (100ms). The distance between 0.1 and 0.9 probability of /ɪ/ identification was 1.20 (22ms centred on the mean 0.5 probability) indicating a relatively sharp boundary. Due to the effects from consonant closure, the duration range for the 0.5 line was 3.45 to 4.18 (94 to 107 ms), and for the 0.1 and 0.9 lines 2.85 to 4.78 (84 to 119ms).

A graphical representation of the Japanese listeners' mean Japanese /iː/-/i/ and English /ɪ/-/i/ boundary is given in Figure 6. The 6ms difference between the mean English and Japanese 0.5 probability points falls within both the English and Japanese 0.1 to 0.9 probability ranges. From these results it can be concluded that the Japanese participants' identified English /ɪ/ and /i/, and Japanese /i/ and /iː/ using a very similar or identical categorical vowel duration boundary.

2.3. Conclusion

The present study found that Japanese listeners used the same duration criteria to identify English /ɪ/ and /i/ and Japanese /i/ and /iː/. This finding supports the hypothesis that Japanese listeners identify English /ɪ/ and /i/ via assimilation to Japanese /i/ and /iː/ proposed by Morrison (2002b).

Speaking rate was found to have no effect on the Japanese listeners' English /ɪ/-/i/ categorical boundary and only minimal effect on their Japanese /i/-/iː/ boundary. This supports the proposal that the near perfect correct identification rate for English /ɪ/ and /i/ in Ingram & Park (1997) was not due to speaking rate normalisation.

References

Akamatsu, Tsutomu. 1997. *Japanese phonetics: Theory and practice*. München, Newcastle: Lincom Europa.

Chen, Matthew. 1970. Vowel length variation as a function of the voicing of the following consonant environment. *Phonetica 22*: 129-159.

Guion, Susan., Flege, James E., Akahane-Yamada, Reiko., & Pruitt, Jesica C. 2002. *Predicting Japanese adults' discrimination of English and Japanese vowels.* Manuscript in preparation.

Ingram, John C. L., & Park, See-Gyoon. 1997. Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics 25*: 343-370.

Kluender, Keith R. Diel, Randy L., & Wright, Beverly A. 1988. Vowel-length differences before voiced and voiceless consonants: An auditory explanation. *Journal of Phonetics 16*: 153-169.

Morrison, Geoffrey Stewart. 2002a. Perception of English /i/ and /ɪ/ by Japanese listeners. In S. Oh, N. Sawai, K. Shiobara, & R. Wojak (eds.) *University of*

British Columbia Working Papers in Linguistics: Vol. 8 (pp 113-131). Vancouver, BC: University of British Columbia, Department of Linguistics.

Morrison, Geoffrey Stewart. 2002b. *Effects of L1 duration experience on Japanese and Spanish listeners' perception of English high front vowels*. MA thesis in preparation.

Nearey, Terrance M., & Assmann, Peter F. 1986. Modelling the role of inherent spectral change in vowel identification. *Journal of the Acoustical Society of America, 80*, 1297-1308.

Strange, Winifred, Akahane-Yamada, Reiko., Kubo, Reiko, Trent, Sonja. A., Nishi, Kanae, & Jenkins, James J. 1998. Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics 26*: 311-344.

Toda, Takako K. 1999. Development of speech discrimination by learners of Japanese as a second language: Remarks from a longitudinal study. In P. Robinson (Ed.), *Representation and Process: Proceedings of the 3rd Pacific Second Language Research Forum: Vol. 1* (pp 207-233). Tokyo: Pacific Second Language Research Forum.

Tsukada, Kimiko. 1999. *An acoustic phonetic analysis of Japanese-accented English*. Unpublished PhD dissertation, Macquarie University, Sydney, Australia.

Department of Linguistics
University of Ottawa
P.O. Box 450, Stn. A
Ottawa, Ontario K1N 6N5
CANADA

gsm@alumni.sfu.ca

Sentential Negation and Verb Movement in Bantu Languages*

DEO NGONYANI
Michigan State University

0. Introduction

This paper explores a well-known asymmetry between negation marking main clauses and subordinate clauses in Bantu languages (Güldemann 1999:551, Meeussen 1967:114). It is noted that pre-initial negation marking is usually restricted to main clauses, while post-initial marking is rarely restricted. Various studies have explored the diachronic origins of the various strategies. This paper focuses on how the strategies are constrained by clause structure. It is argued that negation projects a NegP as an element of Infl. Asymmetrical negation marking is due to two NegPs, one selecting TP, and the other selected by TP.

The paper is organized into five sections. Section 1 presents examples of strategies of marking negation in main clauses, while Section 2 contains examples of negation in relative clauses. Section 3 discusses verb movement to C° and Section 4 proposes an analysis of NegPs. Section 5 concludes the discussion.

1. Strategies of Marking Negation in Main Clauses

Bantu languages exhibit a wide range of negation marking strategies in matrix clauses. Most languages mark negation by means of an affix on the verb. Affixal negation markers are located in one or two slots on the Bantu verbal template.

(1) Elements of the Bantu Verb (adopted from Meeussen 1967:108)

	Pre-initial	Initial	Post-initial	Tense & Aspect	Infix	Stem	Final	Post-final
Function	REL/NEG	SM	NEG	TNS/ASP	OM	Verb	ASP	NEG
Negation	+		+					+

The three slots for negation marking together with independent negative words derive six strategies of marking negation in main clauses as illustrated below.

* This paper is based on work supported in part by the Michigan State University Intramural Research Grant Program.

(i) Preverbal particle as in Hung'an H.42 (Takizala 1972:127)[1]

(2) lo i-meen kiti[2]
NEG I-saw 7-chair
'I didn't see the chair.'

(ii) Pre-initial Affix as in Luba L.31 (Yukawa 1992:309, 317)

(3) a. tw-áka-mu-p-a
we-PT-1OM-give-FV
'we gave him/her'
b. ká-tw-aká-mú-p-a
NEG-we-PT-1OM-give-FV
'we did not give him/her'

(iii) Post-initial Prefix as in Nyakyusa (Ngonyani personal notes)

(4) a. tw-a-piy-ile
We-PT-cook-PF
'We cooked.'
b. tu-ka-a-piy-ile
we-NEG-PT-cook-PF
'We did not cook.'

(iv) Post-final Suffix as in Nkoya L.62 (Yukawa 1987:138)

(5) a. na-mon-ene
I-see-PT
'I saw.'
b. ná-mon-enĕ-ha
I-see-PT-NEG
'I did not see.'

[1] A letter and a number after the name of a language refer to the reference number assigned in Guthrie's 1968-71 classification of Bantu languages.

[2] Abbreviations:

CON	Conditional Marker	DEM	Demonstrative	FV	Final Vowel
FT	Future Tense	INF	Infinitive Marker	NEG	Negation
OM	Object Marker	PF	Perfective Marker	PT	Past Tense
REL	Relative Marker	SM	Subject Marker	SUBJ	Subjunctive Marker

The numbers in the glosses stand for noun classes.

(v) Post-verbal Particle as in Ngoni P.13 (Ngonyani field notes)

(6) a. m-geni_ a- tol-ili u-gimbi
1-guest 1SM-take-PF 14-beer
'The guest took beer.'
b. m-geni a- tol-i <u>lepa</u> u-gimbi
1-guest 1SM-PT-take-FV NEG 14-beer
'The guest did not take beer.'

(vi) Pre-initial Prefix + Affix as in Modern Standard Swahili G.42[3]

(7) a. tu-li-ondok-a
we-PT-leave-FV
'We left.'
b. <u>ha</u>-tu-<u>ku</u>-ondok-a
NEG-we-NEG+PT-leave-FV
'We did not leave'

The nature of double marking of negation as found in Swahili, Zulu (S.42), Luvale (K.14), Venda (S.21) and other languages is not discussed in this paper and is deferred to future research.

The interesting feature is that when it comes to relative clauses, languages exhibit the same patterns. For example, languages with preverbal negative particles have periphrastic negative markers in relative clauses regardless of whether they are closely related or not. The patterns of negation in relative clauses are presented in the following section.

2. Negative Marking Strategies in Relative Clauses

The wide range of strategies is narrowed in relative clauses as shown in this section. Using the sample languages cited in the preceding section, we notice some interesting restrictions.

In root clauses, Hung'an negation is marked by a preverbal negative particle. In relative clauses, negation is marked with a periphrastic element that is inflected for tense and subject agreement. The same applies to Mbukushu (K.43).

(8) a. kit ki a-swiim-in Kipes zoon (Takizala 1973:128)
7chair 9REL 1SM–buy-PT Kipese yesterday
'the chair that Kipese bought yesterday'
b. kit ki <u>a-khoon-in</u> Kipes ku-suum
7chair 7REL 1SM-fail-past Kipese INF-buy
'the chair that Kipese didn't buy'
*kit ki lo a-swiim-in Kipes

[3] Standard Swahili examples are provided by the author.

The negative marker appears between the relative marker and the subject of the relative clause.

Luba and languages such as Ganda (J.15), Haya (J.22), Shambala (G.23) and Nyoro-Tooro (J.11) mark negation in pre-initial morphological slot in main clauses, but use the post-initial slot in relative clauses as shown the Luba example below.

(9) a. mú-úntú ú-twa-aká-biik-íle (Yukawa 1992:341)
1-person 1REL-we-PT-call-PF
'the person whom we called'
b. mú-úntú ú-tu-<u>di</u>-ká-tu-yi-biik-íle
1-person 1REL-we-NEG-PT-we-3OM-call-PF
'the person whom we did not call'

Languages that use the post-initial slot in matrix clauses, such as Pangwa (G.64), Lozi (K.21), and Tuki (A.64), use the same strategy in the relative clauses as illustrated with Nyakyusa below.

(10) a. un-dindwana u-yo a-piy-ile mbatata (Ngonyani field notes)
1-girl 1-REL 1SM-cook-PF 10potato
'the girl who cooked potatoes'
b. un-dindwana u-yo a-<u>ka</u>-piy-ile mbatata
1-girl 1-REL 1SM-NEG-cook-PF 10potato
'the girl who did not cook potatoes'

The negative marker is in the post-initial slot, the same as in the matrix clauses.

It was noted that Nkoya marks negation by means of a postfinal marker in the main clause. In relative clauses, Nkoya marks negation by means of an inflected auxiliary word.

(11) a. mu-ntu é-mon-ené (Yukawa 1987:161)
1-person 1REL-see-PT
'a person whom he saw'
b. mu-ntu bá-bul-ilé ku-môn-a
1-person ?2SM-lack-PF INF-see-FV
'the person who they did not see'

The periphrastic negative marker is derived from *kubula* 'to lack'. Languages that use a post-final negation marker in matrix clauses use a periphrastic form in the relative clause.

Kingoni, Kindendeule, and Kimatuumbi mark negation the same way in the root clauses as well as in relative clauses, by means of a postverbal particle.

(12) a. m-geni_ mwe-a- tol-ili u-gimbi (Ngonyani: field notes)
1-guest 1REL-1SM-take-PF 14-beer
'the guest who took the beer.'
b. m-geni mwe-a- tol-ili lepa u-gimbi
1-guest 1REL-1SM-PT-take-FV NEG 14-beer
'the guest did not take beer.'

The asymmetries show that while there are six different strategies for marking negation in the main clauses, there are only three for relative clauses. In the following sections, I attempt to provide a structural account for the asymmetries.

3. V-to-I-to-C

The asymmetry is reminiscent of word order asymmetries found in V2 languages such as Swedish. Consider the position of negation in root clauses and relative clauses in Swedish.

(13) a. Ulf köpte inte boken (Holmberg and Platzach 1995:45)
Ulf bought not the-book
b. att Ulf inte köpte boken
that Ulf not bought the-book

In the root clause (13a), the negative appears after the verb while in the subordinate clause (13b), the negative marker appears before the verb. This asymmetry is accounted for by verb movement to C° in matrix clauses, while in the subordinate clause C° is filled by the complementizer. Therefore, the verb does not move to C° in the subordinate clause. In languages where there is no V movement to C° such as English, there is no asymmetry between negation in main clauses and subordinate clauses. The negation marker appears between the auxiliary and the verb in both clauses.

A good starting point for Bantu languages, therefore, is to investigate the contrast between languages that have verb movement to C° and those that do not have such movement. Based on evidence from VP ellipsis and VP adverbs in Swahili and Ndendeule[4], Ngonyani (1996, 2000) argues that the verb moves to Infl. Demuth & Harford (1999) find evidence of verb movement to C in object relative clauses of several Bantu languages. They show that Bantu languages with independent relative markers do not exhibit subject-verb inversion, while languages with an affixal relative marker do as the following contrast between Sotho (14a) and Shona (14b) shows (Demuth & Harford 1999:42).

(14) a. Setula seo basadi ba-se-rek-ile-ng kajeno
7chair 7REL 2women 2AGR-7OBJ-buy-PERF-RL today
'the chair which the women bought today'

[4] Not classified by Guthrie (1967-71) but its closest relative is Ngindo (Ngonyani 2001a).

b. Mbatya dza-v-aka-son-era vakadzi mw-enga
10clothes 10REL-2AGR-TN-sew-APL 2women 1bride
'clothes which the women sewed for the bride'

The SV order is maintained in Sotho where the relative marker is an independent word, while in Shona, where the relative marker is an affix, there is subject-verb inversion. Assuming that the relative marker occupies C° in Sesotho, there is no movement of the verb because the landing site is occupied. In (14b), however, the position is empty and subsequently filled by the verb.

It is tempting to hypothesize that in languages without subject-verb inversion (that is languages without verb movement to C°), there will be no asymmetry between main clauses and relative clauses because the verb does not have to cross Neg. This, however, is not borne out, as the following example from Ganda shows.

(15) Omusajja Petero gwe a-labye mu-somesa (Walusimbi 1996:37)
man Peter REL 1SM-see-PF 1-teacher
'The man that Peter has seen is a teacher.'

This example of object relativization shows two important features of Ganda relative clauses: (a) an independent word for a relative marker, and (b) no subject-object inversion. These two facts suggest there is no verb movement to C°. However, negation marking is asymmetrical as the following examples show.

(16) a. ye a-<u>ta</u>-kol-a (Ashton et. al. 1954:138)
1REL 1SM-NEG-work-FV
'(he) who does not work'
b. ebi-kopo <u>te</u>-byt-ise
8-cup NEG-crack-PF
'The cups are not cracked.'

Example (16a) shows that the negative marker in the relative clause appears after the subject marker (post-initial) different from what is found in the root verbs (16b), a pre-initial affix. Therefore, the asymmetries cannot be explained in a straightforward manner simply by appealing to V-to-I-to-C.

A closer look at the data reveals that in relative clauses, the relative marker is always associated with finiteness and more specifically TP. The relative marker, whether prosodically independent or not, is followed by a subject marker and a tense marker on the verb. Assuming an Agr-less clause structure (Chomsky 1995), we find the relative marker adjacent to the tense marker. This suggests selectional relations between the C of the relative clause and TP. In languages with pre-initial negation or independent particle before the verb in main clauses, the structure cannot be embedded as a relative clause because the C of the relative marker will fail to govern TP. Another strategy is used instead. The selection of

TP by C is very widespread among Bantu languages. Let us consider how this feature accounts for the asymmetries in the various languages.

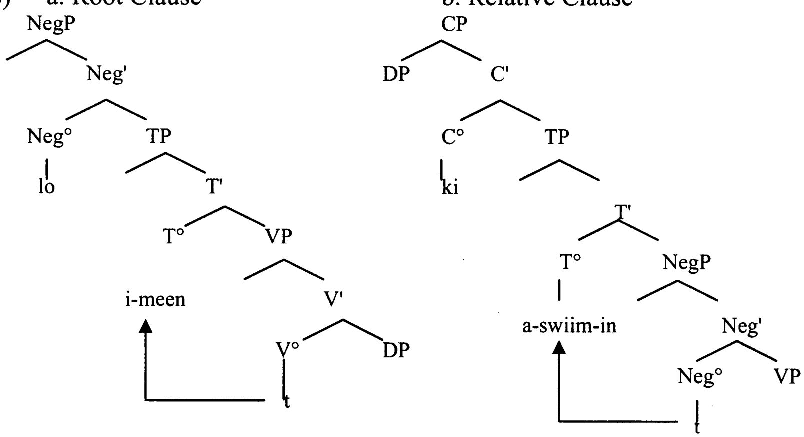
4. Analysis

The correlation between the tense marker, the relative marker, and the negative marker suggests that Neg blocks government of TP by C°. Therefore, negation projects a NegP which selects TP. In subordinate clauses, the negative marker cannot intervene between the relative marker and the tense marker. The asymmetry between matrix clauses and relative clauses can be characterized as NegP selecting TP in matrix clauses, and NegP selected by TP in relative clauses. This is illustrated in the following abbreviated tree diagrams for Hungan sentences repeated here for verbs cited earlier.

(17) a. lo i-meen kiti
NEG I-saw 7-chair
'I didn't see the chair.'

b. kit ki a-khoon-in Kipes ku-suum
7chair 7REL 1SM-fail-past Kipese INF-buy
'the chair that Kipese didn't buy'

(18) a. Root Clause b. Relative Clause



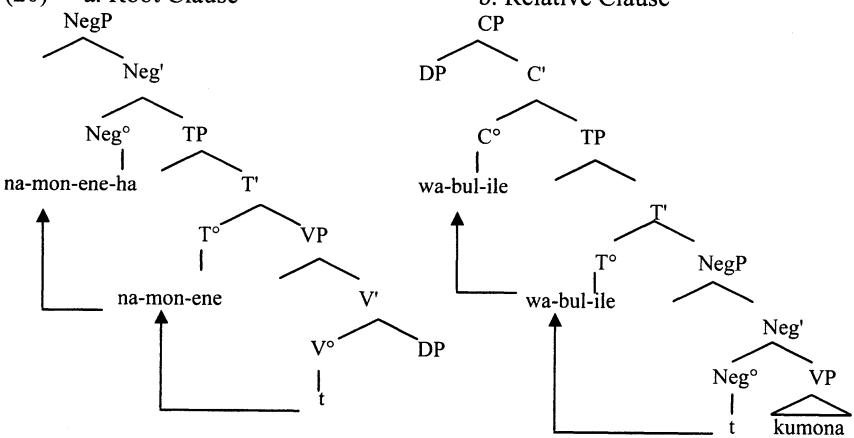
In the matrix clause, Neg appears as the highest Infl which selects TP. This accounts for the preverbal negation *lo*. The verb, which is lower than TP, raises and incorporates onto T in an antisymmetric fashion (Kayne 1994). In the relative clause, however, CP does not select NegP. It selects TP, and negation has to be expressed in the form of a periphrastic marker generated as a verb and raised to T in the regular V-to-I manner.

The inability of the negative marker to license TP is also found in languages with pre-initial negation such as Oruhaya and postfinal negation such as Nkoya. Let us consider Nkoya negation since it requires a periphrastic negative marker in relative clauses in the same way Hung'an does. In matrix clauses, negation selects TP. This negative marker attracts Tense resulting in the incorporation of the inflected verb to Neg. This explains the appearance of the negative *ha* in verb final position. In relative clauses, however, CP selects TP. NegP between CP and TP blocks this relation. Thus, the regular negation is eliminated and negation is introduced by a negative verb *buli* 'fail'. This is shown in the following pair of tree diagrams.

(19) a. ná-mon-enĕ-ha
I-see-PT-NEG
'I did not see.'

b. mu-ntu bá-bul-ilé ku-môn-a
1-person ?2SM-lack-PF INF-see-FV
'the person whom they did not see'

(20) a. Root Clause b. Relative Clause



This account provides a simple explanation for the appearance of the negative suffix at the end of the verb phrase as in (20a). In such cases, the entire VP raises to Spec, NegP.

(21) a. ná-mon-ene-há ngombe
I-see-PT-NEG 9cow
'I did not see the cow'

b. Ná-mon-ene ngombĕ-ha
I-see-PT 9cow-NEG
'I did not see the cow.'

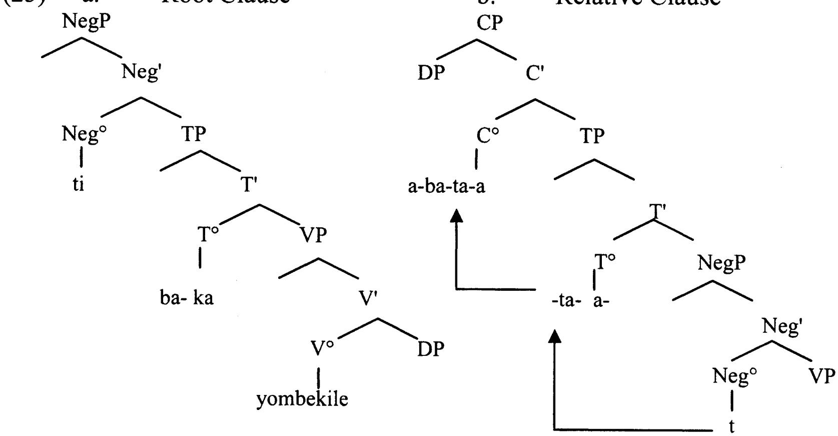
These two sentences show the two alternative slots for the negative affixes, namely, as a verbal suffix or as an affix on the final element of the verb phrase.

Note that the two negative markers in the matrix clauses and the relative clauses are phonologically different suggesting different origins. The different origins explain also the different affixes used in Haya where negation is a pre-initial affix is different from the post-initial negation marker of the relative clause (Rubanza 1988:104-105).

(22) a. Aba-ntu ba Kanyigo ti-ba-ka-yombek-ile shule
2-people of Kanyigo NEG-2SM-RP-build-PF school
'Kanyigo people have not build the school.'

b. A-ba-ta-a-shom-e ti-ba-a-sing-e mitiani
REL-2SM-NEG-IP-study-SUBJ NEG-2SM-IP-pass-SUBJ exams
'those who will not study will not pass the exams'

(23) a. Root Clause b. Relative Clause



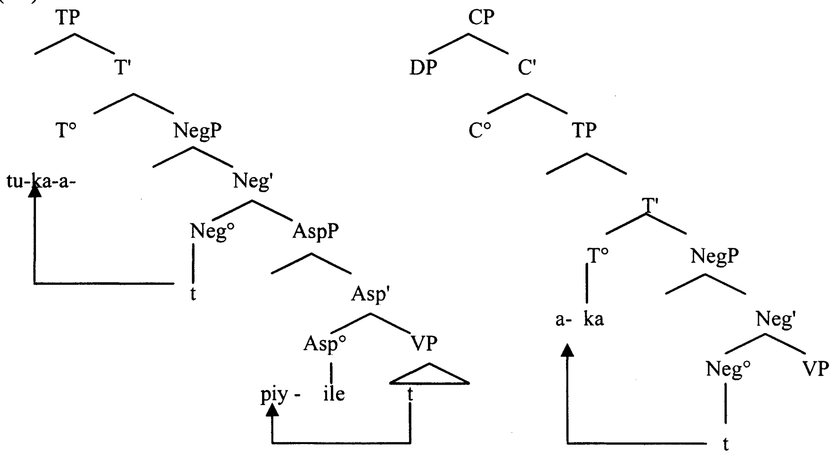
In root clauses, the negative marker appears as a clitic that is attached to the verb and is a head that selects a TP. In relative clauses, the CP selects TP, and negation is generated lower than TP. The negation marker raises and incorporates onto T° and finally raises to C° where the relative marker is marked under Spec-Head.

The languages in which the negation marker is in post-initial position in both root and relative clauses as in Nyakyusa, the TP selects NegP in both. Neg then incorporates onto T° as the following two structures from Nyakyusa show.

(24) a. tu-ka-a-piy-ile
we-NEG-PT-cook-PF
'We did not cook.'

b. un-dindwana u-yo a-ka-piy-ile mbatata
1-girl 1-REL 1SM-NEG-cook-PF 10potato
'the girl who did not cook potatoes'

(25) a. Root Clause b. Relative Clause



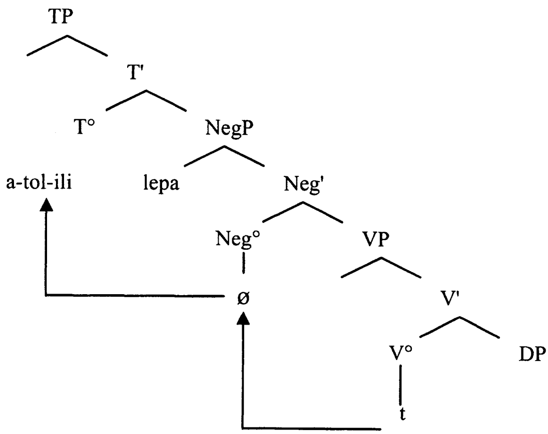
The post-verbal negative particle does not exhibit the root clause-relative clause asymmetry. Ngonyani (2001b) noted that the negative particle in Ndendeule does not prevent movement of the verb. Assuming Ngonyani's analysis in which the particle is generated higher than the VP in order to have scope over it, more specifically, in the Spec, NegP with an empty Neg°, the structure below is proposed.

(26) a. m-geni a- tol-i lepa u-gimbi
1-guest 1SM-PT-take-FV NEG 14-beer
'The guest did not take beer.'

b. m-geni mwe-a- tol-ili lepa u-gimbi
1-guest 1REL-1SM-PT-take-FV NEG 14-beer
'The guest did not take beer.'

The verb raises to the empty Neg° and subsequently to T°. As in other Bantu languages, the subject marker is an agreement marker, a feature assigned under Spec-Head configuration. The symmetric negation marking is due to TP selecting NegP in both root and relative clauses.

(27)



5. Conclusion

To conclude, in section 2, two questions were raised: (a) What is the status of Neg?, and (b) Where is Neg located in the clause structure?

For the first question, data indicate that negation projects a NegP consistent with Pollock's (1989) proposal. This is based on two considerations. One is the selectional relations between Neg and Tense which suggest in most of the languages, the negative markers are Neg. The second consideration is that the asymmetries result from avoiding the root Neg from blocking or preventing C° from governing TP. As for the second question, we find two Neg positions. One NegP selects TP, and the other TP selects NegP. This supports findings in other languages such as Romance where Zanuttini (1997) has observed that several NegPs can be found in the clause structure.

This study made use of data from several languages which are not cited here. However, the analysis presented in 4 predicts the negation marking in them. Further research is needed to determine whether the predictions are borne out in all languages.

References

Ashton, E., E. Mulira, E. Ndawula, & A. Tucker. 1954. *A Luganda Grammar*. London: Longmans, Green and Co.

Chomsky, Noam. 1995. *The Minimalist Program*, Cambridge, Mass.: MIT Press.

Demuth, Katherine & Carolyn Harford. 1999. Verb raising and subject inversion in Bantu relatives. *Journal of African Languages and Linguistics* **20**(1):41-61.

Güldemann, Tom. 1999. The genesis of verbal negation in Bantu and its dependency on functional features of clause types. In Jean-Marie Hombert & Hyman, Larry (eds.) *Bantu Historical Linguistics*. Stanford: CSLI.

Guthrie, Malcolm. 1967-71. *Comparative Bantu.* Vols. 1-4. Farnborough: Gregg Press.

Holmberg, Anders & Christer Platzach. 1995. *The Role of Inflection in Scandinavian Syntax.* New York: Oxford University Press.

Kayne, Richard. 1994: *The Antisymmetry of Syntax,* Cambridge, MA: MIT Press.

Meeussen, Achilles E. 1967. Bantu grammatical reconstructions. *Africana Linguistica* III: 79-121.

Ngonyani, Deo. 2001a. The evolution of Tanzanian Ngoni. *Sprache und Geschichte in Afrika* **16/17**: 321-353.

Ngonyani, Deo. 2001b. The relevance of tense for sentential negation in Kindendeule. Paper presented at the 32nd Annual Conference on African Linguistics, University of California, Berkeley.

Ngonyani, Deo. 2000. VP ellipsis in Kindendeule. In Kulikoyela Kahigi, Yared Kihore, & Maarten Mous (eds.) *Languages of Tanzania*, 215-232. Leiden University: Research School CNWS.

Ngonyani, Deo. 1996. The Morphosyntax of Applicatives. PhD dissertation, University of California Los Angeles.

Pollock, Jean-Yves. 1989. Verb movement, universal grammar and the structure of IP. *Linguistic Inquiry* **20**: 365-424.

Rubanza, Yunus, I. 1988. Linear Order in Haya Verbal Morphology: Theoretical Implications. PhD dissertation, Michigan State University.

Rubongoya, L. T. 1999. *A Modern Rutoro-Runyooro Grammar.* Köln: Rudiger Köppe.

Takizala, Alexis. 1973. Focus and relativization. In John Kimball (ed.) *Syntax and Semantics Vol. 2.*, 123-149. New York: Seminar Press.

Walusimbi, Livingston. 1996. *Relative Clauses in Luganda.* Köln: Rüdiger Köppe Verlag.

Yukawa, Yasutoshi. 1987. A tonological study of Nkoya verbs. In *Studies in Zambian Languages*, 129-184. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

Yukawa, Yasutoshi. 1992. A Tonological study of Luba verbs. In *Studies in Cameroonian and Zairean Languages.* Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

Zanuttini, Raffaella. 1997. Negation and verb movement. In Liliane Haegeman (ed.) *The New Comparative Syntax*, 214-245. London: Longman.

Department of Linguistics and Germanic, Slavic, Asian and African Languages
A-614 Wells Hall
Michigan State University
East Lansing, MI 48824

ngonyani@msu.edu

Time Course of the First Formant Bandwidth

HANSANG PARK
The University of Texas at Austin

0. Introduction

Phonetic characteristics of modal or non-modal voice qualities, such as open quotient, spectral tilt, the first formant bandwidth, and the degree of noise, have been measured in terms of the relative amplitude difference between the first two harmonics, the relative amplitude difference between the first harmonic and formant peaks, the first formant bandwidth, and harmonic-to-noise ratio, or subjective noise ratings. It was observed that one of the characteristics of breathy voice is an increased first formant bandwidth (Klatt & Klatt 1990) and that breathiness varies over time within a vowel (Silverman *et al.* 1995). It is predicted from these observations that the first formant bandwidth varies over time as well as across phonation types. This paper investigates the time course of the first formant bandwidth of the vowel /a/ in tokens of type /Ca/ and /VCa/, where C is 8 Korean coronal obstruents /t, t^h, t*, c, c^h, c*, s^h, s*/, to see the behavior of the first formant bandwidth in those contexts.

0.1. Theoretical background

Formant bandwidth, which is measured by taking the width of the band forming 3 dB down from the resonance peak, is a measure of resonance. Resonance, or formant, is described in terms of the peak frequency and its bandwidth. As is well known, the vocal tract transfer function can be represented as a function of formant frequency and bandwidth (Fant 1960:53).

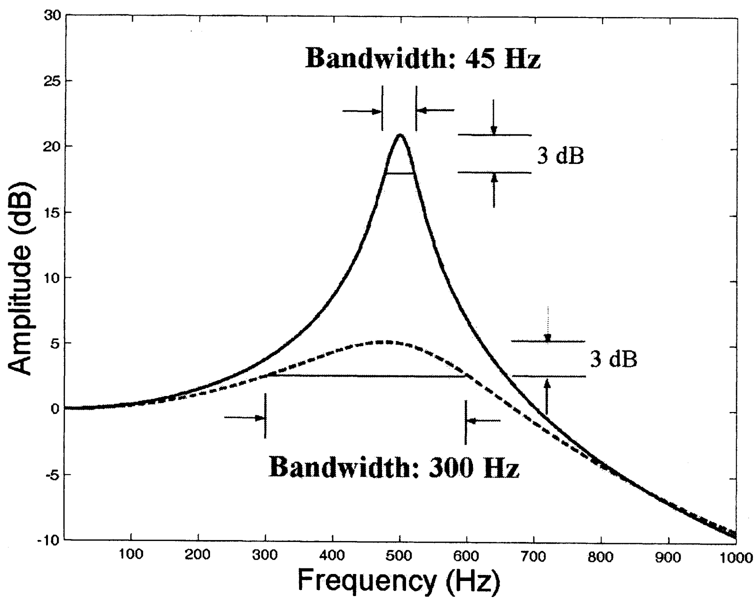
(1) Vocal tract transfer function

$$H_n(f) = \frac{F_n^2 + (B_n/2)^2}{\sqrt{(f-F_n)^2 + (B_n/2)^2}\sqrt{(f+F_n)^2 + (B_n/2)^2}},$$

where F_n is the frequency of the n^{th} formant and B_n the bandwidth of the n^{th} formant.

Formant bandwidth is associated with the amplitude of the formants such that the wider the bandwidth, the lower the peak amplitude, hence the amplitude level of the harmonics. The relationship between formant bandwidth and amplitude is illustrated in (2) below.

(2) Bandwidth and amplitude (adapted from (1))



The abscissa represents frequency and the ordinate amplitude. Two bandwidths, 45 Hz (solid line) and 300 Hz (dashed line), are illustrated with the first formant frequency set at 500 Hz. Other things being equal, the first formant peak decreases at the approximate rate of 6 dB/octave as the first formant bandwidth increases.

Formant bandwidth is determined by acoustic energy losses in the vocal tract. In general, variation in bandwidth results from the power dissipated at frequency *f*. The general scheme of formant bandwidth variation is represented in (3). (Adler 1957, Stevens 1999)

(3) The general scheme of formant bandwidth variation

$$B = \frac{power\ \ dissipated\ \ at\ \ frequency\ \ f}{2\pi \times total\ \ stored\ \ energy\ \ at\ \ frequency\ \ f}$$

There are a number of factors affecting the power dissipation, such as radiation, rigidity of the vocal tract walls, heat conduction, and viscosity, airflow at a constriction, and glottal opening. The contributions from these factors account for the numerator in the general scheme.

The contribution to the formant bandwidth due to the radiation impedance depends on the configuration of the vocal tract. If the length of a uniform tube is 17.7 cm and the cross-sectional area is 3 cm^2, then for the formants at 500, 1500, 2500, and 3500 Hz, the contribution of the radiation loss to the bandwidth of a formant is about 2, 16, 50, and 90 Hz (Stevens 1999). The contribution of the radiation resistance to the bandwidths of the front cavity resonances is particularly important. The contribution of the radiation loss to bandwidth is mainly dependent on the natural frequency of the front cavity resonance, particularly at higher frequencies.

The resistive component of the wall impedance causes energy loss in the vocal tract and an increase in the bandwidths of the formants, particularly the first formant. The bandwidth contribution of the wall resistance for the first formant of a circular uniform tube of length 17.7 cm and diameter 2.0 cm is roughly equal to 11 Hz (Stevens 1999). When the vocal tract configuration is a uniform tube, the reactive component of the wall impedance has a negligible effect on the frequencies of the formants. The contribution to the first formant bandwidth of the resistive component of the wall impedance for a Helmholtz resonator is a function of the first-formant frequency. This frequency is manipulated by changing the size of the opening in the front part of the tube. The bandwidth component decreases with increasing frequency.

The propagation of sound in a tube is accompanied by two kinds of losses at the walls of the tube, in addition to the losses due to the yielding walls (Fant 1960, Flanagan 1972). One type of loss is caused by viscous friction at the wall of the tube. Another source of loss arises because energy is drawn from the acoustic wave by heat conduction at the walls of the tube. The contributions from both viscosity and heat conduction are proportional to the square root of frequency. When the lowest natural frequency of a circular cylindrical tube of cross-sectional area 3 cm^2 is 500 Hz, the contributions from viscosity and heat conduction are 5 Hz and 3 Hz, respectively (Stevens 1999). The contribution from both viscosity and heat conduction become larger as the cross sectional area decreases, assuming the shape of the cross section of the tube does not change as the cross-sectional area changes. These components also increase as the surface area of the tube increases, for example, if the surface contains undulations or irregularities. The contributions to the formant bandwidths arising from viscosity and heat conduction are much less than those from other sources.

In summary, the contributions of viscosity, heat conduction, radiation, and vocal tract walls to the bandwidth depend on frequency and the dimensions of the resonator. Radiation contributes the most to the formant bandwidths at high frequencies, whereas the other three components are more important at low frequencies. Each of these loss mechanisms depends on frequency, with the wall losses tending to dominate at low frequencies and radiation loss dominating at high frequencies.

Airflow through a constriction can introduce acoustic loss in the form of a nonlinear resistance. This resistance can have a significant effect on the bandwidths of vocal tract resonances and hence on the overall shape of the spectrum of the sound produced by various sources and configurations. The contribution of the constriction to the lowest formant bandwidth is inversely proportional to the cross sectional area of the constriction. The resistance at a short and narrow constriction can have a significant effect on the bandwidths of resonances of cavities anterior or posterior to the constriction in addition to that of the low-frequency Helmholtz resonance. The effect of the constriction on the bandwidth is greater than that of the radiation resistance at low frequencies and less at high frequencies.

During vocal fold vibration, the acoustic impedance of the glottis varies with time. The glottal impedance has the greatest influence on the bandwidth of the first formant. The contribution to the bandwidth of a formant due to losses at glottis for a uniform tube of length 17.7 cm, a cross-sectional area of 3.0 cm^2, and a subglottal pressure of 8 cm H_2O is approximately proportional to cross sectional area of the glottis (Stevens 1999). During normal voicing a typical value for the average glottal opening during the open phase is about 0.06 cm^2, and this opening would contribute about 120 Hz to the average first formant bandwidth in this part of the glottal cycle. When the average glottal opening is greater, as it might be during an aspirated consonant, the first formant bandwidth is considerably greater. The bandwidth may be so great that the perturbation method would probably lead to a bandwidth that is comparable to the frequency of the first formant. A wide average bandwidth would also be expected during breathy voicing, when the average glottal area over the entire cycle of the glottal vibration may be 0.06 cm^2 or more.

The effect of glottal impedance on the bandwidth of a formant depends to some extent on the configuration of the vocal tract above the glottis. The effect can be particularly large when the vocal tract is constricted in the region immediately above the glottis (Fant 1979). These contributions to formant bandwidth depend on the glottal configuration used by the speaker and on the subglottal pressure, and may vary somewhat from speaker to speaker. The reduction in cross-sectional area in the pharyngeal region can influence the bandwidth or the degree of spectral prominence of particular formants. For a low vowel with a backed tongue body, the first formant is often associated primarily with the back cavity, particularly for adult male speakers, since most of the acoustic energy at the frequency of the first formant is stored in this pharyngeal region. The contribution of these losses due to pharyngeal constriction to the formant bandwidths is roughly the same for different speakers, except for differences in scale. (Stevens 1999)

0.2. Goals of the present study

Klatt and Klatt (1990: 825) stated, "Additional characteristics of a breathy vowel include the possibility of increased first-formant bandwidth and/or the appearance of tracheal poles and zeros in the vocal tract transfer function due to the greater glottal opening." In other words, a breathy voice is accompanied by a wider glottal opening that causes an increased first formant bandwidth as well as a coupling of the supraglottal and subglottal cavities. This statement suggests that the breathier the vowel, the wider the first formant bandwidth.

Silverman *et al.* (1995) showed that in Jalapa Mazatec, which has modal, breathy, and laryngeal voices, contrastive breathiness lasts for only 43 % of the vowel duration, giving way to modal vibration for the remainder. It is predicted from this observation that the first formant bandwidth varies over time within the breathy vowel itself. It is evident that phonetic characteristics of different phonation types vary over time. Blankenship (1997) showed that phonation type varies over time in terms of H1-H2, H1-F2, and cepstral peak prominence in Tagalog, Mazatec, Chong, Mpi, and Navajo. Ahn (1999) also indicated that voice qualities are different at different time points within a vowel as well as across phonation types in Korean.

Stevens and Hanson (1995) and Hanson (1997) employed first formant bandwidth as well as H1*-H2*, H1*-A1, H1*-A3* in their studies of voice qualities of female speakers. They estimated the first formant bandwidth by measuring the decay rate of the first formant waveform during the early part of the glottal period, where the glottal area is expected to be smallest, under the assumption that F1 oscillation is a damped sinusoid in the form $e^{-\alpha t}\cos 2\pi f$ where f is the first formant frequency and α is related to the first formant bandwidth by the equation $B1 = \alpha/\pi$.

First of all, they selected 8 consecutive pitch periods from a relatively stable portion of the vowel, generally at the middle. Then, they isolated F1 oscillation by applying a bandpass filter of 600 Hz centered at the first formant frequency. They determined the rate of decay from the change in the peak-to-peak amplitude in the first two cycles of the F1 oscillation. Finally, they estimated the decay rate for 8 consecutive pitch periods and then average 40 estimates to obtain a mean value for each speaker.

They restricted estimation of the first formant bandwidth to a relatively stable portion of a vowel /æ/, providing a caveat that there must be, in order to obtain accurate estimates, a high enough first formant frequency and a long enough pitch period to get at least two oscillations during the closed part of the cycle. This caveat suggests that the estimation method is not adequate for speech signals with a lower F1 or a short pitch period where two oscillations are not available during the closed part of the cycle. This means that the closed part of the glottal cycle, $(1-OQ)T_0$, where OQ is open quotient and T_0 is a fundamental period, must be greater than double the period of F1 oscillation, T_{F1}. If T_{F1} is considered to be the period of the harmonic closest to the F1, and therefore, the frequency of the F1 oscillation $1/T_{F1}$ can be represented as a multiple of the fundamental frequency, this relation can be roughly expressed as in (4).

(4) Condition for B1 estimation

$$2T_{F1} < T_0(1-OQ), \quad where \quad nT_{F1} = T_0$$

$$0 < (1-OQ) - 2/n$$

where, n is the index of the harmonic closest to the first formant frequency. According to this condition, their method is applicable to rather a restricted range of speech samples. It is hard to apply this method to speech samples with a low harmonic index due to a high fundamental frequency, with a low harmonic index due to a high vowel with a low first formant frequency, or to breathy voices with a high OQ. It is, in reality, not easy to obtain accurate estimates from female voices which usually have a higher fundamental frequency, hence a low harmonic index, or from breathy voices which have a higher open quotient, since it is difficult to obtain at least two oscillations during the closed part of the fundamental period. As a corollary, they did not estimate B1 for the marginal part of the vowel where quite different voice qualities are likely to be observed.

The present study does not apply the method provided in Stevens and Hanson (1995) and Hanson (1997), since the present study is interested in quite different voices qualities, particularly breathy voice that is highly likely to be observed at the onset of the vowel after an aspirated consonant. This study, instead, simply measures first formant bandwidths from LPC spectra. Even though the LPC algorithm may have its inherent problems, it could be a reasonable way to investigate variation in the first formant bandwidth if the first formant bandwidth is consistently measured for all tokens within a given setting. The present study investigates time course of the first formant bandwidth of a Korean vowel /a/ from the onset of the vowel to the offset, based on the observation that the earlier part of the vowels after different consonant types in Korean have different voice qualities (Ahn 1999). This study may provide yet another profile of phonation type variation.

1. Method

1.1. Subjects

Three male native speakers of Standard Korean in their mid-twenties participated in the experiment. All subjects were students at the University of Texas at Austin who had been in the United States for less than 6 months. Two of them were native Seoulites born to parents who were also native Seoulites while one was born and raised in Kyunggi province, an area surrounding Seoul, to parents who had also been born and raised in the same province. None of the subjects had any history of speech disorders.

1.2. Materials

The material used in the present study is represented in (5).

(5) The material used in the present study

a. Tokens

[t^hada]	'to ride'	[tada]	'to be *ta*'
[t^*ada]	'to pick'	[$tʃ^h$ada]	'to kick'
[tʃada]	'to sleep'	[$tʃ^*$ada]	'to squeeze'
[s^hada]	'to buy'	[s^*ada]	'to wrap'

b. Carrier Sentence

/$\underline{\sigma_1\sigma_2}$ ka anila $\underline{\sigma_2\sigma_1}$ta/ 'It is not $\sigma_1\sigma_2$ but $\sigma_2\sigma_1$'

c. Sample Sentence

[**<u>t^hada</u>**gaaniɾa **<u>tat^{h}a</u>**da]

All tokens provided in (5a) are infinitival forms of real Korean verbs except *tata*. However, *tata* is also often spoken in real speech, meaning '(The answer) is *ta*'. The vowels in all tokens are phonologically short. Tokens are different from each other only in their initial consonants. Two-syllable words were embedded in a carrier sentence, as is shown in (5b). Utterance-initial CV sequences were available in the first slot of the carrier sentence. The two syllables in the second slot were derived by a metathesis of the token in the first slot of the sentence to obtain VCV sequences. The reason why the present study investigates both utterance-initial CV sequences and non-utterance-initial VCV sequences, even though the latter properly includes the former, is that the present study was interested in positional variation of the first formant bandwidth. An example of the tokens embedded in a carrier sentence is presented in (5c).

1.3. Recording and Digitization

Subjects were asked to read a list of randomly ordered sentences to the effect that each sentence was repeated 10 times. Subjects were given the instruction that each sentence should be separated by a long pause. Recording was conducted in the Phonetics Lab in the Department of Linguistics at the University of Texas at Austin. Speech data were recorded onto analogue tapes via a Marantz SuperScope. The analogue signals were digitized on SoundScope at a sampling rate of 22,050 Hz.

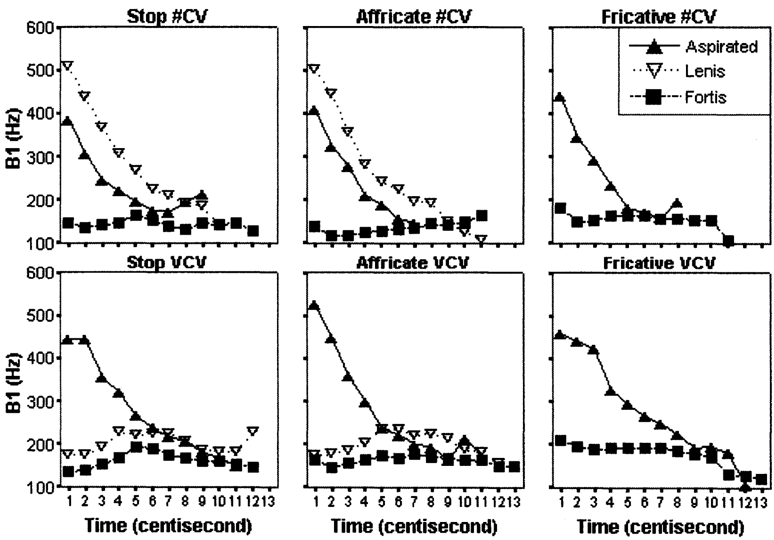
1.4. Measurements

First formant bandwidth was measured from pre-emphasized LPC spectra of 20 ms Hamming windows. The first window was placed so that the center of the window was set at 10 ms from the onset. Measurements were taken from the windows at the interval of 10 ms, overlapping 10 ms with the preceding window. Measurement was not performed for any window whose size was less than 20 ms around the offset of the vowel.

2. Results

Time course of the first formant bandwidth by manner of articulation and position is provided in (6) below.

(6) Time course of the first formant bandwidth by manner and position



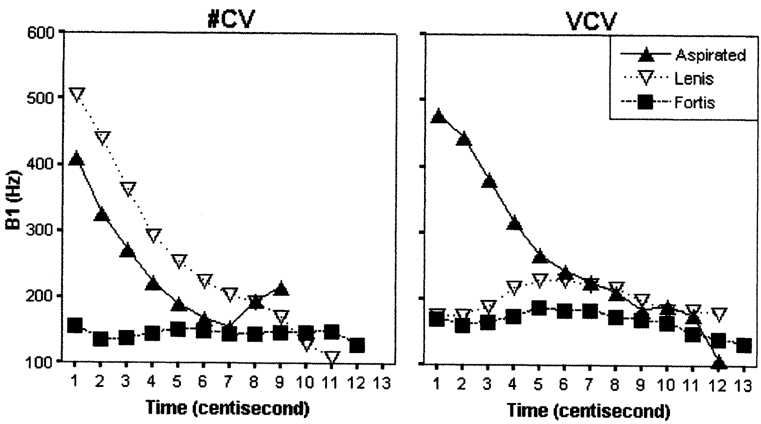
In (6), the upper panels indicate the utterance-initial CV sequence (#CV) while the lower panels the intervocalic VCV sequence (VCV). In each position, panels indicate manners of articulation from the left: Stop, Affricate, and Fricative. The abscissa represents time in centisecond from the onset of the vowel, while the ordinate the first formant bandwidth (B1). In each panel, filled triangles stand for the vowel after aspirated obstruents (Aspirated), empty upside-down triangles for the vowel after lenis obstruents (Lenis), and rectangles for the vowel after fortis obstruents (Fortis). It should be noted that Fricative has only two series in Korean.

At word onset, B1 was greatest in Lenis tokens (around 500 Hz) in Stops and Affricates, intermediate in Aspirated tokens (around 400 Hz), and least in Fortis tokens (around 150 Hz). Intervocalically, B1 was greatest in Aspirated tokens (around 500 Hz), intermediate in Lenis tokens (around 200 Hz), and least in Fortis tokens (around 150 Hz).

A rapid decrease was observed over time for both Aspirated and Lenis tokens in #CV, while only for Aspirated ones in VCV. There was no substantial change in B1 for Fortis tokens in either position, nor for Lenis tokens in VCV. With a greatest difference in B1 at the onset, the lines representing different phonation types merged around 50% of the vowel duration (around 5th or 6th window). Lenis patterned with Aspirated in #CV, while with Fortis in VCV.

We can summarize the results by presenting time course of B1 by position alone, which is illustrated in (7) below.

(7) Summary of the time course of the first formant bandwidth



A striking difference was observed at the onset of the vowel. The difference was reduced drastically, so that there was no discernible distinction after 50 ms from the onset. Lenis patterned with Aspirated in #CV, while with Fortis in VCV.

3. Discussion and Conclusion

The striking difference in the first formant bandwidth across phonation types at the onset of the vowel suggests that this difference may be a function of the glottal state in the prevocalic phase. In the utterance-initial CV sequence, a wide bandwidth in Aspirated is associated with an open glottis in the prevocalic phase. In other words, vocal fold vibration is initiated with a considerable glottal opening. On the other hand, a narrow bandwidth in Fortis is related to a closed glottis in the prevocalic phase. Vocal fold vibration begins with a completely closed glottis. This closed glottis can be achieved by a voluntary adduction of vocal folds.

In the intervocalic VCV sequence, a wide bandwidth in Aspirated and a narrow bandwidth in Fortis are each associated with an open glottis and a closed glottis in the prevocalic phase, as in the utterance-initial CV sequence. However, a narrow bandwidth in Lenis in the intervocalic VCV sequence is also related to a closed glottis in the prevocalic phase. In Lenis, vocal fold vibration begins with a completely closed glottis. This closed glottis seems to be achieved by a natural process of intervocalic voicing. In the particular case of Korean, intervocalic voicing of lenis obstruents seems to be a function of closure duration. The intervening interval between two vowels is so short that the voicing of the previous vowel is extended to the following vowel, maintaining vocal fold vibration throughout the interval. In this case, a closed glottis is ready even before the following vowel begins. This glottal state, in turn, leads to a low bandwidth at the onset of the following vowel.

This view is supported by Kagaya (1974). In his fiberscopic study of glottal width of Korean obstruents, Kagaya (1974) showed that glottal width is greatest in Aspirated, intermediate in Lenis, and least in Fortis in the prevocalic phase. This observation is comparable to Kim (1970) for stops. In addition, Kagaya (1974) provided the time point of a complete contact of the vocal folds. For the isolated /CV/ environments, a complete contact of the glottis is observed around voice onset in Aspirated and Lenis while a complete contact of the vocal processes is achieved early before the voice onset in Fortis. For the intervocalic /VCV/ environments, the same pattern occurred in both Aspirated and Fortis as in the /CV/ environments. However, a complete contact of the vocal processes is observed in Lenis all through the utterance. This observation is consistent with the generally accepted allophonic variation of this type of Korean consonants, that is, intervocalic voicing. The vocal processes contact each other again at a point from 40-100 ms before the voice onset as in the case of /CV/ environments.

A wide first formant bandwidth does not abruptly give way to modal vibration but decreases gradually over time in Aspirated in both positions and in Lenis in the utterance-initial CV sequence. The monotonic decrease in the first formant bandwidth during the earlier part of the vowel indicates a gradual yielding to modal vibration. This pattern suggests an interpolation between two target glottal states. This observation is agreement with Silverman *et al.* (1995) who, as noted earlier, observed that in Jalapa Mazatec contrastive phonation types last only for 43% of the vowel duration, giving way to modal phonation for the rest of the vowel duration.

This inference seems to explain why the pattern of variation in the first formant bandwidth is consistent for Aspirated and Fortis across positions, while Lenis patterns with Aspirated in the utterance-initial CV sequence while with Fortis in the intervocalic VCV sequence. However, we did not provide an answer to why the first formant bandwidth is greater in Lenis than in Aspirated in the utterance-initial CV sequences and than in Fortis in the intervocalic VCV sequences. One possible answer may be that Lenis has a wider glottal opening than Aspirated in the utterance-initial CV sequences and than in Fortis in the intervocalic VCV sequences. This inference requires a presupposition that the first formant bandwidth is the only indicator of glottal opening and glottal opening is the only source of an increased first formant bandwidth. No evidence, however, is provided supporting this idea. In the previous section, It was noted that bandwidth is affected by many factors including radiation, vocal tract wall, heat conduction and viscosity, airflow at a constriction, glottal opening. Among these, radiation, vocal tract wall, and heat conduction and viscosity have a minor effect on the first formant bandwidth. It seems to be clear that glottal opening is major source of an increased first formant bandwidth in the experimental setting of the present study. An alternative is that there is some other source than glottal opening that increases bandwidth. This remains to be explored.

In summary, a spontaneous transition of glottal opening from the prevocalic phase, a voluntary glottal closing, and extension of voicing through the short interval of the intervening consonant creates two possible glottal states at the onset of the vowel: a closed glottis and an open glottis. The default glottal state in the utterance initial position is open while that in the intervocalic position is either closed or open. Vocal fold vibration begins with either of the default states consistently in Aspirated and Fortis regardless of position, whereas the default state varies with position in Lenis. The interaction of these two default glottal states with the positional context results in a distinct behavior of the first formant bandwidth in Lenis.

References

Adler, R. B., L. J. Chu, & R. M. Fano. 1957. *Electromagnetic energy transmission and radiation*. New York: Wiley.

Ahn, Hyunkee. 1999. Post-release phonatory processes in English and Korean: Acoustic correlates and implications for Korean phonology. PhD dissertation, The University of Texas at Austin.

Blankenship, Barbara. 1997. The time course of breathiness and laryngealization in vowels. PhD dissertation, UCLA.

Fant, Gunnar. 1960. *Acoustic Theory of Speech Production*. The Hague: Mouton.

Fant, Gunnar. 1979. Glottal source and excitation analysis. Speech Transmission Laboratory Quarterly Progress and Status Report **1**, Royal Institute of Technology, Stockholm, 85-107.

Flanagan, J. L. 1972. *Speech analysis, synthesis, and perception*. Berlin: Springer-Verlag.

Hanson, Helen M. 1997. Glottal characteristics of female speakers: Acoustic correlates. *Journal of the Acoustical Society of America* **101**: 466-481.

Kagaya, Ryohei. 1974. A fiberscopic and acoustic study of the Korean stops, affricates and fricatives. *Journal of Phonetics* **2**: 161-180.

Kim, Chin-Wu. 1970. A theory of aspiration. *Phonetica* **21**: 107-116.

Klatt, Dennis H. & Laura. C. Klatt. 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America* **87**(2): 820-857.

Silverman, D., B. Blankenship, P. Kirk, & P. Ladefoged. 1995. Phonetic structures in Jalapa Mazatec. *Anthropological Linguistics* **37**(1): 70-88.

Stevens, Kenneth N. 1999. *Acoustic Phonetics*. Cambridge: MIT Press.

Stevens, Kenneth N. & Helen M. Hanson. 1995. Classification of glottal vibration from acoustic measurements. In O. Fujimura & M. Hirano (eds.) *Vocal fold physiology: Voice quality control*, San Diego: Singular, 145-170.

Hansang Park
Department of Linguistics
The University of Texas at Austin
Austin, TX 78712

phans@mail.utexas.edu

Debunking 'Discourse Topic': Ideational Momentum and the Unpredictable Trajectory of Conversation[1]

MISCHA PARK-DOOB
University of Chicago

0. Introduction

There is a long-standing notion that in most instances of language use, the participants (whether speaking or writing) organize their production into coherent stretches identifiable as 'topics'. There is the further notion that these topics have easily definable beginnings and ends, allowing them to be labeled and strung together into a list or hierarchy which forms a useful schematic of what occurred during a given stretch of discourse. The preliminary rationale goes as follows:

> The conversationalists stop talking about 'money' and move on to 'sex'. A chunk of conversational discourse, then, can be treated as a unit of some kind because it is on a particular 'topic'. The notion of 'topic' is clearly an intuitively satisfactory way of describing the unifying principle which makes one stretch of discourse 'about' something and the next stretch 'about' something else, for it is appealed to very frequently in the discourse analysis literature. (Brown & Yule 1983:70)

It was along these lines that Keenan & Schieffelin (1976) introduced the 'discourse topic' as the "question of immediate concern", "a proposition (about which some claim is made or elicited)". This intuition seems to argue for parsing discourse in terms of separate topics, each identifiable by a single phrase or sentence.

As Brown & Yule and many later analysts have recognized, such a view is far too simplistic. Their own concept of a "topic framework" is defined as the contextual framework "within which the topic is constituted", consisting of "*activated features of context* . . . those aspects of the context which are directly reflected in the text, and which need to be called upon to interpret the text"

[1] This paper is basically a summary of my (unpublished) Master's thesis, *Deconstructing 'topic': Relevance, consciousness, and the momentum of ideas* (Park-Doob 2001). I'd like to thank my advisor David McNeill, and my colleagues in the McNeill Lab, including Karl-Erik McCullough, Chris Corcoran, Sue Duncan, KaLynne Harris, Nobuhiro Furuyama, Irene Kimbara, Fey Parrill, Mika Ishino, Amy Franklin, Arika Okrent, Nicla Rossini, and Gale Stam. Recent work has been supported by NSF KDI grant No. BCS-9980054, "Cross-Modal Analysis of Signal and Sense: Multimedia Corpora and Tools for Gesture, Speech, and Gaze Research", Francis Quek, P.I.

(Brown & Yule 1983:75). This formulation presents a few problems. In particular, the features of context which are salient for participants, at the time of interaction, may not all be later identifiable by the analyst, nor will they necessarily all be directly reflected in the "text". In an effort to go beyond both the text-based tradition and the speech-focused tradition, a growing body of work studying face-to-face narrative and conversation has been exploring the complex ways in which people communicate using the visuo-spatial and verbal modalities in synchrony (cf. McNeill 1992, papers in McNeill 2000). With a high-resolution audio/video recording of an interaction, the analyst has a view to the spontaneous hand gestures, facial expressions, body postures, and eye-gaze used in synchrony with speech by the participants, in addition to information about their physical environment. Given that even the best possible audio/video recording still fails to provide the analyst with a great deal of the information available at the time to the participants themselves, it seems clear that using just audio recordings and, worst of all, just written transcriptions to study conversation will provide limited insights (or even worse, result in false conclusions). Brown & Yule's "topic framework", though fairly bogged down in a tradition of research using written transcripts, does at least highlight the critical necessity of examining both the fluid shared context in which any conversation is situated, as well as all available data from the conversation thus far, and points to the necessity for conversational participants to make their contributions *relevant* to the currently salient elements of the context and to the currently "activated" elements from the conversation thus far. Relatively few scholars have emphasized the importance of the interactively constructed, dialogic nature of language phenomena (Clark (1996) and Linell (1998) stand out favorably). Especially important is Rommetveit's (1974) concept of *intersubjectivity* in interaction, "in which there exists a partial complementarity, temporary reciprocity, and above all a shared consciousness among the interlocutors" (O'Connell & Kowal, in press).

In this paper I will assume that participants in a conversation express *ideas* of some sort, and that their contributions are *relevant*, to a greater or lesser degree, to what has recently been, or is currently being, expressed or talked about. I use the word *idea* here in a purposefully vague sense, to mean any expressible notion, able to be communicated in any way. Jakobson (1960) provides a classic introduction to the multiple functions language is capable of handling simultaneously.

1. Two modern attempts to improve on the 'discourse topic'

Rather than parse discourse by labeling chunks with individual topics, a system has been devised by Barbara Grosz and colleagues (cf. Nakatani et al. 1995) that assumes segments of spoken discourse have a specific *purpose* underlying them, agreeing with Brown & Yule's statement that "any consideration of topic involves asking why the speaker said what he said in a particular discourse situation" (1983:77). The system described in Nakatani et al. (1995) involves parsing a monologue into a hierarchical structure of large segments with broad purpose, each embedded with smaller nested subsegments of increasingly specific purpose.

The success of such a purpose hierarchy analysis requires that the segments and subsegments have clearly identifiable beginnings and ends, and that they each have clearly identifiable purposes and nested subpurposes. Each subpurpose must succeed in supporting the broader purpose(s) above it in the hierarchy.

Herb Clark differs from many analysts by making explicit the claim that all language is a form of *joint action* (Clark 1996). He acknowledges the cooperative, emergent, and opportunistically created nature of conversation, and wisely remarks that transcripts are only "footprints in the sand" (337), imperfect records of the dynamic processes of language in use. Naturally, he finds that "discourse topic is a static notion that doesn't do justice to the dynamic course of joint actions and what develops out of them" (342). To get around the question of 'topic' in interaction, he suggests joint actions are generally part of *joint projects*, which are structured hierarchically by smaller subprojects. What we get is a project hierarchy analysis which rather resembles the purpose hierarchies of Grosz and colleagues, but differs crucially in that the projects and subprojects are viewed as dynamic *processes*, not static objects with set purpose. In the coding scheme illustrated by Clark (343), each project and subproject requires a clear beginning and end, consciously agreed upon by the participants, and subprojects are placed at specific levels of embeddedness. Thus the two coding systems are similar in that each segment requires clear boundaries, and must have clear placement in a hierarchy. In the following sections we will discuss data that cooperate with these requirements, as well as data that illustrate their pitfalls.

2. The experiment

Before presenting data from recorded conversations, a few introductory remarks are in order. I originally recorded all the examples discussed below as part of a McNeill Lab study meant to elicit spontaneous gestures in two-person conversations (approximately 65 pairs of participants were recorded in 1999 and 2000; the participants cited here were all undergraduates at the University of Chicago). The present work grew out of my attempts to parse the results into purpose hierarchies inspired by Nakatani et al. (1995). Those attempts were done as a first pass, without special emphasis on examining gestures, and likewise I will not raise specific points about gestures in this paper. It should be noted, however, that all my judgments were made based on repeated observation of the video taped data, which allowed me to make qualitative observations about the participants' physical interaction with their environment and each other (including their gestures) and undoubtedly increased the accuracy of my judgments.

The experiment proceeded as follows: participant B remained outside the experiment room while participant A examined a sheet of instructions pertaining to a plastic model city placed in front of his or her chair. The model was about three feet square and had five buildings: a movie theater, train station, church, and two houses (called #33 and #35). The instructions, which were the same for each pair of participants, outlined the following rather comical scenario: a family of intelligent wombats had taken up residence in the abandoned movie theater of the tiny town of Arlee, and it was the participants' job to recruit helpers from the town, flush out the wombats, capture them, and send them back to Australia. Participant A was instructed to describe the scenario to participant B, and together they were to design a creative plan of action (using the model city as a guide), and decide on what materials would be needed and how they would be used.

3. When the methodology works, and when it fails . . .

Example (1) below is a small sample of a purpose hierarchy, the first draft of which was created by KaLynne Harris and me during the summer of 2000. Each segment is headlined by a verb phrase describing its purpose, and each line represents one *intonation unit*, a single communicative pulse which is also definable in prosodic terms (see Chafe 1994:57-70). Please refer to the transcription conventions listed at the end of the paper (paraphrased from Chafe 1994 and 1997, with a few additions based on McNeill Lab conventions). Given that the transcription is a mere shadow of the real interaction, it is mainly designed for maximum readability. I have attempted to keep the data relatively free of clutter, while still trying to provide hints to what the speech actually sounded like.

(1) **6.5 = discuss roles of A and B after wombats are rounded up in pen**
A # an' só==,
A # whíle,
  6.5.1 = go into details about B's role
  A ... since yóu're out front doing nothing but bànging your pót .. {^{1}[nn]^{1} yer sh}—
  B .. ^{1}[yeah.]^{1}
  A .. with your stíck,
    6.5.1.1 = refute A, clarify role
    B well Í'm ^{2}[gonna]^{2} be rùnning ^{3}[thróugh.]^{3}
  A ^{2}[you'll—]^{2} ... ^{3}[you'll—]^{3} you'll— you'll come $^{@}$back óut$^{@}$ ^{4}[when—]^{4}
  B ^{4}[yeah.]^{4}
  A .. when thèy're thére.
A # and,
A ... yóu kèep an èye on thèm an' Í'll go wìth= the péople in hòuse—
A .. from e—
A w— thirty-thrèe or thirty-five,
  6.5.2 = discuss calling authorities
  A # an' we'll cáll up lìke,
  A ...(1.2) ánimal contròl 'r whatéver.
6.6 = (unclear purpose)
A ... an' then
A ... we'll==,
  6.5.2 (return) = discuss calling authorities
  A ... or=,
  A ... y'know.
  A ... the authórites in Austrália n' we'll make a phóne càll an' say we have these wómbàts that need to go back to Austrália or—
  A # (*quite loud:*) or we'll cáll a local zóo.
  A ...(1.1) an' jùst be like we've càught these wómbàts and we néed you to còme táke thèm==
  A like bàck to Austrália or whatèver.

As can be seen, the above example fits fairly well into a purpose hierarchy analysis, for it consists mainly of a cinematic or play-by-play description, thoroughly dominated by participant A, of the plan for rounding up the wombats (see Park-Doob 2001 for a lengthier excerpt). Her description outlines a series of subtasks within subtasks, all geared toward the larger task of capturing the wombats, and all described in chronological order. Each subtask has an easily identifiable purpose functioning as a clear subpart of the larger mission goal, thus creating a *hierarchy* of tasks and nested subtasks, and so it is a straightforward matter to organize her speech into an identical hierarchy of purposes, each labeled according to what she is trying to describe. Participant B interferes very little, so this part of their dialogue behaves much like a *monologue*. Given that the discussion is also highly goal oriented, it's not surprising that it works well in a Grosz-style system, which in the case of Nakatani et al. (1995) was designed specifically to deal with goal oriented, task-driven monologue.

However, if intended for use with dialogues in general (and even many monologues), the treatment above poses some severe problems. First of all, the participants may *begin* a subsegment with some clear subpurpose to a broader purpose, but then lose track of this broader purpose as the subsegment progresses. This is due to our limited capacity to keep track of information and, in dialogic interaction, to our tendency towards unpredictable digressions and divergence (these issues will be discussed in detail below). The system of indentation is therefore potentially misleading: it places talk at absolute levels of nestedness, disregarding the fact that participants may lose track of the broader purpose, or digress repeatedly without ever returning to the previous discussion. In some conversations, we would be faced with a coding nightmare and find ourselves indenting beyond the right margin of the page! Fundamentally, speakers often shirk the broader purpose that led to their current discussion, even though that broader purpose may have been critical a few seconds before, and there may not have been a specific point of 'purpose shift' during the intervening talk. Since there may not be any *absolute* level of hierarchical nestedness in the minds of the participants, we have no justification in coding a transcript this way. The best we can do is note the *relative* changes in nestedness at their moment of occurrence, and allow that the resulting hierarchical relationship may lose credence as talk progresses.

The other critical problem with this analysis lies in its explicit requirement that talk be segmented into discrete, clearly delineated sections, each labeled with a specific 'purpose'. As we shall see, it is sometimes impossible to pinpoint where such segments begin or end. Clark's (1996) system of joint project hierarchies affords greater flexibility, but basically suffers from the same weaknesses: while participants may find themselves interacting jointly as if engaged in a particular project, there may be no clear spot where that particular project began if it has diverged smoothly away from the project of a few seconds before. Clark's hierarchical structure is also hampered by the same dependence on absolute levels of embeddedness. Consider the following extended excerpt, from a second pair of participants:

(2) B ... # an' máybe y'know whìle we're thère we còuld .. y'know,
stày with thèm an',
... pàrty with thèm an',
... èat 1[their fóo=d an',]1
A 1[brìng a few—]1 ... bèers,
B ... fr— yèah bring a few b—
.. ingrátiàte oursèlves with thèe uh,
with the tòwn= ... prèacher,
A ... 1[m-hm.]1
B ... 1[#]1 with a few bèers,
... @2[@@]2 @ ... #
A 2[@@]2 .... #
B mày—
.. Ì don' know .. maybe we could gèt some diví=ne help,
.. to gèt rìd of the wó=mbàts=.
1[m—]1
A ... oh 1[y]1éah.
yòu're Cátholic áren't you.
B # ... yéah.
we could hà—
... like wè could .. s=èe if they had any hóly water on hànd an',
A .. @
B ... {w}ait .. màybe they're demónic @.. wòm=bàts@ 1[@@@@]1
A 1[@@@@]1
... @and crú@cifixes.
B .. @
A .. #—
B .. # and crúcifixes.
.. and gár1[lic.]1
A 1[thà]1t would be rèally gòod.
B # gárlic we could ùse as bàit,
mày2[be wómbàts like gárlic.]2
A 2[yéah I b— .. I bèt th]2ere're vámpìres lìving in the—
... abàndoned movie theater.
B vå3[mpire wòmbats.]3
A # 3['cause you knów where— #]3 where there're báts= there're vámpìres tóo.
... an' ònce we find their còffins we can búrn'em.
B ... yés==.
A or expóse'em to the lìght of dày.
#
B ... yés==.
... 1[àbso]1lùtely.
A ... 1[yeah.]1

B # becàuse y'knòw,
  # vàmpires and wòmbats are wònt to lìve in .. old abàndoned móvie theaters.
A ^{2}[.. m-hm.]^2
B ^{2}[#]^2 in smàll tòwn= USÀ.
A .. @
B ... @
  # um=,
A ... or—
  ... or hélicòpters.
  # maybe we could just còver the entíre—
  # like using hélicopters we could còver the entíre—
  # tówn with a hùge nét.
    (continues. . .)

This excerpt poses a serious problem for an analyst wishing to parse it into a hierarchy of purposes, projects, or topics. While the participants actually do keep their discussion within the bounds of the task at hand, we face trouble once we try to parse it beyond this single over-arching project. Their talk begins with a discussion on *partying with the townsfolk*, and then follows a smoothly curving trajectory all the way to *killing vampires*. There are no 'kinks' in this trajectory. In other words, if we looked at any single moment during this stretch, the contributions before and after would all seem part of the same 'topic'. Then at the moment when participant A begins talking about helicopters, we have a sharp change easily identifiable by any traditional analysis as a point of 'topic shift'. The fundamental point is that the difference between *partying with the townsfolk* and *killing vampires* is certainly just as great as the difference between *killing vampires* and *using helicopters*. Clearly, the participants are able to drastically alter the course of their discussion without relying on specific points of major transition. A traditional analysis which cuts discourse into discrete, bounded segments will by definition focus on the *boundaries*, ignoring the fact that one of these same segments may experience subtle yet continuous internal changes which result in a net difference just as great as that found at the boundaries.

The constantly changing conversational trajectory illustrated in (2) is in no way abnormal. On the contrary, such a phenomenon is predictable given certain key aspects of human consciousness, which we examine next.

4. The nature of consciousness

For this discussion, we will follow Wally Chafe's model of consciousness:

> At any given moment the mind can focus on no more than a small segment of everything it "knows." I will be using the word *consciousness* here to refer to this limited activation process. Consciousness is an active focusing on a small part of the conscious being's self-centered model of the surrounding world. (Chafe 1994:28)

Chafe goes on to describe certain key aspects of consciousness. First of all, consciousness has a *focus* which is "restless, moving constantly from one item of information to the next" (29). This movement is reflected linguistically in sequences of intonation units, as each unit "verbalizes the speaker's focus of consciousness at that moment" (63). Most of what is available to be focused upon will not be in a person's active consciousness at a given moment:

> This limited activation allows a person to interact with the surrounding world in a maximally productive way, for it would hardly be useful to activate everything a person knew at once. Aside from the burden such a process might place on neural resources, most of that vast store of information would be irrelevant to one's interests at any particular time. (29)

Consciousness also maintains, as a context for its current focus, a "periphery of *semiactive* information" (29). Once it is no longer in active focus, a given idea does not usually drop out of one's head immediately, but rather remains for a while in the semiactive realm. Each new focus then "find[s] coherence in the contexts provided by the surrounding semiactive information" (30).

Furthermore, consciousness requires an *orientation*, which gives individuals information about themselves with respect to the crucial domains of "space, time, society, and ongoing activity" (30). How a person's consciousness is oriented can be said to be a part of the peripheral context which aids the understanding of each new focus. Brown & Yule (1983:60) suggest that interlocutors try to understand each other's utterances through as little processing as possible, and that they do this by way of an assumption of *local interpretation* (an interpretation based on the most recently put forth utterances and ideas). We can infer that a contribution intended to have a non-local interpretation will require more processing time and effort to produce and interpret. Chafe (1994:138) remarks that boundaries between "episodes" of talk are often observable as points of sudden *reorientation*, with the strength of the boundary depending on how much reorientation is necessary. It seems clear then that reorientation requires *work*. In order for a reorienting contribution to be coherent, the speaker must provide extra signals to aid the addressees in reorienting toward the new ideas, and the addressees must pick up on these signals and do the work of matching the speaker's new orientation.

5. Chafe's notion of 'discourse topic'

Equipped with an understanding of Chafe's notion of consciousness, we can move on to discuss his definition of 'discourse topic':

> We can think of each . . . topic as an aggregate of coherently related events, states, and referents that are held together in some form in the speaker's semiactive consciousness. A topic is available for scanning by the focus of consciousness, which can play across the semiactive material, activating first one part and then another until the speaker decides that the topic has been adequately covered for whatever purpose the speaker may have in mind. (Chafe 1994:121)

While this notion is appealing, and a great deal of real conversation may occur this way, Chafe seems to assume that conversations *always* have such a structure: "One of the things that seems intuitively true of conversations is that they focus on different topics . . . at different times, moving from one topic to another" (Chafe 1994:121). The problem, similar to what we faced with the frameworks of Herb Clark (1996) and Barbara Grosz and her colleagues (Nakatani et al. 1995), is that we will fail to identify clear 'topics' in actual stretches of discourse if the talk proceeds haphazardly and without clear points where any speaker deems the talk has "adequately covered" the ideas whose full discussion required organization within a 'topic'. After a speaker initiates discussion that he or she hopes will remain within a particular bounded realm of semiactive consciousness, the participants have limited resources with which to ensure the mutual continued awareness of those boundaries, and they may not have even agreed on the boundaries in the first place. During the course of "scanning" across the semiactive material with their ever-restless foci of consciousness, one or more participants may begin to lose track of the originally intended boundaries, allowing their understanding of the "semiactive realm" of *current* discussion to gradually move farther and farther from the original. Since consciousness has only a single, constantly restless focus and a limited ability to maintain peripheral attention on salient pieces of semiactive information, staying 'on topic' requires the tough work of continually re-foregrounding ideas as they pass out of active consciousness, become faded in semiactive consciousness, and begin to slip out of a speaker's control.

In the realm of semiactive consciousness, certain elements may remain highly salient while others fade away. Physical context is particularly important in this respect, as it plays the main role in maintaining orientation in space and time. A constant physical orientation provides the necessary stimuli by which many other elements of an interaction remain in the foreground of attention. For example, in (2) above, the participants allow their talk to meander without any clear 'topic' structure, yet all of their discussion remains confined to the realm of the model city and the task at hand. Their very physical presence in the lab, among its lights and cameras, and the model city placed before them, keep them constantly oriented towards these physical aspects and each other. This orientation, in turn, keeps the directed task near the foreground of their consciousness. What we have then are conversations which illustrate how the conscious mind is both constrained by its orientation and simultaneously left free to make innovative and unpredictable associations on the fly. Orientation can be thought of as a stabilizing, guiding force, from which our cognitive creativity may attempt to spring free.

6. Ideational momentum

In order to properly understand the coherent flow of ideas in examples such as (2) above, regardless of whether there appears to be an identifiable 'topic', I propose the notion that ideas have *inertia*, and that the act of communicating ideas imparts conversation and the ideas themselves with a sort of *ideational momentum*. The notion of momentum implies that the participants' contributions will carry both *magnitude* and *direction*, meaning that the trajectory of conversation can be altered either gently or violently, in any direction, resulting in a path which maps a much more descriptive picture of the course of a conversation than a simple framework of topics divided by boundaries.

Chafe has recognized that a conversation can attain "momentum", but he describes this momentum as something that "sustains topics" (Chafe 1994:127). This is one of the things momentum is capable of, but I would rather describe it in more general terms, as the strength with which ideas cause conceptual associations to other ideas, whether or not these associations are 'on topic'. The human mind seems to involuntarily draw links from one idea to the next, in a sort of internal "association engine". Once communicated in a conversation, an idea becomes foregrounded in the consciousnesses of the participants, who each interpret it in their own way. Depending on the momentum it imparts to each person, the participants energetically or weakly draw associations to various related ideas, including those which may not be at all relevant to the current discussion. In many interactional situations, participants are expected to ignore these irrelevant associations and base their contributions not only on what was just in focus, but also on salient pieces of information still in semiactive consciousness. The likelihood of divergence increases with the number of participants, both because of the increased variety in conceptual associations being made, and because the separate consciousnesses of separate minds, even as they continually reorient to each other, still have a limited ability to synchronize.

Ideational momentum and a flexible notion of conversational trajectory still account for traditional points of 'topic shift': they are simply the points where the momentum and trajectory change most suddenly. What have so often been called 'topics', we could assert, are fundamentally segments of smoothly changing ideational momentum. Whether or not the discussion remains 'on topic' throughout one of these segments is a consequence not wholly under the participants' control. A successful 'topic' in Chafe's sense requires an actual desire to maintain, for a time, a discussion with conscientiously maintained boundaries, and also depends on the ability of the participants to overcome potentially divergent contributions.

We noted earlier that Chafe observed stronger or weaker boundaries between sections of talk to be points of stronger or weaker reorientation, respectively. To account for this observation, we can claim that the degree of required reorientation, and the amount of effort exerted, is proportional to the degree of change in ideational momentum. Given the limited attention of consciousness, shifts in trajectory that occur gradually over large periods of time will be less noticeable (and require less sudden reorientation) than shifts that occur quickly, and they may not even be perceived by the participants. Along the same lines, a shift from silence to talk forces the magnitude of ideational momentum to shift suddenly from zero to a positive value, requiring reorientation. When the momentum has reached zero (as can occur during long silences or when the participants seem to "run out of steam"), equal effort is required to restart talk in *any* direction—there is no longer a bias towards following the trajectory of the most recent talk. As expected, points of zero momentum are typical places for speakers to suddenly alter the discussion. Park-Doob (2001) describes many cases not shown here, but one spot in (2) above should suffice to illustrate the point. It is hard to depict on paper, but right before A brings up the idea of helicopters, the participants "run out of steam", a feeling which B intones with his low *um=*. Participant A takes up this cue and cautiously starts discussion down a new trajectory. Her repetition of *or—* signals her efforts to begin reorienting herself and her interlocutor to deal with brand new material.

7. Conclusion

'Discourse topics', in any sense, are simply not a *requirement* for coherent language use. It takes special effort to bring about and maintain them, and their use arises out of a social need to discuss specific groups of ideas thoroughly and completely, rather than out of a cognitive requirement that discourse proceed this way. With the concept of ideational momentum, I have attempted to describe how participants in conversation navigate from one idea to the next, following a trajectory that is largely unpredictable, yet illustrates many creative cognitive associations and processes that seem to form an integral part of human consciousness.

Transcription conventions *(paraphrased from Chafe 1994 and 1997)*

´	Primary accent (pitch deviation accompanied by loudness or lengthening)
`	Secondary accent (for syllables with qualitatively less-than-primary stress)
..	A very brief (on the order of 0.1 seconds or less) pause or break in timing
...	A "typical" pause (between around 0.1 and 0.7 seconds for most speakers, at times up to one second in places where this is still unmarked)
...(2.4)	A longer pause (measured in seconds)
#	A pause coinciding with an audible intake of breath
'	Marks points where letters have been omitted from normal orthography in order to represent the pronunciation more accurately (beyond a generic reading in American English); also used for standard contractions
=	Lengthening of the preceding vowel or consonant sound (ignoring "silent" elements in the orthography; for example, 'there=' 'yeah=' 'uh=')
,	Marks the end of an intonation unit as having a terminal contour which is not sentence-final
.	Marks the end of an intonation unit as having a sentence-final falling pitch
—	Marks an intonation unit or word as aborted or suspended
¿ ?	Encloses stretches of speech that have a rising pitch contour typical of a yes-no question (includes many instances of "up-talk")
@	A pulse of laughter or giggling
@ @	Encloses words spoken while laughing or giggling
^{2}[ ]^{2}	A segment of speech that overlaps with another segment uttered by a different speaker (indexed by common superscripts).
{ }	Encloses a doubtful transcription

References

Brown, Gillian & George Yule. 1983. *Discourse Analysis*. Cambridge: Cambridge University Press.

Chafe, Wallace. 1994. *Discourse, Consciousness, and Time: The flow and displacement of conscious experience in speaking and writing*. Chicago: The University of Chicago Press.

Chafe, Wallace. 1997. Polyphonic topic development. In Talmy Givón (ed.) *Conversation: Cognitive, communicative and social perspectives*. Amsterdam: John Benjamins.

Clark, Herbert H. 1996. *Using Language*. Cambridge: Cambridge University Press.

Givón, Talmy (ed.). 1997. *Conversation: Cognitive, communicative and social perspectives*. Amsterdam: John Benjamins.

Jakobson, Roman. 1960. Closing statement: Linguistics and poetics. In Thomas A. Sebeok (ed.) *Style in Language*. Cambridge, MA: The MIT Press. Excerpts reprinted in Adam Jaworski & Nikolas Coupland (eds.) *The Discourse Reader*. New York: Routledge.

Jaworski, Adam & Nikolas Coupland (eds.). 1999. *The Discourse Reader*. New York: Routledge.

Keenan, Elinor Ochs & Bambi Schieffelin. 1976. Topic as a discourse notion: A study of topic in the conversations of children and adults. In Charles N. Li (ed.) *Subject and Topic*. New York: Academic Press.

Linell, Per. 1998. *Approaching Dialogue: Talk, interaction and contexts in dialogical perspectives*. Amsterdam: John Benjamins.

McNeill, David. 1992. *Hand and Mind: What gestures reveal about thought.* Chicago: The University of Chicago Press.

McNeill, David (ed.). 2000. *Language and Gesture*. Cambridge: Cambridge University Press.

Nakatani, Christine H., Barbara J. Grosz, David D. Ahn & Julia Hirschberg. 1995. Instructions for annotating discourses. Technical Report TR-21-95, Center for Research in Computing Technology, Harvard University.

O'Connell, Daniel C. & Sabine Kowal. In press. Psycholinguistics: A half century of monologism. *American Journal of Psychology*.

Park-Doob, Mischa. 2001. *Deconstructing 'topic': Relevance, consciousness, and the momentum of ideas*. Unpublished Master's thesis, University of Chicago.

Rommetveit, Ragnar. 1974. *On message structure: A framework for the study of language and communication*. London: Wiley.

5610 Hammersley Rd
Madison, WI 53711

maparkdo@alumni.uchicago.edu

Tonal Polarity as Phonologically Conditioned Allomorphy in Mundurukú

GESSIANE PICANÇO
University of British Columbia, Canada/Museu Paraense Emílio Goeldi, Brazil

0. Introduction

The notion of tonal polarity refers to a phenomenon where a morpheme is assigned a tone opposite to an adjacent tone. At issue is whether polarization equals dissimilation. Schuh (1978) and Newman (1995) propose that morphemes may be polar by nature; i.e. their surface tone is determined exclusively by the context in which they occur. If there exists evidence to presume that these morphemes have an underlying tone, then the process is one of dissimilation. For example, in the Guddiri dialect of Hausa, the diminutive *ɗan* (masc.) shows polarity: *ɗán ràagóo* 'a small ram'/*ɗàn yáaròo* 'a small boy'. However, the fact that the diminutive exists in the language as an independent (H-tone) noun *ɗán* 'son' suggests that this is a case of dissimilation rather than true tonal polarity (Newman 1995; Schuh 1978). At another extreme, Kenstowicz, Nikiema, and Ourso (1988) propose that polar tones are underlyingly H in all languages and the apparent polarity is in fact the result of dissimilatory rules. Pulleyblank's (1986) analysis of tonal polarity in Margi treats polarizing morphemes as having floating H tone underlyingly, but they are lexically marked as extratonal. Extratonality in conjunction with a rule of H-deletion generates the polarity effect. A more recent, constraint-based account has been proposed by Suzuki (1998). Tonal polarity is a dissimilatory process that results from the requirements of two Generalized OCP constraints – one prohibits a sequence of H-tones, (*H...H), while the other prohibits a sequence of L-tones, (*L...L), in a given domain.

The question "true tonal polarity or dissimilation?" seems far from having a definitive answer and I will not pursue it here. Tonal polarity, although common in tone languages, is never a general phonological rule; only few items, mostly affixes, participate in this process. In this paper, I examine tonal polarity in Mundurukú, a Tupi language spoken in Brazil. Mundurukú contains a set of nouns that show polarity in a particular context, but L otherwise. After examining its properties, I propose that the phenomenon is best captured in terms of phonologically conditioned allomorphy (Kiparsky 1994). My proposal asserts that

Optimality Theory (Prince & Smolensky 1993) can properly account for the distribution of allomorphs. I will demonstrate that selection of morpheme variants is determined by PARSE-MORPH (Akinlabi 1996), and that Alignment constraints (McCarthy & Prince 1993) and constraint conjunction (Crowhurst & Hewitt 1997) are required to ensure that allomorphs are selected according to their appropriate environments. Finally I will compare the analysis with that advocated in Suzuki (1998), and show that the allomorphy approach is to be preferred because it successfully explains tonal polarity as well as other tonal processes observed in Mundurukú.

1. Tonal polarity in Mundurukú

Every vowel in Mundurukú has either high or low tone on the surface. At the underlying level, however, there is a three-way contrast, /H, L, ∅/ (Picanço 2002).[1] Mundurukú contains a small number of inalienable nouns that surface on a tone opposite to that of an adjacent syllable. Tonal polarity exhibits certain properties, as described below.

1.1. Tonal polarity is idiosyncratic

Mundurukú monosyllabic inalienable nouns may be divided into two basic tonal groups: nouns that have L tone and those that show tonal polarity.[2] For instance, *təp* 'leaf/CLS', *tə̨y* 'tooth', and *ʔa* 'head/CLS' are H following L and L following H, as shown in (1-3).

(1) (a) ákò də́p 'banana leaf'
banana leaf/CLS
(b) bórǫ́ də̀p 'cotton leaf'
cotton leaf/CLS

(2) (a) ò-nə̨́y 'my teeth'
1sg-tooth
(b) òcé-nə̨̀y 'our teeth'
1pl.excl.-tooth

(3) (a) tópà-ʔá 'his/her forehead'
face-CLS
(b) wìtá-ʔà 'stone'
stone-CLS

They differ from the other group, which is always L-toned.

[1] Abbreviations: CLS=classifier; NOM=nominalization; DIM = diminutive; RED=reduplication; pl=plural; sg=singular; excl=exclusive; v́=high tone; v̀=low tone; y̨=nasality; y̰=laryngealization.

[2] There are few inalienable nouns that seem to have H-tone: *-kə́* 'cultivated field' and *–có* 'basket'.

(4) (a) pà̰ 'arm/CLS' → ò-bà̰ 'my arm'
1sg-arm
òcé-bà̰ 'our arms'
1pl.excl.-arm
(b) ȩ̀n 'flesh' → bíò ȩ̀n 'tapir's flesh'
tapir-flesh
sàpòkáy ȩ̀n 'hen's flesh'
hen-flesh

1.2. Tonal polarity is peripheral

Pulleyblank (1986:214) observes that "polarity effects occur at the edges of a domain." (See also Archangeli & Pulleyblank (1994).) Mundurukú confirms the assumption that tonal polarity is peripheral. Many of the inalienable nouns can incorporate to the verb, as the subject in the case of intransitive or descriptive verbs and as the object in the case of transitive verbs (Gonçalves 1987), or be combined to form compounds. When attached to the verb, polarizing nouns appear in preverbal position where they do not manifest tonal polarity, surfacing L toned by default; for instance, *təp* 'leaf/CLS' in (5).

(5) (a) ò-tə̀p-cócó 'I saw a leaf'
1sg-leaf-see.RED
(b) tə̀p-bòŋ át tə̀p 'big leaf'
leaf-be.big NOM leaf

Similarly, in combinations of two or more polar nouns, only the rightmost element polarizes while the preceding ones surface L toned. As seen in (6a), *-bə* 'finger/CLS' occurs at the right periphery where it is H following the possessive prefix *o-* '1sg'; if another morpheme is added, for example *-nə̧* 'nail' in (6b), *-bə* surfaces L and *-nə̧* H. The polarity effect is blocked by *-ʔítʔít* 'Diminutive' in (6c), causing *-nə̧* to surface L.

(6) (a) ò̰-bə́ 'my finger'
1sg-finger/CLS
(b) ò-bə̰̀-nə̧́ 'my fingernail'
1sg-finger/CLS-nail
(c) ò-bə̀-nə̧̀-ʔítʔít 'my little fingernail'
1sg-finger-nail-DIM

The important generalizations are: (i) tonal polarity affects a small number of morphemes, (ii) which surface H only at the right periphery of a given domain, but (iii) L otherwise. I will refer to the domain where tonal polarity shows up as being the phonological word (PhWd), which is defined here as the domain that coincides with morphosyntactic boundaries within which phonological processes apply (Hall 1997). The examples below illustrate polarization in two contexts. In (7a), *təp* occurs within the word as a classifier (leaf-like object); in (7b), it is the head of a noun phrase. For present purposes, I assume that either construction may constitute a phonological word.

(7)	(a) j̹́wáp-tə̀p	'ray'
	wèrè-də́p	'mushroom'
	(b) bórǫ́ də̀p	'cotton leaf'
	ákò də́p	'banana leaf'

1.3. Tonal polarity does not distinguish between underlying and derived tones

Tonal polarity does not function as a general phonological rule. First, it applies only to a subset of inalienable nouns; i.e. it is morpheme-specific. The phenomenon must be distinguished from another, more general process, involving dissimilation of L tones. Lexical L-tones trigger dissimilation of a following L, changing it to H, to satisfy the Obligatory Contour Principle (Leben 1973; Goldsmith 1976; McCarthy 1986; Odden 1986). This is illustrated in (8b). Derived L tones are inert as triggers, as shown in (8a).

```
(8)   (a) ako-pa̰   →  ákòbà̰     'banana'
          |    |
          H    L       (banana-CLS)
      (b) waje-pa̰  →  wàjèbá     'cacao'
            \|  |
             L  L      (cacao-CLS)
```

Conversely, polarization does not distinguish between lexical and derived tones. For example, *təp* 'leaf/CLS' is H following not only lexical L-tones (9b) but also derived ones (9a).[3]

```
(9)   (a) ako + təp  →  ákò də́p   'banana leaf'
          |
          H
```

[3] Newman (1995) reports that in Standard Hausa the 'stabilizer' *nee/cee* shows polarity whether or not the preceding word has tonal variants.

(a) jàakíi nèe 'it's a donkey'

(b) rìigáa cèe 'it's a gown'

(c) kèekè née = kèeké(e) nèe 'it's a bicycle'

Similar pattern is found in Margi as well (Pulleyblank 1986; Hoffmann 1963).

(b) wąję + təp → wàjè dә́p 'cacao leaf'
L

1.4. Tonal polarity does follow from general patterns

The general pattern in Mundurukú is that sequences of H-tones originating from distinct morphemes require no strategy to repair the OCP violation, as (10) illustrates. Despite this, we can assume that the OCP still plays a significant role in Mundurukú though in a more restrictive way.

(10) (a) ərə̰-ʔit.RED → ə́rə̰́ʔítʔít 'little maracanã'
H H (maracanã bird-DIM)

The core properties of tonal polarity in Mundurukú inalienable nouns are summarized below.

(11) Tonal polarity in Mundurukú
- (a) It is idiosyncratic (i.e. morpheme-specific), applying only to nouns that are marked to undergo it.
- (b) It is peripheral; i.e, it is restricted to the right edge of PhWd.
- (c) It does not distinguish between lexical and derived tones.
- (d) It does not follow from general constraints which tones are or are not subject to in the language.

2. Tonal polarity as phonologically-conditioned allomorphy

We have seen that Mundurukú tonal polarity applies arbitrarily to a small group of nouns rather than applying as a general phonological rule. Besides, these nouns manifest polarity in a particular context, surfacing L otherwise. The distribution of polarizing nouns in Mundurukú could be stated as follows: "the H-tone variant occurs at the right edge of PhWd, after L; the L-tone variant occurs elsewhere". Statements like this describe often-cited cases of phoneme variants (allophones) in Phonology or morpheme variants (allomorphs) in Morphology. Dealing with tonal polarity or similar phenomena as allomorphy may be a plausible alternative. I will here explore this possibility, asserting that Mundurukú tonal polarity is the result of phonologically-conditioned allomorphy.

2.1. Selection of allomorphs

Kiparsky (1994), who adopts the selection method advocated in Lieber (1982) and Zwicky (1986), argues for a model where allomorphs are lexical items with one unmarked, default alternant, and other, lexically marked, that is restricted to appropriate contexts. He characterizes the factors that condition allomorphy in terms of contextual or internal selection. Contextual selection may be either morpholexical (e.g. oxen), or phonological because "just as a 'morpheme' can be restricted to a particular *phonological* environment, so can an 'allomorph'."[p.17].

I will adopt Kiparsky's contextual phonological selection to account for the distribution of polarizing nouns in Mundurukú by arguing that these nouns have two input forms – one is toneless and the other is H-toned, illustrated in (12) – which must obey conditions imposed by the phonology in order to be realized. The specific/marked case is the H-tone variant, which must be selected only to the right edge of PhWd, after L. The toneless form is the general/unmarked case, occurring elsewhere.

(12) leaf/CLS = {tə́p, təp}
finger/CLS = {pə́[ʔ], pə[ʔ]}[4]

The relationship between the general case and the specific case is expressed by a dominance relation. In optimality-theoretic terms, allomorphs are ranked with respect to one another (Kager 1996). By analogy with the Pãnini's Theorem on Constraint-ranking (Prince & Smolensky 1993), I suggest that if the specific and general allomorphs are in conflict, then the more specific must dominate the more general case, as in (13). The factors that will determine which allomorph must be taken to be more specific or more general are language-specific products.[5]

(13) tə́p >> təp

The question then is how the selection of allomorphs should be handled by a constraint-based approach such as the Optimality Theory. The distribution of allomorphs, if phonologically conditioned, cannot be unrestricted; on the contrary, it must be accomplished by universal considerations. Viewed in this way, allomorphs, like features or featural affixes, need to be licensed to be phonetically realized. Their surface realization depends upon the restrictions imposed by the phonology to, generally, the more specific case. If such conditions are not satisfied, then the general case must occur instead.

I hypothesize that it is imperative to ensure that input forms of morpheme variants are realized in the output. This requirement is compelled by a family of constraints, namely PARSE-MORPH (Akinlabi 1996: 247), formally defined below, with particular formulations in (15).

(14) PARSE-MORPH – A morph must be realized in the output.

[4]Mundurukú has a process of laryngealization in which certain morphemes trigger laryngealization on the preceding vowel under certain circumstances. For purpose of this study, I will distinguish this process by using [ʔ] in the underlying form of the morpheme.

[5] See Hargus (2000) for a different proposal. According to her, the preferred allomorph is always the phonologically shortest variant, which is selected by a universal constraint called BREVITY.

(15) (a) PARSE-tə́p - tə́p must be realized in the output.
(b) PARSE-təp - təp must be realized in the output.

Ranking of PARSE-MORPH constraints must reflect dominance relations between allomorphs. Thus, the H-tone variant, e.g. PARSE-tə́p, dominates the toneless varaint, e.g. PARSE -təp, in Mundurukú.

(16) PARSE -tə́p >> PARSE -təp

The tableau in (17) shows selection of the more specific over the general case.

(17) Selection of the specific case: PARSE -tə́p >> PARSE -təp

ako {təp; təp} | | H H	MAX PATHH	PARSE-tə́p	PARSE-təp
a) ☞ ákò də́p			*
b) ákò də̀p		*!	
c) àkó də̀p	*!	*	

The domination relation PARSE-tə́p >> PARSE-təp says that the H-tone allomorph must be selected first, as seen in the optimal candidate (17a). Candidate (17b) is penalized by selecting the unmarked allomorph instead. MAXPATHH prohibits loss of tone specifications, eliminating candidate (17c).[6]

The tableau in (17), however, does not show entirely how the restrictions on allomorph distribution are handled. There are two conditions on the distribution of the specific case in Mundurukú: (i) it is restricted to the right edge of PhWd; and (ii) it obeys the OCP constraint prohibiting a sequence of H-tone (*HH). The former is expressed here as an Alignment constraint (McCarthy & Prince 1993) demanding coincidence of a H-tone allomorph with the right edge of PhWd.

(18) ALIGN-tə́p – ALIGN (tə́p, Right, PhWd, Right)
tə́p must be aligned with the right edge of a phonological word

While PARSE-tə́p demands the realization of tə́p in the output, ALIGN-tə́p limits its occurrence at a given edge. This is illustrated in the following tableau. ALIGN must dominate PARSE in order to ensure that the specific case does not surface if it is not aligned with the right edge of PhWd.

---

[6] MAXPATHH – Any input path between H and an anchor must have a correspondent path in the output. (See Pulleyblank (1996))

(19) Selection of the general case

{təp; təp}-boŋ at | | | H L H	MAX PATHH	ALIGN-tə́p	PARSE-tə́p	PARSE-təp
a) tə́pbòŋ át		*!		*
b) ☞ tə̀pbòŋ át			*	

Note, however, that ranking ALIGN-tə́p above PARSE-tə́p does not prevent the specific case from being selected after a H tone, predicting, for example, **bóró̜ də́p* instead of the actual form *bóró̜ də̀p* 'cotton leaf'. This is because there is an additional restriction on the distribution of the specific case in Mundurukú, namely that it obeys the OCP constraint *HH. To capture this requirement, suppose that PARSE is conjoined with the OCP constraint *HH. The proposal is not new, constraint conjunction (Smolensky 1997; Itô & Mester 1996; Alderete 1997) has played an important role in accounting for phonological facts that could not be properly explained otherwise. Crowhurst & Hewitt (1997) argue for a model of disjunction as positive conjunction where a candidate *passes* a conjunction if and only if it passes every conjunct. If a candidate fails to satisfy one of constraints from the conjunct, then it fails to satisfy the conjunct (cf. Smolensky 1997). Following Crowhurst & Hewitt, I propose the following conjunction:

(20) PARSE-tə́p ^ *HH

PARSE-tə́p requires the specific allomorph to be realized in the output, but this requirement is accomplished by the OCP constraint *HH. The proposed conjunction achieves the correct results: it fails to select the H-tone allomorph when adjacent to a H-tone base, and it captures the fact that only certain morphemes must obey *HH. The candidates in (21a, b) are non-optimal as both violate the conjunct. Candidate (21a) violates PARSE-tə́p and (21b) violates *HH. Since neither candidate passes the conjunction, they are evaluated by PARSE-təp which selects the default case instead.

(21) Selection of the general case

boro̜ {təp; təp} H H	ALIGN-tə́p	PARSE-tə́p ^ *HH	PARSE-təp
a) ☞ bóró̜ də̀p		*	
b) bóró̜ də́p		*	*!

When an input form contains two or more polarizing nouns, the hierarchy ALIGN >> PARSE-SC (specific case) >> PARSE-GN (general case) predicts the attested results. Although PARSE-SC dominates PARSE-GC, no dominance relations can be established between one particular instantiation of PARSE-SC/ PARSE-GC and another instantiation of PARSE-SC/ PARSE-GC.

Consider, for instance, the case of *ò-bə̰̀-nə̹́* 'my fingernail' in which both *-bə* and *-nə̹* are polar. We cannot say that constraints evaluating *pə́* dominate those evaluating nə̹ or vice-versa. Instead, let us assume that these constraints are left unranked. This is illustrated in the tableau in (22). Each set of allomorphs is evaluated by particular formulations of the general hierarchy, but no one is ranked over the other for the following noun. The specific case {pə́[ʔ]} cannot be selected in (22a,b) because it is not at the right edge of a phonological word. Candidates (22c,d) both fail PARSE-pə́, but (22c) is optimal because it satisfies the conjunction established for {nə̹́[ʔ]}.

(22) Input: /o-{pə́[ʔ];pə[ʔ]}-{nə̹́[ʔ]; nə̹[ʔ]}/ → [òbə̰̀nə̹́] 'my fingernail'

Hierarchy:	ALIGN-SC		PARSE-SC ^ *HH		PARSE-GC	
	ALIGN-pə́	ALIGN-nə̹́	PARSE-pə́ ^ *HH	PARSE-nə̹́ ^ *HH	PARSE-pə	PARSE-nə̹
a) ò̰bə́nə̹̀	*!			*	*	
b) ò̰bə́nə̹́	*!		*	*	*	*
c) ☞ òbə̰̀nə̹́			*			*
d) òbə̀nə̹̀			*	*!		

In this section I showed that Mundurukú tonal polarity can be analyzed as phonologically conditioned allomorphy. I proposed that rankings of allomorphs conforms the Pãnini's principle on constraint rankings – the more specific dominates the more general case. The selection of allomorphs is determined by PARSE-MORPH along with Alignment constraints and constraint conjunction. The interaction of these constraints handles the disjunctive pattern of allomorph distribution, selecting each variant to a context that is compatible with it.

3. A comparison with the GOCP approach to tonal polarity.

Suzuki (1998) claims that tonal polarity derives from the combination of two Generalized OCP (GOCP) constraints, given in (23).

(23) GOCP constraints (adapted from Suzuki 1998: 142)
$(*H...H)_{\text{Domain}}$ – A sequence of H-tone is prohibited within a given domain.
$(*L...L)_{\text{Domain}}$ – A sequence of L-tone is prohibited within a given domain.

The GOCP approach makes wrong predictions in Mundurukú. First, we have seen that the OCP prohibits a sequence of L-tone, but only lexical L-tones trigger dissimilation of a following L (e.g. wàje-bá 'cacao' *but* ákò-ba̰ 'banana'). This generalization is obscured in Suzuki's proposal, which predicts, for example, that the L-tone of *-bà̰* in [ákò-bà̰] 'banana' should be realized as H.

(24) Mundurukú and the GOCP approcah

áko + təp	(*L...L) (*H...H)
a) ☞ ákò dəp	
b) ákò də̀p	*!
áko-pà̰	
c) ☞ ákò-bá	
d) ☹ ákò-bà̰	*!

Second, it predicts tonal polarity everywhere and, as I showed, polarization is restricted to the right edge of a phonological word in Mundurukú. In other contexts, polarizing nouns get L by default. Take again the case of [ò-bə̰̀-nə̰́] 'my fingernail'. Three candidates are listed in the tableau in (26). The GOCP constraints select (26a) over the actual output (26c).

(26) Sequences of polar morphemes

o+pə[ʔ]+nə̰[ʔ]	(*L...L) (*H...H)
a) ☞ òbə́nə̰̀	
b) ☹ òbə̰̀nə̰́	*!
c) òbə̀nə̰̀	*!*

The analysis pursued thus far, on the other hand, provides not only a good account of tonal polarity in Mundurukú but also generalizes to other tonal processes observed in the language. For instance, the possessive prefix has two surface realizations: *é-* in word-initial position, (27a); *è-* otherwise, (27b).

(27) (a) àyácát é-kò̰bé 'woman's canoe'
woman POSS-canoe

(b) w-è-kò̰bé 'my canoe'
1-POSS-canoe

Under the proposal presented here, the analysis of the possessive prefix is straightforward. The prefix has two allomorphs /é-; e-/ ranked as é- >> e-. The specific allomorph, é-, must be left-aligned with a word. ALIGN-é- dominates PARSE-é-, which dominates PARSE-e-, yielding the ranking ALIGN-é- >> PARSE-é- >> PARSE-e-. The specific case shows no other restriction, consequently PARSE-é- requires no conjunction with a markedness constraint. As shown in the following tableau, the same analysis proposed for polarizing nouns also accounts for the distribution of variant forms of the possessive prefix.

(28) Selection of {é-; e-}

N {é-; e-}-N	ALIGN-é-	PARSE-é-	PARSE-è-
a) N è-N		*!	
b) ☞ N é-N			*
PREF-{é-; e-}-N			
c) ☞ PREF-è-N		*	
d) PREF-é-N	*!		*

4. Conclusion

I argued that the generalizations involving tonal polarity in Mundurukú can be captured in terms of phonologically conditioned allomorphy. The marked allomorph is the H-tone variant and is to be selected to an appropriate context; the default variant occurs elsewhere. I suggested that selection of allomorphs is determined by the PARSE-MORPH family of constraints demanding that each variant be realized in the output. Ranking of PARSE constraints conform that of allomorphs, namely the specific case must dominate the general case. The hierarchy ALIGN-SC>> PARSE-SC>> PARSE-GC not only successfully accounts for tonal polarity in Mundurukú - including sequences of polar morphemes – but also generalizes to other cases of allomorphy observed in the language.

References

Akinlabi, A. 1996. Featural affixation. *Journal of Linguistics* **32**: 239-289.

Alderete, John. 1997. Dissimilation as local conjunction. *NELS 27*: 17-31.

Archangeli, D. & D. Pulleyblank. 1994. *Grounded Phonology*. MIT Press.

Crowhurst, M. & M. Hewitt. 1997. Boolean Operations and Constraint Interaction in Optimality Theory. ROA-229.

Goldsmith, J. 1976. *Autosegmental Phonology*. Garland Press.

Gonçalves, C.H. 1987. *Concordância em Mundurukú*. Editora da Unicamp, Campinas.

Hall, T.A. 1997. The phonological word: A review. In T. A. Hall & U. Kleinhenz (eds.), *Studies on the Phonological Word*. John Benjamins Publishing Company.

Hargus, S. 2000. Prosodically Conditioned Allomorphy in Optimality Theory: Brevity in Witsuwit'en. Ms., University of Washington.

Hoffmann, C. 1963. *A Grammar of the Margi Language*. London: OUP.

Itô, J. & A. Mester. 1996. Rendaku 1: Constraint Conjunction and the OCP. ROA-144.

Kager, R. 1996. On affix allomorphy and syllable counting. In U. Kleinhenz (ed.), *Interfaces in Phonology*. Akademie Verlag.

Kenstowicz, M., E. Nikiema, & M. Ourso. 1988. Tone polarity in two Gur languages. *SLS* **18**(1): 77-103.

Kiparsky, P. 1994. Allomorphy or morphophonology? In R. Singh (ed.), *Trubetzkoy's Orphan*. Current Issues in Linguistic Theory 144, 13-31.

Leben, W. 1973. Suprasegmental Phonology. PhD dissertation, MIT, Cambridge, Mass.

Lieber, R. 1982. Allomorphy. *Linguistic Analysis* **10**(1): 27-52.

McCarhty, J. & A. Prince. 1993. Generalized Alignment. Ms., University of Massachusetts, Amherst and Rutgers University.

McCarthy, J. 1986. OCP Effects: Gemination and antigemination. *Linguistic Inquiry* **17**: 207-264.

Newman, P. 1995. Hausa tonology: Complexities in an "easy" tone language. In J. Goldsmith (ed.), *The Handbook of Phonological Theory*. Blackwell.

Odden, David. 1986. On the role of the Obligatory Contour Principle in phonological theory. *Language* **62**(2): 353-382.

Picanço, G. 2002. Tones in Mundurukú Nouns. Ms., University of British Columbia.

Prince, A. & P. Smolensky. 1993. Optimality Theory: Constraint Interaction in Generative Grammar. Ms., Rugters University and University of Colorado at Boulder.

Pulleyblank, D. 1986. *Tone in Lexical Phonology*. Reidel.

Pulleyblank, D. 1996. Neutral vowels in Optimality Theory: A comparison of Yoruba & Wolof. *Canadian Journal of Linguistics* **41**: 295-347.

Schuh, R.G. 1978. Tone rules. In V. Fromkin (ed.), *Tone: A Linguistic Survey*. Academic Press.

Smolensky, P. 1997. Constraint interaction in generative grammar II: Local conjunction, or random rules in universal grammar. Paper presented at Hopkins Optimality Theory Workshop/University of Maryland Mayfest, Baltimore MD.

Suzuki, K. 1998. A Typological Investigation of Dissimilation. PhD dissertation, University of Arizona.

Zwicky, A. 1986. The general case: basic form versus default form. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, 305-314.

Department of Linguistics
Buchanan E-270 – 1866 Main Mall
Vancouver, BC, Canada – V6T-1Z1
University of British Columbia

globato@interchange.ubc.ca

Harmony Drivers: No Disagreement Allowed

DOUGLAS PULLEYBLANK
University of British Columbia

0. Introduction

'Harmony' is a widely attested pattern in natural language, a configuration where within some domain all eligible anchors for some feature bear the same feature value. Typically, a harmony system exhibits a choice between two feature values. Either all anchors within some domain D bear the feature value F or all anchors within D bear the opposite value G. Depending on the theory of features and harmony, both harmonic values may be overtly specified or one may be indicated by the absence of featural specification.

Within derivational autosegmental theory (Clements 1981, etc.), harmony was generally achieved by requiring (a) that lexical representations choose between specifications of either *F* or *G/∅*, and (b) that any lexically specified value be realized throughout domain D via a harmonic rule of spreading. A fundamental distinction was posited between assimilation/harmony and processes such as dissimilation and polarity. Where harmony was analyzed as a type of assimilation (and assimilation was by spreading), dissimilation and polarity were the result of the OCP (McCarthy 1986, etc.), typically resulting from delinking/deletion plus redundancy. Hence assimilation/harmony and dissimilation/polarity were opposites, accounted for by different mechanisms.

With the advent of Optimality Theory (Prince & Smolensky 1993), the consensus on the mechanism for deriving harmony has disappeared, with very little agreement in the literature as to the formal analysis of harmony. The following constitutes a sample of the constraint types proposed to drive harmony.[1] In a large body of work (Kirchner 1993, Akinlabi 1996, etc.), it has been suggested that *alignment* constraints can be applied to features, producing harmony by requiring that some harmonic feature be aligned with the left and right edges of a phonologically or morphologically defined domain: Align(PCat,

[1] Numerous issues in the analysis of harmony are not addressed in this paper, for example, how to characterize the eligible targets and the eligible sources for harmony, how to characterize possible blocking segments and possible transparent segments, how to achieve directionality in harmonic patterns. For formal definitions of the constraints mentioned here, see the cited references.

L/R; MCat/PCat, L/R). A related variant of this approach is *optimal domains theory* (Cole & Kisseberth 1994, etc.). This approach elevates the formal status of harmonic domains, establishing abstract harmonic domains via alignment and requiring that every anchor within such a domain be affiliated with the harmonic feature. While alignment-based accounts provide a direct analog of directionality, *agreement* approaches (Baković 2000, etc.) are inherently directionless, requiring that adjacent segments within some domain have the same value for the harmonic feature. It is quite possible, however, to posit constraints comparable to agreement constraints but with directionality directly encoded. *Spread* constraints (Walker 1998, etc.) do exactly this: Spread-L/R([F], D). A rather different approach to encoding harmony has been proposed in work on *featural prohibitions and positional faithfulness* (Beckman 1995, etc.). By ranking general prohibitions on a harmonic feature above faithfulness (*+F >> *–F >> IDENT[F]), a general pattern is obtained where only one feature value is tolerated (i.e. –F). By also ranking a positional faithfulness constraint (e.g. IDENTσ_1[F]) above the prohibition on +F, harmony is achieved. The last harmony driver to be mentioned was sketched in Smolensky (1993). Smolensky notes that harmony can be achieved by prohibiting disharmony. That is, a sequence of opposite values for some harmonic feature can be prohibited (*F G), with the result that segment sequences must have the same value for the harmonic feature.

This multiplicity of harmony drivers raises interesting questions concerning the properties and learnability of harmony systems. If multiple drivers are possible, then there will be a significant degree of indeterminacy in the analysis of harmony. In this regard, note that several of the approaches to harmony described briefly above involve formal properties of a stipulatory nature. For example, the optimal domains approach requires the postulation of abstract featural domains, while the alignment approach requires equating featural domains with prosodic or morphosyntactic domains, requiring that the 'edges' of featural spans be aligned with such domains. In general, one must ask the question of *why* such constraints should exist. Why is it desirable to have one feature agreeing with another, or to have one feature spread throughout some domain? To what extent does the functional motivation for such a constraint relate to the constraint's formulation?

In exploring here the option of deriving harmony by prohibiting feature disharmony, the functional motivation is that the resetting of articulatory targets costs the grammar. That is, inertia is claimed to be the central functional principle, with changes of articulatory setting prohibited: *ATR RTR, *RTR ATR; *NASAL ORAL, *ORAL NASAL. Several arguments support such an account of harmony (see also Hansson 2002). The no-disagreement account permits the unification of the driver for harmony with the Obligatory Contour Principle, providing a unified account of vowel harmony and consonant harmony.[2] Both opacity and

[2] Note, however, Hansson's (2001) arguments for distinguishing the two harmony types. If the unified approach to the two harmony types is correct, the differences discussed by Hansson must be attributed to properties other than the harmony driver.

transparency are accounted for, as are patterns of edge-conditioned harmony where eligible undergoers of harmony behave as transparent because they are medial. Cases of incomplete harmony can be accounted for, where target features do not match source features. Before turning to the phonological arguments in favor of no-disagreement, I begin by illustrating the basic mechanisms involved.

1. Let's not disagree: the basic harmonic effect

The simplest harmony system requires agreement for the harmonic feature within a domain such as the word. Consider the tongue root harmony system in a language like Degema (Elugbe 1984, Kari 1995, 1997, Pulleyblank et al 1995). This language exhibits a basic 5-vowel system cross-cut by a tongue root distinction. The ten resulting vowels are illustrated in (1), in forms that show the basic harmonic imperative with alternating prefixes and suffixes.

(1) a.	*Advanced*: [i u e o ə]		b.	*Retracted*: [ɪ ʊ ɛ ɔ a]	
[i]	u-bí-ə̄	*state of being black*	[ɪ]	á-kɪ̄	*pot*
[u]	u-pú-ə̄m	*closing*	[ʊ]	ʊ-fʊ́-ā	*state of being white*
[e]	u-dér-ə̄m	*cooking*	[ɛ]	ɔ-ɗɛ́ɗɛ̄	*chief*
[o]	i-sór-ə̄	*passing liquid faeces*	[ɔ]	ʊ-bɔ́m-ām	*beating*
[ə]	o-gədəgə́	*mighty*	[a]	ɔ-kpakɪraká	*tough*

Roots may be either advanced (1a) or retracted (1b); the tongue root value of an affix is the same as the root to which it is attached.

To achieve harmony, prohibitions on sequences of vowels with different values for the tongue root feature mush outrank faithfulness conditions.

(2) *<u>ATR</u>-C_0-RTR: Ignoring consonants, an ATR segment may not be immediately followed by RTR.[3]

*RTR-C_0-<u>ATR</u>: Ignoring consonants, an ATR segment may not be immediately preceded by ATR.

I assume a correspondence approach to faithfulness (McCarthy & Prince 1995) with featural values ensured by 'MAX/DEP FEATURE' and 'MAX/DEP PATH' constraints (Pulleyblank 1996, etc.).[4] MAX/DEP FEATURE constraints are instantiations of the schema in (3), where F-elements range over features and nodes such as {ATR, RTR, ...} and domains include {Root, Word, ...}.

(3) [MAXG]$_D$: Within domain D, an F-element G in the input must correspond to an F-element G in the output.

[3] Violations are assessed by universally quantifying over one of the members of the sequential constraint, in this case, ATR. That is, in this case one violation is assigned per advanced segment violating the constraint. The ATR-orientation is indicated in the constraint name by underlining.

[4] The basic arguments concerning a no-disagreement approach to harmony could also be achieved with segment-based MAX/DEP in combination with 'IDENT'. Violations would be assessed differently but the optimal candidates would be the same for the cases under consideration.

To achieve root control harmony, as seen in Degema, the harmonic constraints must outrank the faithfulness constraints, and within the faithfulness class, root domain constraints must outrank word domain constraints: *ATR-C_0-RTR, *RTR-C_0-ATR >> [MAXATR]$_{RT}$, [MAXRTR]$_{RT}$ >> [MAXATR]$_{WD}$, [MAXRTR]$_{RT}$.[5] Where the distinction is not relevant, I will abbreviate [MAXATR]$_{ROOT}$ and [MAXRTR]$_{ROOT}$ as '[MAX]$_{ROOT}$', and [MAXATR]$_{WORD}$ and [MAXRTR]$_{WORD}$ as '[MAX]$_{WORD}$'.

Consider a case where the root vowel is advanced. Root faithfulness ensures that advancement is preserved (ruling out (4e)); the prohibitions on disagreeing tongue root sequences rule out any candidate with retracted vowels (4b-d).

(4) /U-[der]-Am/	*ATR-C_0-RTR	*RTR-C_0-ATR	[MAX]$_{ROOT}$	[MAX]$_{WORD}$
a. ☞ [uderəm]				
b. [uderam]	*!			
c. [ʊderəm]		*!		
d. [ʊderam]	*!	*!		
e. [ʊdɛram]			*!	*

Similar considerations ensure that all vowels are retracted if the root is retracted.

The basic exigencies of harmony are straightforwardly captured by positing constraints prohibiting featural disagreement. If features cannot disagree, and if the prohibition on disagreement outranks faithfulness, the result is a span with a uniform featural specification.[6]

2. The formal basis for no-disagreement constraints

Analyzing harmony through no-disagreement constraints permits the unification of harmony with the Obligatory Contour Principle ('OCP'). In essence, the OCP can be analyzed as a sub-case of a general prohibition on sequences of elements. Consider the formulation of the Generalized OCP in Suzuki (1998):

(5) *Generalized OCP*: *X…X: A sequence of X is prohibited

Proximity: The closer the elements are the stronger the interaction.

Adjacency: *X…X = {*XX >> *X-C_0-X >> *X-μ-X >> … >> *X-∞-X}

Similarity: The more similar the elements are the stronger the interaction.

---

[5] In line with work such as Mohanan 1993, Archangeli & Pulleyblank to appear, I assume that constraints hold more strongly of small domains than larger domains, deriving the observed root/affix asymmetry. Root control could also be achieved by assuming a constraint on root faithfulness that would outrank a constraint on affix faithfulness (McCarthy & Prince 1995).

[6] For reasons of space, I do not examine issues here concerning 'Richness of the Base' (Prince & Smolensky 1993). It can be established, however, that the results obtained via no-disagreement constraints are consistent with assuming unconstrained inputs. Whether inputs are fully specified or underspecified to some degree, whether features are monovalent or binary, whether features are free or linked – such variables will affect the details of the analysis but not the basic effect of no-disagreement constraints. For concreteness, I will assume representations in which outputs are fully specified for harmonic features. Where the harmonic value for a morpheme is entirely predictable from its context, I will assume an input that is unspecified for the harmonic feature.

By removing the identity requirement between the two elements assessed by this constraint, the OCP becomes just one type of sequential prohibition, with *proximity* and *similarity* conditions holding of all such constraints.

(6) *Sequential prohibitions*: *X...Y: A sequence of X, Y on a tier is prohibited
 OCP: X = Y
 Harmony via No-Disagreement: X ≠ Y

I propose that this class of sequential prohibitions applies to sequences of elements within tiers (Sagey 1986, etc.). Hence an ATR specification cannot be in sequence with another ATR specification (the OCP) nor can it be in sequence with an RTR specification (deriving harmony); a NASAL specification cannot be in sequence with another NASAL specification (the OCP) nor can it be in sequence with an ORAL specification; and so on. Just as the OCP holds more strongly of elements that are close together, so is harmony more rigorously enforced between elements that are close to each other. Just as the OCP holds more strongly between segments sharing multiple features, so does harmony apply more strongly in such cases of shared features.

A sample of the similarities between the OCP and harmony with respect to proximity are sketched in (7), with OCP examples cited from Suzuki (1998).

(7) *Proximity*	*X = Y*	*Example*	*X ≠ Y*	*Example*
Close	*XX	Ainu rhotics	*XY	Local nasal assimilation
Medium	*X-C_0-X	Kera low vowels	*X-C_0-Y	ATR harmony with opacity
Distant	*X-∞-X	Japanese: Lyman's Law	*X-∞-Y	ATR harmony with transparency

In the following sections, I argue for the no-disagreement approach to harmony while presenting cases illustrating different properties of the proposal. To begin with, I compare examples of opacity and transparency contrasting two dialects of Yoruba. In spite of the considerable attention that the opacity/transparency distinction has received in the literature, a reexamination of the question seems warranted. While it is not within the scope of this paper to examine previous approaches in any detail,[7] the proposal here falls into the class where opacity vs. transparency is directly encoded: it falls out directly from the

[7] Opacity may be distinguished from transparency both directly and indirectly. A sample of the direct approaches includes distinctions in rule formulation (strictly local vs. optionally nonlocal) (Jensen 1974), the presence vs. absence of a blocking autosegment (Clements 1981), the selection of a harmonic value that matches/fails to match the value of the transparent class (Goldsmith 1985), the definition of alignment constraints (Pulleyblank 1996, Orie 2001). Various approaches have also been proposed that involve mediated reference to the neutral class – see discussion below. For example, harmony may apply to all segments but be followed by a rule of absolute neutralization (Lightner 1965), the optimal candidate may be 'sympathetic' to a fully harmonic candidate (McCarthy 1999), or 'targeted' constraints may be postulated (Baković & Wilson 2001).

ranking of constraints, with no postulation of abstract stages or representations.

3. Opacity vs. Transparency: Ọyọ Yoruba vs. Ifẹ Yoruba

Standard Yoruba has seven oral vowels {i, e, ɛ, a, ɔ, o, u} and exhibits a pattern of tongue root harmony involving opacity of the consistently advanced high vowels (Archangeli & Pulleyblank 1989, 1994). Orie (2001, to appear) examines two dialects of Yoruba, Ọyọ and Ifẹ. She demonstrates that Ọyọ exhibits a pattern extremely close to that of the standard language, while Ifẹ differs markedly in that high vowels are transparent rather than opaque.

In both Ọyọ and Ifẹ, sequences of mid vowels must agree in their specification for tongue root advancement/retraction.

(8)

Ọyọ Yoruba				*Ifẹ Yoruba*			
oko	*farm*	ọkọ	*husband*	oko	*farm*	ọkọ	*husband*
òkè	*hill*	ọ̀kẹ́	*large bag*	òɣe	*proverb*	ẹ̀dọ̀	*liver*
ehoro	*rabbit*	ọpọlọ	*brain*	ehoro	*rabbit*	ọ̀rọ́rọ̀	*type of bag*
ekòló	*earth-worm*	ẹ̀rẹ̀kẹ́	*cheek*	ògèdè	*incanta-tions*	ọ̀gẹ̀dẹ̀	*banana, plantain*

Such harmonic agreement results from an analysis largely analogous to that of Degema given in §1. One crucial difference, which will be motivated below, is that disharmony can be forced by the need to maintain root RTR specifications while satisfying substantive constraints on retraction and advancement.

(9) [MAXRTR]$_{Root}$ >> *<u>ATR</u>-C$_0$-RTR, *RTR-C$_0$-<u>ATR</u> >> [MAXATR]$_{Root}$

For roots containing mid vowels, the effect of this ranking is full harmony. Moreover, because faithfulness to RTR outranks faithfulness to ATR, harmony initiated by a lexically specified RTR value will be achieved at the expense of any specified ATR value (cf. Archangeli & Pulleyblank 1989). I illustrate this by an example, /ɔgɛde/, where I have arbitrarily assumed that the first two vowels are retracted in the input while the final vowel is advanced.

(10) /ɔgɛde/	[MAXRTR]$_{RT}$	*<u>ATR</u>-C$_0$-RTR	*RTR-C$_0$-<u>ATR</u>	[MAXATR]$_{RT}$
a. [ɔgɛde]			*!	
b. [ɔgede]			*!	
c. [ogede]	*!			
d.☞ [ɔgɛdɛ]				*

The specific cases of interest for the analysis of neutrality are high vowels. Unlike their mid vowel counterparts, high vowels show a single value for the tongue root, namely advancement (Archangeli & Pulleyblank 1989, 1994).

(11) *HI/RTR: A high vowel must be advanced, not retracted.

To absolutely rule out high retracted vowels, *HI/RTR must outrank [MAXRTR]$_{RT}$;

*HI/RTR must also outrank the harmonic constraint *<u>ATR</u>-C$_0$-RTR since harmony is not enforced following high vowels (on mid-high sequences, see below).

(12)	*Ọyọ Yoruba &*	igbe	*noise*	ìgbẹ́	*excrement*
	Ifẹ Yoruba	igbó	*bush*	ilẹ̀	*ground*
		eruku	*dust*	itọ́	*saliva*
		ìrèké	*sugarcane*	ìrọ̀lẹ́	*evening*

The tableau in (13) shows how surface disharmony can result from the appearance of a high vowel in a form with a lexical RTR specification. Whether the high vowel is underlyingly specified as ATR, as in (13), or underlyingly unspecified for a tongue root value, this does not affect the outcome.

(13) /ìgbɛ́/	*HI/RTR	[MAXRTR]$_{RT}$	*<u>A</u>-C$_0$-R	*R-C$_0$-<u>A</u>	[MAXATR]$_{RT}$
a.☞ [igbɛ]			*		
b. [ɪgbɛ]	*!				*
c. [igbe]		*!			

Of interest for neutrality are cases where a high vowel appears flanked by mid vowels. One option for such words should be to have all vowels advanced, the result with inputs with no RTR specifications. Such words exist in both dialects.

(14)	*Ọyọ Yoruba &*	ògùrò	*stick for stirring*
	Ifẹ Yoruba	eúrò	*bitter leaf*
		oríwo	*boil, tumor*
		èbúté	*harbour*

Of more interest are cases discussed by Orie where a mid-high-mid sequence involves retraction. In Ọyọ, as in Standard Yoruba, only the final mid vowel can be retracted. Ifẹ shows a different pattern, with the initial and final mid vowels systematically harmonizing – although the intervening high vowel is unaffected.

(15)	*Ọyọ Yoruba:*	eúrẹ́	*goat*	*Ifẹ Yoruba:*	ẹúrẹ́	*goat*
	opacity	èlùbọ́	*yam flour*	*transparency*	ẹ̀lùbọ́	*yam flour*
		òtítọ́	*truth*		ọ̀títọ́	*truth*
		odídẹ	*parrot*		ọdídẹ	*parrot*

The analysis presented above derives the opacity pattern of Ọyọ.

(16) /ElUbɔ/	*HI/RTR	[MAXRTR]$_{ROOT}$	*RTR-C$_0$-<u>ATR</u>	*<u>ATR</u>-C$_0$-RTR
a.☞ elubɔ				*
b. ɛlʊbɔ	*!			
c. ɛlubɔ			*!	*
d. elubo		*!		
e. ɛlubo			*!	

The high vowel cannot be advanced (16b); the lexically specified RTR value cannot be deleted (16d); an RTR specification must be as far to the right as possible (16c,e). Note that $[\text{MAXRTR}]_{\text{ROOT}}$ governs the retention of RTR but does not govern its location. Hence (16a) emerges as optimal no matter where an RTR value is specified in the input form, provided that the harmony constraints outrank constraints on location (for example, constraints of the 'MAXPATH' class (e.g. Pulleyblank 1996)). The ranking of *RTR-C_0-<u>ATR</u> over *<u>ATR</u>-C_0-RTR ensures the RTR feature occurs on the rightmost vowel (16a), not the leftmost vowel (16e).[8]

Since the analysis so far established derives opacity, some modification is needed to achieve the transparency observed in Ifẹ. Consider three instantiations of the generalized sequential prohibitions with respect to tongue root values:

(17) *Proximal vs. distant sequential prohibitions*

<u>Close</u>:	*RTR <u>ATR</u>:	An ATR segment may not be immediately preceded by RTR.
	*<u>ATR</u> RTR:	An ATR segment may not be immediately followed by RTR.
<u>Medium</u>:	*RTR-C_0-<u>ATR</u>:	Ignoring consonants, an ATR segment may not be immediately preceded by RTR.
	*<u>ATR</u>-C_0-RTR:	Ignoring consonants, an ATR segment may not be immediately followed by RTR.
<u>Distant</u>:	*RTR-∞-<u>ATR</u>:	An ATR segment may not be preceded by RTR.
	*<u>ATR</u>-∞-RTR:	An ATR segment may not be followed by RTR.

In its most local manifestation, the sequential prohibition would disallow any immediately adjacent sequence of differing tongue root specifications (relevant data will be discussed below). In its medium distance form, these prohibitions derive the pattern of opacity seen in Ọyọ. I propose that the extension of the sequential prohibition to long-distance environments is the crucial innovation required to derive transparency of the type seen in Ifẹ.

Consider a tableau for Ifẹ – like (16) but with *distant* prohibition constraints:

(18) /ElUbɔ/	*HI/RTR	$[\text{MAXRTR}]_{\text{RT}}$	*<u>ATR</u>-∞-RTR	*RTR-∞-<u>ATR</u>
a. elubɔ			**!	
b. ɛlubɔ	*!			
c.☞ ɛlubɔ			*	*
d. elubo		*!		

The medium distance no-disagreement constraint *<u>ATR</u>-C_0-RTR assigns exactly one violation to elubɔ in (16): two ATR segments precede the final RTR segment

[8] Two types of candidates should also be mentioned. First, gapped representations must be ruled out: a feature cannot link to two anchors while skipping an intervening anchor. I assume that this is due to the universal ill-formedness of a gapped representation (Archangeli & Pulleyblank 1994, Pulleyblank 1996, Gafos 1996, Ní Chiosáin & Padgett 1997) although it could also be achieved here by a 'NoGap' constraint (Itô, Mester & Padgett 1995). Second, given the high ranking of $[\text{MAXRTR}]_{\text{ROOT}}$, it is also important that the OCP on RTR (*RTR-∞-RTR) outrank root faithfulness. This prevents a form like [ɛlubɔ] from being optimal for an input form having two RTR values.

but only the medial vowel is in the local relation required to violate the constraint. In contrast, the distant formulation of *ATR-∞-RTR is violated twice by elubɔ (18a): each ATR vowel is followed (locally or at a distance) by an RTR vowel. By adopting the distant formulations for Ifẹ, and by ranking *ATR-∞-RTR above *RTR-∞-ATR, the transparent candidate in (18c) is correctly assessed as optimal.

A problem emerges, however. In order for (18c) to be optimal, it is crucial that *ATR-∞-RTR outrank *RTR-∞-ATR – if the two constraints were ranked in the opposite order, then elubɔ (18a) would win over ɛlubɔ (18c). The problem is that this ranking would actually evaluate *ɛlubo as optimal (a candidate not included in (18)). Compare the tableau for Ọyọ in (16) where by ranking *RTR-∞-ATR over *ATR-∞-RTR (the opposite to Ifẹ), elubɔ (16a) wins over *ɛlubo (16e). A careful consideration of the two constraint rankings shows that one ranking will prefer the appearance of RTR on the rightmost vowel (*RTR-∞-ATR >> *ATR-∞-RTR) while the other ranking (*ATR-∞-RTR >> *RTR-∞-ATR) prefers RTR on the leftmost vowel. Neither ranking derives transparency!

A consideration of this type of case reveals an interesting property of the proposed analysis of transparency. By themselves, a ranked pair of distant no-disagreement constraints will simply produce edge-orientation to the left or the right. To achieve transparency, either an additional constraint is needed or a restriction on the effect of the prohibition on RTR values preceding ATR values. For Ifẹ, I propose a constraint on edge-anchoring:[9]

(19) *RTR-∞-ATR]$_{\mathrm{WD}}$: An ATR value at the right edge of the word may not follow RTR.

This constraint rules out candidates like (20c) with a single RTR specification at the left edge of a word. The optimal candidate is the transparency candidate (20b).

(20) /ElUbɔ/	*HI/RTR	*R-∞-A]$_{\mathrm{WD}}$	[MAXRTR]$_{\mathrm{RT}}$	*A-∞-R	*R-∞-A
a. [elubɔ]				**!	
b.☞ [ɛlubɔ]				*	*
c. [ɛlubo]		*!			**

Postulating an edge-orientation constraint makes an immediate prediction for Ifẹ that is borne out. Compare the Ọyọ and Ifẹ forms in (21) which involve mid vowels before a string of one or more word-final high vowels (Orie to appear).

[9] Low vowels require a slight modification of this constraint since both [aCe/o] and [aCɛ/ɔ] patterns are possible. Such patterns indicate that the full form of the edge anchoring constraint is *RTR/NONLO-∞-ATR/NONLO]$_{\mathrm{WD}}$. The addition of the nonlow restriction falls into the class where a no-disagreement constraint applies more strongly between segments that share features. Since all ATR vowels are nonlow, the condition applies to vowels sharing the property of being nonlow.

(21)	*Ọyọ Yoruba:*	ẹbí	*family*	*Ifẹ Yoruba:*	ebí	*family*
	pre-high	ẹ̀rù	*fear*	*pre-high*	èrù	*fear*
	retraction	ẹ̀wù	*clothing*	*retraction*	èwù	*clothing*
	possible	ẹ̀bùrú	*shortcut*	*not possible*	èbùrú	*shortcut*

As discussed by Orie, retraction is possible before a high vowel in Ọyọ:

(22) /ɛbUrU/	*HI/RTR	[MAXRTR]$_{\text{ROOT}}$	*RTR-C$_0$-$\underline{\text{ATR}}$	*$\underline{\text{ATR}}$-C$_0$-RTR
a. ☞ [ɛburu]			*	
b. [ɛbʊrʊ]	*!*			
c. [eburu]		*!		

The cognate forms in Ifẹ, however, show advanced vowels before the high vowel string: retraction of mid vowels before a word-final high vowel is systematically excluded. This result is predicted by a constraint including the edge-orientation constraint. Consider the optimal output in Ifẹ, assuming a mid-high sequence with retraction underlyingly, that is, exactly the input form just seen for Ọyọ.

(23) /ɛbUrU/	*HI/RTR	*R-∞-$\underline{\text{A}}$]$_{\text{WD}}$	[MAXRTR]$_{\text{ROOT}}$	*$\underline{\text{A}}$-∞-R	*R-∞-$\underline{\text{A}}$
a. [ɛburu]		*!			**
b. [ɛbʊrʊ]	*!*				
c. ☞ [eburu]			*		

Whether or not a 'mid...high]' form includes an RTR specification in its input representation, the optimal surface form in Ifẹ will have all vowels advanced.

Does the no-disagreement approach makes transparency too easy to achieve? Assuming that opacity is more common cross-linguistically than transparency, the theory must provide an explanation for this skewing. In fact, numerous factors work against transparency. First, representations with transparency invariably violate the OCP. As such, high ranking of the OCP would produce opacity, not transparency. Similarly, transparency can require the appearance in the output of one or more harmonic feature values not present in the input – that is, DEP can be violated. Hence transparency is only possible if DEP is ranked sufficiently low (Pulleyblank 1996). Also working against transparency, the no-disagreement constraint must be set for its 'distant' value. More local instantiations of no-disagreement derive opacity, not transparency. Independent of the issue of transparency vs. opacity, it is clear that featural interaction is maximized in local environments rather than distant ones (Mohanan 1993, Suzuki 1998). A fourth factor working against transparency is factorial rigidity: while four permutations of the basic constraints *Faith*, *ATR-∞-RTR, and *RTR-∞-ATR produce opacity, only the two rankings with faithfulness above both no-disagreement constraints can produce transparency. A final point concerns the need for edge orientation: even if the basic no-disagreement constraints are ranked appropriately, transparency is only possible if some additional consideration such as edge orientation plays a role. Overall, deriving transparency via no-disagreement constraints is straightforward, but limited in its potential by numerous independent factors.

4. Edge-conditioned harmony with transparency of undergoers

There exists a class of harmonic cases where edge orientation plays a fundamental role: harmony only applies from a trigger at the edge of a domain. These cases are of particular interest when they involve transparency since the transparent segments are perfectly eligible recipients of the harmonic feature. Consider the case of C'Lela (Dettweiler 2000), a Benue-Congo language of Nigeria.

C'Lela exhibits an eight-vowel inventory with 3 high and 5 nonhigh vowels: {i, ɨ, u, e, ɛ, a, ɔ, o}. Harmony is with respect to height, hence roots may exist with high vowels only or with nonhigh vowels only.[10]

(24)	a.	d^{ə}tɨndi	*nest*	g.	kwesa	*show*
	b.	c^{ə}rinɨ	*charcoal*	h.	^{ə}ddakso	*palm (of hand)*
	c.	ɨrmɨ	*man*	i.	c^{ə}gyɔmbo	*eyebrows*
	d.	kumu	*get*	j.	soma	*run*
	e.	k^{ə}pɨru	*flower*	k.	d^{ə}vɛso	*broom*
	f.	dwiri	*hyena*	l.	s^{ə}ʔava	*tongs*

In addition to distributional patterns such as seen in (24), affixes such as the pronominal suffixes of (25) provide evidence of height-based alternations: the suffixes are high after a high root and nonhigh after a nonhigh root.

(25)	a.	-mi/me	*1st sg*	in-mi	*my mother*	cet-me	*my father*
	b.	-vu/vo	*2nd sg*	in-vu	*your mother*	cet-vo	*your father*
	c.	-u/o	*3rd sg*	in-u	*her mother*	cet-o	*her father*
				hɨn-u	*his sibling*	waar-o	*his child*

A consideration of suffixal behavior provides evidence that the alternating vowels are underlyingly high. The direct object pronouns of (26) fall into two classes: (i) forms alternating according to the height-based pattern, (ii) forms exhibiting a height value unconditioned by the root to which it attaches.

(26) a.	mi/me	*1st sg*	buz^{ə}k^ə mi	ɛpk^ə me	*chased/bit me*
			sipk^ə mi	wegaka me	*grabbed/indicated me*
			fumt^{ə}k^ə mi	batk^ə me	*pulled/released me*
b.	vu/vo	*2nd sg*	buz^{ə}k^ə vu	ɛpk^ə vo	*chased/bit you*
			sipk^ə vu	wegaka vo	*grabbed/indicated you*
			fumt^{ə}k^ə vu	batk^ə vo	*pulled/released you*

[10] The raised schwa represents a vowel Dettweiler (2000: 4) describes as "a nonphonemic vowel which serves as a short transition between certain occurrences of consonants in clusters." This brief vowel often appears to be "an echo of the full vowel preceding the consonant." If this vowel is to be represented phonologically, I assume its representation is a featureless mora.

c.	o	*3rd sg*	sipk^ə o	wegaka o	*grabbed/indicated him*
d.	na	*1st pl incl*	sipk^ə na	wegaka na	*grabbed/indicated us*
e.	co	*1st pl excl*	buz^{ə}k^ə co	batk^ə co	*chased/released us*
f.	no	*2nd pl*	buz^{ə}k^ə no	batk^ə no	*chased/released you*
g.	e	*3rd pl inan*	hin^{ə}k^ə e	ked^{ə}k^ə e	*uprooted/picked them*

The generalization is that nonalternating suffixes are nonhigh.[11] This suggests that the alternating class corresponds to the 'high' class: an underlyingly high vowel lowers to mid after a nonhigh vowel but remains high after a high vowel.

The basic analysis of C'Lela depends on two no-disagreement constraints. Since a high-nonhigh sequence is impossible within a root, but tolerated between a root and a suffix, a constraint ruling out such a sequence must only apply root internally. Since a nonhigh-high sequence is ruled out generally, a constraint governing such a sequence must apply within the broader domain of the word.

(27) *[...<u>HI</u>-C$_0$-NONHI...]$_{Root}$: Within a root, a high vowel may not be followed by a nonhigh vowel (ignoring consonants) (*[...<u>I</u>-C$_0$-E...]$_{Rt}$)

*[...NONHI-∞-<u>HI</u>]$_{Word}$: A word-final high vowel may not be preceded by a nonhigh vowel (*E-∞-<u>I</u>]$_{Wd}$)

The word-final restriction will be discussed shortly. The harmonic conditions override faithfulness so they must outrank conditions on the retention of height. I assume root and word domain faithfulness for height analogous to the constraints given for the tongue root in (3), ranked below the harmony constraints.[12]

These constraints rule out height disagreement within the root domain.

(28) /suma/	*[...<u>I</u>-C$_0$-E...]$_{Rt}$	*E-∞-<u>I</u>]$_{Wd}$	[MAX]$_{ROOT}$		[MAX]$_{WORD}$	
			NONHI	HI	NONHI	HI
a. [suma]	*!					
b. [soma]				*		*
c.☞ [sumɨ]			*!		*	
d. [somɨ]		*!	*	*	*	*

For reasons that go beyond C'Lela, I assume that faithfulness to nonhigh values outranks faithfulness to high values (Howe & Pulleyblank 2001), though I will not pursue this issue here. Tableau (28) illustrates the result with a high-nonhigh

---

[11] An apparent exception to this is the third person plural nɨ, e.g. fumt^{ə}k^ə nɨ 'pulled them', batk^ə nɨ 'released them', which Dettweiler shows should be analyzed with an epenthetic high vowel. If nɨ has no vowel underlyingly, the generalization is that underlyingly high vowels alternate while nonhigh vowels do not. Epenthetic [ɨ] can be exempted from the harmonic imperative by ranking the markedness conditions against the insertion of nonhigh vowel above harmony. If we assume a constraint prohibiting the epenthesis of a [−high] vowel (DEPμ[−hi]) and a second constraint prohibiting the epenthesis of a [+high] vowel (DEPμ[+hi]), then the high ranking of DEPμ[−hi] will prevent the epenthesis of a nonhigh vowel: DEP[−hi] >> DEP[+hi], *[...NONHI-∞-HI]$_{Word}$.

[12] [MAXNONHIGH/HIGH]$_{ROOT/WORD}$: Within the root/word, a NONHIGH/HIGH value in the input must correspond to a NONHIGH/HIGH value in the output.

sequence; a comparable result would obtain with a nonhigh-high sequence.

Beyond the root, high-nonhigh and nonhigh-high sequences diverge in their properties. When a nonhigh suffix is attached to a high root, the root-internal prohibition of a high-nonhigh sequence is irrelevant because of its domain restriction and the prohibition of a nonhigh-high sequence is irrelevant because there is no nonhigh-high sequence. The result is therefore suffixal disharmony.

(29) /buz$^{\text{ə}}$k$^{\text{ə}}$ co/	*[.$\underline{\text{I}}$-C_0-E.]$_{\text{Rt}}$	*E-∞-$\underline{\text{I}}$]$_{\text{Wd}}$	[MAX]$_{\text{ROOT}}$ NONHI	HI	[MAX]$_{\text{WORD}}$ NONHI	HI
a. ☞ [[buz$^{\text{ə}}$k$^{\text{ə}}$]$_{\text{Rt}}$ co]						
b. [[buz$^{\text{ə}}$k$^{\text{ə}}$]$_{\text{Rt}}$ cu]					*!	
c. [[boz$^{\text{ə}}$k$^{\text{ə}}$]$_{\text{Rt}}$ co]				*!		*

When a high suffix is attached to a nonhigh root, in contrast, the word-level prohibition of a nonhigh-high sequence eliminates the fully faithful candidate.

(30) /wegaka vu/	*[.$\underline{\text{I}}$-C_0-E.]$_{\text{Rt}}$	*E-∞-$\underline{\text{I}}$]$_{\text{Wd}}$	[MAX]$_{\text{ROOT}}$ NONHI	HI	[MAX]$_{\text{WORD}}$ NONHI	HI
a. [[wegaka]$_{\text{Rt}}$ vu]		*!**				
b. ☞ [[wegaka]$_{\text{Rt}}$ vo]						*
c. [[wigɨkɨ]$_{\text{Rt}}$ vu]			*!**		***	

[MAXNONHI]$_{\text{ROOT}}$ selects the harmonic candidate with altered suffixal height (30b) rather than the alternative candidate modifying root values (30c).

So far, all vowels within a root agree for height, and high vowel suffixes harmonize. Cases with multiple suffixes, however, show an interesting restriction on the applicability of harmony: the word-final vowel is subject to harmony, while nonfinal high suffix vowels are neutral and transparent. Consider (31) where the right-hand column shows that the class markers -i/-e & -u/-o harmonize when they are word-final[13] and the left-hand column shows that these morphemes are high when nonfinal (CM = class marker; ADJM = adjectival suffix).[14]

[13] As seen above, epenthetic vowels do not harmonize even when word-final. Hence the final epenthetic vowels in i-po-jɨ 'CM-new-CM' and u-s^{w}a-wɨ 'CM-big-CM' are high even though the roots are nonhigh. The final vowel can be seen to be epenthetic by comparing the forms that include adjectival suffixes: i-po-i-ne 'CM-new-CM-ADJM', u-s^{w}a-u-ne 'CM-big-CM-ADJM'.

[14] The word-final adjectival suffix (-ni/-ne) harmonizes with the root as expected, appearing as -ni after a high root (0a-c) and as -ne after a nonhigh root (0d-f).

(31)	a.	i-zis-i-ni	*CM-long-CM-ADJM*	i-zis-i	*CM-long-CM*
	b.	u-pus-u-ni	*CM-white-CM-ADJM*	u-pus-u	*CM-white-CM*
	c.	u-rim-u-ni [rim-u-ni]	*CM-black-CM-ADJM*	u-rim-u [rim-u]	*CM-black-CM*
	d.	i-rek-i-ne	*CM-small-CM-ADJM*	i-rek-e	*CM-small-CM*
	e.	u-gʲɔz-u-ne	*CM-red-CM-ADJM*	u-gʲɔz-o	*CM-red-CM*
	f.	u-rek-u-ne [rek-u-ne]	*CM-small-CM-ADJM*	u-rek-o [rek-o]	*CM-small-CM*

The fact that the same morphemes harmonize word-finally but fail to harmonize medially indicates that failure to harmonize must be attributed to the morphemes' position in the word, not to some idiosyncrasy of the morphemes in question.

This conclusion is supported by the behavior of a polysyllabic suffix like -ini/-ine 'perfective aspect'. Within a morpheme, we see that the word-final vowel is subject to height harmony while the medial high vowel is neutral and transparent.

(32)	*High root*	a.	sip-ini	*grap-PERF*
		b.	buz-ini	*chase-PERF*
		c.	fumtɨ-ini	*pull-PERF*
	Nonhigh root	d.	ɛp-ine	*bite-PERF*
		e.	bat-ine	*release-PERF*
		f.	wega-ine	*indicate-PERF*

One last class of cases to consider involves class markers that are not high. As expected, a nonhigh class marker itself is invariable: the only potentially relevant constraint would be the no-disagreement constraint prohibiting nonhigh vowels after high vowels, but that constraint is restricted to the root domain. Where alternations are expected is when a nonhigh class marker precedes a high suffix such as the adjectival marker. Irrespective of the root value in such a case, if the high suffix is word-final, the expectation is that the high suffix should surface as nonhigh after the nonhigh class marker. The examples in (33) confirm these predictions; compare these forms with those in (31) where the -ni/-ne suffix is high or nonhigh depending on the root to which it is attached.

(33) *High*	a.	a-rim-a-ne	*CM-black-CM-ADJM*	a-rim-a	*CM-black-CM*
root	b.	a-zis-a-ne	*CM-long-CM-ADJM*	a-zis-a	*CM-long-CM*
Nonhigh	c.	a-rek-a-ne	*CM-small-CM-ADJM*	a-rek-a	*CM-small-CM*
root	d.	a-gʲɔz-a-ne	*CM-red-CM-ADJM*	a-gʲɔz-a	*CM-red-CM*
	e.	a-sʷa-a-ne	*CM-big-CM-ADJM*	u-sʷa-na	*CM-big-CM*

The cases involving sequences of suffixes are accounted for by the analysis presented for single suffixes. Consider examples where an underlyingly high class marker occurs in different positions in the word. In final position, the class marker will lower when following a nonhigh root: input /rek-u/ gives output [rek-o]. Such an example is analogous to comparable cases seen above; see the tableau in (30). When the same suffix occurs medially, it will fail to harmonize:

(34) /rek-u-ni/	*...I̲-C$_0$-E...]$_{Rt}$	*E-∞-I̲]$_{Wd}$	[MAX]$_{ROOT}$ NONHI	[MAX]$_{ROOT}$ HI	[MAX]$_{WORD}$ NONHI	[MAX]$_{WORD}$ HI
a. [[rek]$_{Rt}$ u-ni]		*!				
b.☞ [[rek]$_{Rt}$ u-ne]						*
c. [[rek]$_{Rt}$ o-ne]						**!
d. [[rik]$_{Rt}$ u-ni]			*!		*	

The crucial comparison is between *rekune* and *rekone* (34b/c). Both candidates satisfy *NONHI-∞-H̲I̲]$_{Wd}$ because there is no word-final high vowel; both satisfy root-domain faithfulness. Distinguishing between the candidates is [MAXHI]$_{WORD}$. While it is necessary to violate word-level faithfulness once in order to satisfy *NONHI-∞-H̲I̲]$_{Wd}$, the second violation of *rekone* (34c) is gratuitous: with the final vowel nonhigh, neither of the no-disagreement constraints forces the additional faithfulness violation that results from lowering of the penultimate vowel.

This pattern of transparency has significant implications. Medial vowels are transparent to harmony, but they are not incompatible with the harmonic feature. This makes the C'Lela patterns problematic for most accounts of transparency. In various instantiations, it is typical to define transparent segments through their incompatibility with the harmonic feature (Kiparsky 1981). This may be achieved by allowing a 'gapped' representation to result from the spreading of a feature over the incompatible segment (e.g. Vago 1988); it may be achieved by allowing the harmonic feature to duplicate itself as it skips the transparent segment (e.g. Archangeli & Pulleyblank 1994, Pulleyblank 1996); it may be achieved through comparison of the winner (which respects the incompatibility) to a candidate that is fully harmonic (but violates the incompatibility) (e.g. sympathy theory (McCarthy 1999) and targeted constraints (Baković & Wilson 2001)); it could be achieved through absolute neutralization (e.g. Lightner 1965). All such approaches to transparency depend on having a transparent class that is incompatible with the harmonic feature. Since C'Lela transparent segments exhibit no such incompatibility – that is, no feature such as backness or roundness makes /i/ or /u/ ineligible to receive a nonhigh feature – they cannot be accounted for by a theory of transparency that depends crucially on incompatibility.

5. Conclusion and implications

This paper addresses the multiplicity of drivers for harmonic behavior in the recent optimality literature by arguing that harmony results from languages attempting to minimize the resetting of articulators, by languages giving in to articulatory inertia. Formally, such inertia is encoded through 'no-disagreement' constraints. These constraints are close cousins to the OCP, with both types of constraints prohibiting sequences of elements, where the prohibitions are most rigorously enforced in local environments and between segments exhibiting shared features. The OCP becomes the logical extreme of a continuum defined by similarity, where the segments being compared are so similar as to be identical.

The implications of driving harmony by the prohibition of difference are numerous, and this paper focuses on implications for transparency. The essence of the proposal is twofold. First, violations of no-disagreement may be forced by a more highly ranked constraint requiring that some segment class have a particular value of the harmonic feature. Proposals basing neutrality on incompatibility have been instantiated in a wide variety of ways in both derivational and constraint-based theories of harmony. The second, more novel, proposal is that neutrality may result from inapplicability. No-disagreement may be defined contextually, with only appropriately positioned segments subject to the constraint; segments intervening between a contextually defined segment and some class of prohibited segments need not be incompatible with the harmonic value.

In closing, I note three additional implications of the no-disagreement approach to harmony, implications that space prevents me from addressing in this paper. First, it is possible to account for cases of incomplete harmony, that is, cases where the features harmonically assigned to some target do not match features of the source. When harmony and assimilation are by spreading, some feature or features of the harmonic source must necessarily be shared by the target output; when harmony is by no-disagreement, then cases involving values of a single feature will be comparable, but cases involving two or more features may result in the optimal harmonic output exhibiting some intermediate value between the prohibited segment and the harmonic source. An example is Setswana (Dichabe 1997; Khabanyane 1991, Clements 1991 on the closely related Sesotho) where harmony creates a three-way height distinction in high vowels on the surface where only a two-way distinction exists underlyingly. A second way in which the no-disagreement approach differs from conventional spreading or agreement analyses is in the explicit expectation that harmony will manifest itself more strongly among segments sharing features – a frequently observed property in harmony systems. To cite a single case involving the tongue root, Lango (Woock & Noonan 1979, Noonan 1992, Smolensky 1993, Archangeli & Pulleyblank 1994) shows a preference for harmony between segments sharing height features. A final point of interest is that vowel and consonant harmony can be unified. It has proven problematic to analyze both consonant harmony and vowel harmony by the same mechanism because of the typically local property of vowel harmony and the apparently nonlocal property of consonant harmony (cf. Gafos 1996). With harmony resulting from no-disagreement, both types of processes can be accounted for by the same constraint type.

References

Akinlabi, Akinbiyi. 1996. Featural affixation. *Journal of Linguistics* **32**: 239-289.

Archangeli, Diana & Douglas Pulleyblank. 1989. Yoruba vowel harmony. *Linguistic Inquiry* **20**: 173-217.

Archangeli, Diana & Douglas Pulleyblank. 1994. *Grounded Phonology*. MIT Press, Cambridge, MA.

Archangeli, Diana & Douglas Pulleyblank. To appear. Kinande vowel harmony: domains, grounded conditions, and one-sided alignment. *Phonology*.

Baković, Eric. 2000. *Harmony, Dominance and Control*. Ph.D. dissertation, Rutgers University.

Baković, Eric and Colin Wilson. 2001. Transparency, strict locality, and targeted constraints. *Proceedings of WCCFL 19*. Cascadilla Press, Somerville.

Beckman, Jill. 1995. Shona height harmony: markedness and positional identity. In J. Beckman, L. Walsh Dickey & S. Urbanczyk (eds.) *Papers in Optimality Theory*, 53–75. Amherst: Univ. of Massachusetts Occasional Papers 18.

Clements, G.N. 1981. Akan vowel harmony: a nonlinear analysis. In G.N. Clements (ed.) *Harvard Studies in Phonology* **2**: 108-177. Department of Linguistics, Harvard University.

Clements, G.N. 1991. Vowel height assimilation in Bantu languages. *Working Papers of the Cornell Phonetics Laboratory* **5**: 37-76.

Cole, Jennifer & Charles W. Kisseberth. 1994. An Optimal Domains theory of harmony. Cognitive Science Tech. Report UIUC-BI-CS-94-02, U. of Illinois.

Dettweiler, Stephen H. 2000. Vowel harmony and neutral vowels in C'Lela. *The Journal of West African Languages* **28**: 3-18.

Dichabe, Seipati Bernice. 1997. Advanced tongue root harmony in Setswana. M.A. thesis, University of Ottawa.

Elugbe, Ben O. 1984. Morphology of the gerund in Dẹgẹma and its reconstruction in Proto- Ẹdoid. *Studies in African Linguistics* **15**: 77-89.

Gafos, Adamantios. 1996. *The Articulatory Basis of Locality in Phonology*. Ph.D. dissertation, The Johns Hopkins University.

Goldsmith, John. 1985. Vowel harmony in Khalkha Mongolian, Yaka, Finnish and Hungarian. *Phonology Yearbook* **2**: 253-275.

Hansson, Gunnar. 2001. *Theoretical and Typological Issues in Consonant Harmony*. Ph.D. dissertation, University of California, Berkeley.

Hansson, Gunnar. 2002. Absolute directionality in output-oriented phonology. Paper presented at the University of British Columbia.

Howe, Darin & Douglas Pulleyblank. 2001. Harmony as faithfulness. Ms. University of Calgary & University of British Columbia.

Itô, Junko, Armin Mester & Jaye Padgett. 1995. Licensing and underspecification in Optimality Theory. *Linguistic Inquiry* **26**: 571-613.

Jensen, John. 1974. A constraint on variables in phonology. *Language* **50**: 675-86.

Kari, Ethelbert E. 1995. The Structure of the Degema Verb. M.A. thesis, University of Port Harcourt.

Kari, Ethelbert E. 1997. *Degema*. Lincom Europa, München-Newcastle.

Khabanyane, Khathatso Evelyn. 1991. The five phonemic vowel heights of Southern Sotho: an acoustic and phonological analysis. *Working Papers of the Cornell Phonetics Laboratory* **5**: 1-36.

Kiparsky, Paul. 1981. Vowel harmony. Ms., MIT.

Kirchner, Robert. 1993. Turkish vowel disharmony in Optimality Theory. Paper presented at the Rutgers Optimality Workshop #1, Rutgers University.

Lightner, Theodore. 1965. On the description of vowel and consonant harmony. *Word* **21**: 244-250.

McCarthy, John. 1986. OCP effects: gemination and antigemination. *Linguistic Inquiry* **17**: 207-63.

McCarthy, John. 1999. Sympathy & phonological opacity. *Phonology* **16**: 331-99.

McCarthy, John & Alan Prince. 1995. Faithfulness and reduplicative identity. In Jill N. Beckman, Laura Walsh Dickey, & Suzanne Urbanczyk (eds.) *Papers in Optimality Theory*, Univ. of Massachusetts Occasional Papers **18**: 249–384.

Mohanan, K.P. 1993. Fields of attraction in phonology. In John Goldsmith (ed.) *The Last Phonological Rule: Reflections on Constraints and Derivations*. Chicago University Press, Chicago.

Ní Chiosáin, Máire & Jaye Padgett. 1997. Markedness, segment realization, and locality in spreading. Report LRC-97-01, Linguistic Research Center, UCSC.

Noonan, Michael. 1992. *A Grammar of Lango*. Mouton de Gruyter, Berlin.

Orie, Ọlanikẹ Ọla. 2001. An alignment-based account of vowel harmony in Ifẹ Yoruba. *Journal of African Languages and Linguistics* **22**: 117-143.

Orie, Ọlanikẹ Ọla. To appear. Two harmony theories and high vowel patterns in four Benue-Congo harmony systems. *The Linguistic Review*.

Prince, Alan & Paul Smolensky. 1993. Optimality Theory: constraint interaction and satisfaction. Ms. Rutgers University & University of Colorado at Boulder.

Pulleyblank, Douglas. 1996. Neutral vowels in Optimality Theory: a comparison of Yoruba and Wolof. *Canadian Journal of Linguistics* **41**: 295-347.

Pulleyblank, Douglas, Ping Jiang-King, Myles Leitch, & Ọlanikẹ Ọla. 1995. Typological variation through constraint rankings: low vowels in tongue root harmony. *Proceedings of South Western Optimality Theory Workshop*. The University of Arizona Coyote Papers, 184-208.

Sagey, Elizabeth. 1986. *The Representation of Features and Relations in Nonlinear Phonology*. Ph.D. diss., MIT. [Published 1990, Garland, New York]

Smolensky, Paul. 1993. Harmony, markedness, and phonological activity. Paper presented at the Rutgers Optimality Workshop #1, Rutgers University.

Suzuki, Keiichiro. 1998. *A Typological Investigation of Dissimilation*. Ph.D. dissertation, University of Arizona.

Vago, Robert. 1988. Underspecification in the height harmony system of Pasiego. *Phonology* **5**: 343-362.

Walker, Rachel. 1998. *Nasalization, Neutral Segments, and Opacity Effects*. Ph.D. dissertation, University of California, Santa Cruz.

Woock, E.B. & Michael Noonan. 1979. Vowel harmony in Lango. *Chicago Linguistic Society* **15**: 20-29.

Department of Linguistics
University of British Columbia
Vancouver, BC, Canada V6T 1Z1

Douglas.Pulleyblank@UBC.ca

Change of State Verbs: Implications for Theories of Argument Projection[1]

MALKA RAPPAPORT HOVAV
The Hebrew University of Jerusalem

BETH LEVIN
Stanford University

0. Introduction

Recent work in argument expression has focused on verbs showing multiple argument projection options, often with concomitant shifts in aspectual classification or assignment of so-called "aspectual roles" (e.g., measure or incremental theme). Two examples of these phenomena are given in (1)-(2).

(1) a. Dana read the book. (telic)

b. Dana read from the book. (atelic)

(2) a. Kerry wiped the table clean. (*the table* is the measure)

b. Kerry wiped the crumbs off the table. (*the crumbs* is the measure)

The ubiquity of such verbs has prompted theories of argument projection which adopt one or both of the following hypotheses:

(3) a. HYPOTHESIS I: Argument projection is aspectually determined.

b. HYPOTHESIS II: Argument expression is not lexically determined.

The first hypothesis, that argument projection is aspectually driven, finds perhaps its earliest explicit statement as Tenny's (1987, 1992, 2994) Aspectual Interface Hypothesis and has subsequently been quite widely adopted (e.g., Arad 1998, Borer 1998). Proponents of this hypothesis often establish a connection between direct objecthood and notions such as telicity (van Hout 1996, Ritter & Rosen 1998), incremental theme (Rothstein 2000), measure (Tenny 1994), or subject of result (Borer 1998). This hypothesis is tied explicitly to the phenomena in (1) and (2) by van Hout's 1996 proposal that argument alternations represent

[1] We thank audiences at BLS and the Syntax of Aspect Workshop at Ben Gurion University of the Negev for their questions and comments. This work was supported in part by Israel Science Foundation Grant 832-00 to Rappaport Hovav.

event type-shifting. Indeed, many alternations can be so understood. Arguments that are alternately expresses as direct object and oblique, e.g., (1), reflect alternations between telic and atelic uses of verbs, while alternate choices of direct object, e.g., (2), are a reflection of alternate choices of the argument which determines the telicity of the sentence.

The pervasiveness of multiple argument realization brought the assumption that verbs lexically determine the expression of their arguments under renewed scrutiny. Various researchers, starting with Hoekstra & Mulder (1990), proposed an alternative hypothesis: argument expression is not lexically determined (our Hypothesis II). Proponents of extreme versions of this hypothesis argue that arguments are projected freely onto syntax, with verbs being unspecified for those components of meaning that determine argument expression. The interpretation of a sentence is derived from the meaning of the verb in combination with the way in which its arguments are projected. Although lip service is often paid to the idea that a verb's meaning must be compatible with syntactically determine meaning (Ghomeshi & Massam 1995), it is the free projection of arguments which is stressed and put to work, while the explication of compatibility is taken to be trivial.

The ideas embodied in Hypotheses I and II go together naturally. In fact, many current theories of argument realization assume a conjunction of the two: aspectual properties are compositionally derived in syntax (Arad 1998, Borer 1998, Ritter & Rosen 1998). Verbs project their arguments freely onto syntax; the aspectual roles of arguments and the aspectual interpretation of the sentence are determined by the nature of this projection: aspectual composition is EFFECTED by checking or interpreting aspectual features in functional projections. This represents a departure from traditional theories of aspectual composition, which assume argument expression merely reflects aspectual composition.

Although much recent work incorporates the conjunction of the two hypotheses, they represent two distinct issues: whether argument expression is aspectually driven and whether argument expression is lexically or syntactically determined. It is possible to argue that argument projection is lexically determined and aspectually driven (e.g., Tenny 1987, 1992, 1994) or to argue that projection is not completely lexically determined but not completely aspectually driven either (e.g., Jackenoff 1990). We argue against each individual hypothesis, as well as against their conjunction. We do this through a close examination of the argument expression properties of change of state (COS) verbs—verbs lexicalizing a change of state—and a comparison of these properties with those of aspectually-related verbs. We show that although argument expression is not entirely lexically determined, as stressed in Butt & Geuder (1998b), a verb's lexicalized meaning is nonetheless important to determining or constraining its argument expression options (see also Erteschik-Shir & Rapoport 2000). Furthermore, the relevant facets of meaning don't correspond to well-known aspectual notions. We begin by showing that COS verbs, though uniform in argument expression, aren't uniform aspectually.

1. Uniformity in argument expression is not aspectual uniformity

COS verbs have long been known to exhibit distinctive argument realization properties (Fillmore 1970, 1977, Levin 1993). What is most striking are the severe constraints on their argument realization options. In particular, the patient argument — the entity undergoing the change of state — must be expressed and can only be expressed as a direct object, as we now illustrate.

Although other verbs are found in any of a number of frames with an argument left unexpressed, COS verbs are never found in such frames without their patient. Specifically, they aren't found with unspecified objects, as in (4), nor are they found in nonsubcategorized NP resultatives, as in (5), nor do they allow *out*-prefixation, as in (6). These last two frames resemble the unspecified object frame in that the verb's normal direct object is left unexpressed.[2]

(4) ∗Pat broke/dimmed.

(5) a. ∗My kids broke me into the poorhouse.

b. ∗The stagehand dimmed the scene dark.

(6) a. ∗The two-year old outbroke the three-year old.

b. ∗The stagehand outdimmed the director.

Furthermore, the patient must be the direct object and cannot be an oblique, as in (7). Consequently, COS verbs aren't found in object alternations in which the argument which is normally the direct object "vacates" its position for another NP, being expressed instead as an oblique, as in (8).

(7) a. Alex broke the vase/*Alex broke at the vase.

b. Sam dimmed the lights/*Sam dimmed at/from the lights.

(8) a. Kelly broke my arm

b. ∗Kelly broke me on the arm. (cf. Kelly hit me on the arm.)

The lack of argument alternation also emerges when the interpretation of the sentence pair in (9) with *break* is compared to that of the superficially parallel sentence pair with the non-COS verb *hit* in (10). As Fillmore (1977) points out, the *hit* sentences, as near paraphrases, qualify as an argument alternation. The *break* sentences, however, are not near paraphrases; rather, in each the direct object is understood as the patient.

[2] Goldberg (2001) points out that COS verbs are sometimes found with unspecified objects or in resultatives with nonsubcategorized NPs. However, as Goldberg herself notes, this happens with COS verbs only in generic or habitual contexts, while other verbs appear in these constructions even outside of these contexts. Thus, COS verbs are special, though such data must be accommodated within a full theory of argument realization

(9) a. Sam broke the fence with the stick.
b. Sam broke the stick against the fence.

(10) a. Sam hit the fence with a stick.
b. Sam hit a stick against the fence. (Fillmore 1977:75)

These differences are another manifestation of the constraint that the patient of a COS verb must be its direct object. These examples are noteworthy in another way. Although *the stick* is associated with an entailment of change of location which allows an argument to qualify for direct objecthood (cf. the acceptability of (10b)), as an argument of *break*, it can only be the direct object if it is also the patient.

This restricted behavior is unexpected from the perspective of Hypothesis II, which is often understood to mean that arguments project freely. Nonetheless, if argument expression is taken to be aspectually determined, the uniformity in argument expression of COS verbs might be attributed to a shared aspectual property. However, COS verbs lack a uniform aspectual characterization, at least in terms of traditional notions. When COS verbs take a definite, singular object, they can be necessarily telic (e.g., *break, dry, explode, flatten, freeze*) or either telic or atelic (e.g., *cool, darken, dim, widen*). Variable telicity, in fact, is the distinguishing property of the much-discussed set of COS verbs known as "degree achievements" (Abusch 1986, Dowty 1979, Hay, Kennedy & Levin 1999). Furthermore, when telic, some COS verbs are punctual (e.g., *break, crack, explode*), while others are durative (e.g., *cool, dim, dry, freeze, widen*). Despite these differences in aspectual potential, all COS verbs show the same behavior. We illustrated the properties of COS verbs using the verbs *break* and *dim*, which were chosen because they differ along aspectual dimensions. First, *break* is necessarily telic, while *dim* — a degree achievement — may be telic or atelic. Second, *break* is punctual and *dim* is durative. Yet both verbs show the same argument realization patterns.

COS verbs, then, share a constrained set of argument projection possibilities, but aren't uniform aspectually. These observations suggest that lexical aspectual classification alone does not determine argument expression. Hypothesis I is undermined, unless some other aspectual property can be shown to unify the class of COS verbs. We now address this issue.

2. Probing the contribution of aspect to argument expression further

2.1 Verbs with incremental themes show different argument projections

Much current work suggests that the aspectual notion most relevant to argument projection is "incremental theme" (Dowty 1991) or one of its relatives. Krifka (1992) suggests that incremental theme verbs have an argument that is lexically associated with the property of "mapping to events", that is, parts of the entity denoted by that argument can be mapped onto parts of the event denoted by the verb. For example, when you drink a glass of water, the event is half over when half the water has been consumed. Therefore, Hypothesis I might still be on the right track, with the aspectual notion of incremental theme unifying the verbs showing the pattern of behavior demonstrated by COS verbs. However, as we now show, though the patient of a COS verb acts as an incremental theme, a comparison of COS verbs with other incremental theme verbs suggests that it is not the patient as incremental theme which determines the argument expression profile of COS verbs.

First, we comment on terminology. The term "incremental theme" was originally applied by Dowty (1991) to the argument of certain predicates involved in defining a homomorphism from its own physical extent to the temporal progress of the event it participates in. By this definition, verbs like *read, write*, and *eat* are incremental theme verbs. Dowty also meant this term to cover the patient argument of COS verbs, but its application to these verbs needs clarification. The sentence *Matt closed the door halfway* doesn't entail that half the door was closed, but that the door was halfway closed. The mapping involves a property of the door and not the door's own physical extent. Recent studies (Hay et al. 1999, Krifka 1998, Ramchand 1997, Tenny 1992, 1994) have found ways to provide parallel aspectual analyses to COS verbs and traditional incremental theme verbs. Patients of COS verbs and traditional incremental themes are associated with some property — a scalar property of the object lexicalized by their verb for the former and the physical extent of the object for the latter — which serves as a scale for measuring the temporal progress of the entire event. When the event describes a specified degree of change on the scale, it is telic, and when it describes an unspecified degree of change, it is atelic (Kennedy & Levin 2001). The objects of both traditional incremental theme verbs and COS verbs, then, share an identical aspectual role, which we continue to call "incremental theme," in that both verb types are associated with a scale for measuring the event's progress.

The parallelism can be brought out further. It is well-known that the quantization of the direct object of a traditional incremental theme verb determines the telicity of its sentence, as in (11). The physical extent of the object provides the scale for measuring the progress of such events. If the object is quantized, its physical extent is specified, as is the change on the associated scale, and the sentence is telic. If the object isn't quantized, the scale lacks a specified endpoint, the change on this scale is unspecified, and the sentence is atelic.

(11) a. Dana read poetry for/*in an hour. (nonquantized object; atelic)

b. Dana read the newspaper for/in an hour. (quantized object; telic)

The telicity of a sentence with a COS verb also depends on the nature of the change on the associated scale. The relevant change, however, is determined by a scalar property lexicalized in the verb, not directly by the quantized nature of the direct object. The verb *warm* is associated with a temperature scale, and a sentence with this verb is telic if the change in temperature is specified and atelic otherwise, as in (12); see Kennedy & Levin (2001) for more discussion.

(12) a. Sandy warmed the solution for three minutes. (atelic)

b. Sandy warmed the solution five degrees in three minutes. (telic)

Although parallel aspectual analyses are available for both verb types, traditional incremental theme verbs are more flexible in their argument expression properties than COS verbs. First, the argument that serves as the incremental theme when these verbs are used transitively need not be expressed. These verbs permit unspecified objects, as in (13); they also allow nonsubcategorized NP objects, either in a resultative construction or via *out*-prefixation, as in (14) and (15) (cf. (4)-(6)). Furthermore, this same argument need not be expressed as direct object, as in (16), though it is then no longer an incremental theme (cf. (7)). Thus, verbs that have an incremental theme do not show uniform argument expression properties.

(13) Dana read/ate/wrote.

(14) a. The teacher read us into a stupor.

b. My kids ate me into the poorhouse.

c. I wrote myself out of a job.

(15) Pat outread/outate/outwrote Chris.

(16) a. Dana read the book./Dana read from the book.

b. Chris ate the apple./Chris ate from/of the apple.

c. I wrote my book./I wrote at my book.

2.2 Comparison with potential incremental theme verbs

Verbs like *read, write*, and *eat*, which invariably allow their direct object to be interpreted as an incremental theme, may be contrasted with potential — or "latent" (Tenny 1992:20) — incremental theme verbs, a class we exemplify with surface contact verbs (e.g., *comb, rub, scratch, shovel, sweep, wipe*). Such verbs do not require their "normal" direct object, even when quantized, to be analyzed as an incremental theme.[3] That is, in the presence of a quantized object, they may pattern as telic or atelic with respect to standard telicity tests, as in (17) and (18).

(17) a. Lee scrubbed the tub for hours. (atelic)

b. Lee scrubbed the tub in three minutes flat. (telic)

[3] In fact, traditional incremental theme verbs (e.g., *read, write, eat*) also show atelic readings with quantized objects (Hay et al. 1999, van Hout 1996, Tenny 1994). *Pat read the newspaper for an hour* is certainly marginally acceptable, and some speakers even find it fully acceptable. In contrast, speakers agree that a potential incremental theme verb with a quantized object is perfectly acceptable with an atelic interpretation.

(18) a. Lee is scrubbing the tub and has scrubbed it for the last hour. (atelic)

b. Lee is scrubbing the tub and still hasn't finished. (telic)

These verbs may be atelic because they describe processes that can be applied indefinitely to a surface. Their telic uses most likely arise because the processes they describe are usually carried out with specific intended results, though these verbs do not entail the achievement of any result (Talmy 2000). The intended result gives rise to an associated scale. With *scrub* two scales are possible. One is provided by the tub's surface area, with the process being complete when the scrubbing has covered the entire tub. Alternatively, the desired result may be a clean tub, with the scale being one of cleanliness. On either interpretation, the object is an incremental theme: on the former, the event is over when the whole tub is scrubbed, and on the latter, the event is over when the tub's state reaches the point of cleanliness.

The "normal" direct object of surface contact verbs is considered a location in a traditional semantic role analysis, but these verbs may also take as their object an argument describing material found at this location (Levin & Rappaport Hovav 1991), giving rise to an argument alternation, as in (19). When the material is the object, it too can be the incremental theme by virtue of its physical extent: it determines telicity when quantized, as in (20).

(19) Lee scrubbed the tub./Lee scrubbed the stains off the tub.

(20) a. Lee scrubbed blood off the tub for ten minutes. (atelic)

b. Lee scrubbed the blood off the tub in ten minutes. (telic)

Surface contact verbs need not express either the material or the location. They don't require the expression of an object, as in (21), and can take nonsubcategorized NP objects, either in resultative constructions or via *out*-prefixation, as in (22)-(23).

(21) Lee swept/wiped/scrubbed.

(22) Cinderella scrubbed her fingers to the bone.

(23) a. Cinderella outswept/outscrubbed her stepsisters.

b. This hairdresser outcombed that one.

Finally, potential incremental theme verbs don't impose semantic restrictions on their direct object: it can be the material, as in *Lee scrubbed the stains*, the location as in *Lee scrubbed the tub*, or a nonargument, as in (22)-(23). The material and location arguments, though potential incremental themes, can be expressed as obliques as well as objects, though the oblique expression of locations is dispreferred.

(24) a. Kerry scrubbed at the stain. (material)

b. ? Kerry scrubbed at the counter. (location)

(25) a. Lee scratched at the mosquito bites. (material)

b. ? Lee scratched at her arm. (location)

In conclusion, COS verbs share an important aspectual property — an incremental theme — with other verbs, but don't share their argument expression properties.

2.3 Lexical influences on the choice of incremental theme

Traditional and potential incremental theme verbs have been used to support the conjunction of Hypotheses I and II: verbs are not lexically specified for an incremental theme and an argument or other NP projected onto the direct object position (or whatever its formal definition is taken to be, e.g., Spec, AgrO) is interpreted as the incremental theme. Indeed, the argument or other NP chosen as the object of these verbs is construed (or construable) as their incremental theme.

Although this is true of some verbs, it isn't true of COS verbs. As discussed in section 2.1, the patient of a COS verb MUST be construed as the incremental theme of its sentence, and no other argument or NP may be so construed. COS verbs, at least, seem to have lexically specified incremental themes. Construal as an incremental theme, then, isn't always the result of the free projection of an argument onto a specific syntactic position.

If some verbs are lexically specified as taking a particular argument as incremental theme, while others are not, it may be possible to maintain a form of Hypothesis I, while abandoning Hypothesis II: with COS verbs, the entity associated with the change of state entailment is lexically constrained to be the incremental theme, with their limited argument projection options following from this assumption. This appears to be the gist of Dowty's (1991) suggested analysis of these verbs. The other classes of incremental theme verbs would not lexically specify a particular argument as incremental theme and, thus, would show more argument expression options. On this approach, a lexically specified aspectual property still determines argument expression.

This proposal, however, is predicated on two assumptions: incremental themes are constrained to be direct objects and in the presence of an incremental theme no other argument may be the direct object. However, neither assumption is correct. First, Dowty (1991:570) and Jackendoff (1996:313) have pointed out incremental themes which are not direct objects, as in *The train crossed the border* and *The parade passed the mayor*. Second, in some instances one argument may serve as an incremental theme even when another is expressed as direct object. This situation is found in the dative alternation, shown by verbs of transfer. With these verbs, the theme normally determines telicity, as in (26).

(26) a. Dana read poetry to employees/her niece for an hour. (atelic)

b. Dana read the story to employees/her niece in an hour. (telic)

Despite Arad's (1998) claim that the dative alternation is aspectually motivated, the theme is still the incremental theme when the recipient becomes the direct object in

References

Abusch, Dorit. 1986. Verbs of change, causation, and time. Report CSLI-86-50, Center for the Study of Language and Information, Stanford University.

Arad, Maya. 1998. VP-structure and the syntax-lexicon interface. PhD dissertation, University College London.

Baker, Mark C. 1997. Thematic roles and syntactic structure. In L. Haegeman (ed.) *Elements of Grammar*, 73-137. Dordrecht: Kluwer.

Borer, Hagit. 1998. Passive without theta grids. In S.G. Lapointe, et. al (eds.) *Morphological Interfaces*, 60-99. Stanford, CA: CSLI Publications.

Butt, Miriam & Willi Geuder (eds.). 1998a. *The projection of arguments: Lexical and compositional factors*. Stanford, CA: CSLI Publications.

Butt, Miriam & Willi Geuder. 1998b. Introduction. In Butt & Geuder (eds.), 1-20.

Dowty, David. 1979. *Word meaning and Montague Grammar*. Dordrecht: Reidel.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67: 547-619.

Erteschik-Shir, Nomi & T.R. Rapoport. 2000. Bare aspect: A theory of syntactic projection. International Round Table 'The Syntax of Tense and Aspect', Université de Paris 7.

Fillmore, Charles J. 1970. The grammar of *hitting* and *breaking*. In Roderick A. Jacobs & Peter S. Rosenbaum (eds.) *Readings in English Transformational Grammar*, 120-133. Waltham, MA: Ginn.

Fillmore, Charles J. 1977. The case for case reopened. In P. Cole & J.M. Sadock (eds.) *Grammatical relations*, 59-81. New York: Academic Press.

Ghomeshi, Jila & Diane Massam. 1995. Lexical/syntactic relations without projection. *Linguistic Analysis* 24: 175-217.

Goldberg, Adele E. 2001. Patient arguments of causative verbs can be omitted: The role of information structure in argument distribution. *Language sciences* 23: 503-524.

Hale, Ken & Jay Keyser. 1998. The basic elements of argument structure. *MIT working papers in linguistics* 32: 73-118.

Hay, Jennifer, Chris Kennedy & Beth Levin. 1999. Scalar structure underlies telicity in 'degree achievements'. *SALT* 9: 127-144.

Hoekstra, Teun & René Mulder. 1990. Unergatives as copular verbs: Locational and existential predication. *The Linguistic Review* 7: 1-79.

van Hout, Angeliek. 1996. Event semantics of verb frame alternations. PhD dissertation, Tilburg University.

Jackendoff, Ray S. 1976. Toward an explanatory semantic representation. *Linguistic inquiry* 7: 89-150.

Jackendoff, Ray S. 1983. *Semantics and cognition*. Cambridge, MA: MIT Press.

Jackendoff, Ray S. 1990. *Semantic structures*. Cambridge, MA: MIT Press.

Jackendoff, Ray S. 1996. The proper treatment of measuring out, telicity, and perhaps even quantification in English. *NLLT* 14: 305-354.

of state MUST be the direct object and MUST be the incremental theme; in contrast, as we now show, an entity which is lexically entailed to undergo movement CAN be a direct object, but need not be, and it need not be an incremental theme.

We make this point with the locative alternation verbs *splash* and *spray*, which lexically entail the movement of a liquid substance, i.e. they take an argument that qualifies as a theme of change of location. This argument need not be expressed, as in (28), nor does it have to be direct object, as in (29).

(28) Brett splashed/sprayed.

(29) a. Brett sprayed/splashed water on the plants.

b. Brett sprayed/splashed the plants with water.

c. ? Brett sprayed/splashed at the plants.

Although the patient of a COS verb MUST be that verb's incremental theme, the theme of a verb entailing change of location need not be that verb's incremental theme. As locative alternation verbs, *splash* and *spray*, allow either a location or a theme of change of location as their direct object. When the location is the object, these verbs pattern like latent incremental theme verbs since the telicity of their sentence isn't necessarily determined by the quantized nature of the object, but it may be (Dowty 1991, Jackendoff 1996).

(30) a. Bill sprayed the wall with paint for five minutes. (atelic)

b. Bill sprayed the wall with paint in an hour. (telic)

It is noteworthy that the location is a potential incremental theme, though these verbs also have an argument that is a theme of a change of location. COS verbs, in contrast, do not allow their patient argument to abdicate incremental themehood to another argument. The special properties of COS verbs, then, are characteristic of just this semantic class and do not generalize to other semantic classes of verbs, even one that has been given a parallel semantic analysis.

4. Conclusion

The argument expression possibilities of COS verbs appear to be determined by a nonaspectual, lexicalized property — change of state — and can't be handled by purely aspectual nonlexical theories of argument projection. Traditional aspectual classes don't constitute natural classes from the perspective of argument expression.

These data pose problems for approaches which take aspectual notions as the sole determinants of argument expression. They support a theory, as in Rappaport Hovav & Levin (1998), in which constants typed by ontological category are lexically associated with nonaspectually defined event structures, and these, in turn, constrain argument projection, as also espoused by Hale & Keyser (1998).

sions, it seems that the entailment of a change of state and the entailment of a change of location for an argument constrain argument projection in different ways.

the double object construction. In (27) the recipient is the first object — taken to be the double object construction's instantiation of a direct object — and the theme, though the second object, still determines telicity.

(27) a. Dana read her niece poetry for an hour. (atelic)

b. Dana read her niece the story in an hour. (telic)

Therefore, this alternation in direct object choice cannot be aspectually driven (see also Baker 1997), contra Hypothesis I, and a de facto incremental theme "relinquishes" its position as direct object to another argument. Recently, some researchers have proposed that the recipient in a double object construction is not an underlying direct object (the theme is), but rather has become the specifier of a higher functional projection (Baker 1997). Although this analysis is meant to explain why recipients do not show the full range of direct object properties (Baker 1997, Maling 2001), this analysis is incompatible with approaches to argument projection that adopt the conjunction of Hypotheses I and II. On these approaches, movement into a higher functional projection is supposed to be associated with aspectual shifts, while "underlying" direct objects (i.e., complements of V) are explicitly not associated with these properties. Yet, on Baker's analysis, the theme, which can act as an "incremental theme" is the "underlying" direct object, while the recipient moves to a higher functional projection.

Finally, there are some incremental themes which aren't explicitly expressed. For example, the direct object in *out*-prefixation sentences is not an incremental theme. In *Pat outate Chris*, parts of Chris do not correspond to parts of the outeating event. Rather, the incremental theme seems to be the amount of eating that Pat did (e.g., *Pat was halfway towards outeating Chris*), and this notion is not expressed.

Since the incremental theme of a sentence need not be its direct object, the fact that the patient of a COS verb must be expressed as direct object can't be attributed to its having a lexically specified incremental theme. The distinctive argument expression properties of COS verbs appear NOT to follow from any aspectual property of these verbs, whether lexically specified or not.

3. Comparison with verbs having a change of location entailment

Most researchers draw a parallel between themes of changes of location and patients of changes of state. These parallels were first drawn in localist theories (Jackendoff 1976, 1983), which conceptualize changes of state as instances of changes of location.[4] As discussed in section 2.1, an entity lexically entailed to undergo a change

[4] Jackendoff (1990) proposes that changes of state and changes of location have different predicates in their conceptual structure. He thereby moves away from the strict versions of the Localist Hypothesis adopted in his earlier work (1976, 1983). More recently, this parallel has been used to unify these arguments under the same aspectual concept (Hay et al. 1999, Levin & Rappaport Hovav 1995, Ramchand 1997, Tenny 1992, 1994). Although drawing the parallel may be useful for determining aspectual properties of verbs and for understanding certain kinds of metaphorical meaning exten-

Kennedy, Chris & Beth Levin. 2001. Telicity corresponds to degree of change. Handout, 75th Annual LSA Meeting, Washington, DC.

Krifka, Manfred. 1992. Thematic relations as links between nominal reference and temporal constitution. In Sag & Szabolcsi (eds.), 29-54.

Levin, Beth. 1993. *English verb classes and alternations.* Chicago, IL: University of Chicago Press.

Levin, Beth. 1999. Objecthood: An event structure perspective. *CLS 35; Part 1*: 223-247.

Levin, Beth and Malka Rappaport Hovav. 1991. Wiping the slate clean: A lexical semantic exploration. *Cognition* 41: 123-151.

Levin, Beth and Malka Rappaport Hovav. 1995. *Unaccusativity: At the Syntax-Lexical Semantics Interface.* Cambridge, MA: MIT Press.

Maling, Joan. 2001. Dative: The heterogeneity of the mapping among morphological case, grammatical functions, and thematic roles. *Lingua* 111: 419-464.

Ramchand, Gillian C. 1997. *Aspect and predication.* Oxford: Clarendon Press.

Rappaport Hovav, Malka & Beth Levin. 1998. Building verb meanings. In Butt & Geuder (eds.), 97-134.

Ritter, Elizabeth & Sara T. Rosen. 1998. Delimiting events in syntax. In Butt & Geuder (eds), 135-164.

Rothstein, Susan. 2000. Secondary predication and aspectual structure. *Approaching the grammar of adjuncts.* ZAS Papers in Linguistics **17**: 241-264.

Sag, Ivan and Anna Szabolcsi (eds.). 1992. Lexical matters. Stanford, CA: CSLI.

Talmy, L. 2000. *Towards a cognitive semantics II.* Cambridge, MA: MIT Press.

Tenny, Carol L. 1987. Grammaticalizing aspect and affectedness. PhD dissertation, MIT.

Tenny, Carol L. 1992. The Aspectual Interface Hypothesis. In Sag & Szabolcsi (eds), 1-27.

Tenny, Carol L. 1994. *Aspectual roles and the syntax-semantics interface.* Dordrecht: Kluwer.

(Rappaport Hovav)
Department of English
The Hebrew University of Jerusalem
91905 Jerusalem, Israel

(Levin)
Department of Linguistics
Stanford University
Stanford, CA 94305-2150

mhovav@mscc.huji.ac.il
bclevin@stanford.edu

Language Typology and Tonogenesis in Two Atlantic Creoles

YOLANDA RIVERA-CASTILLO
University of Puerto Rico-Río Piedras

0. Introduction

Languages belong to three basic types when analyzed according to the assignment of tonal features within phonological units (Hyman 1978, and 1992): intonational (phonological phrases), pitch-accent (fixed position in phonological words), and tonal languages (feet and syllables). Regarding the lexically distinctive nature of tone in languages, Hyman (1992:166) states that

> [..] a tone language is a language in which both pitch phonemes and segmental phonemes enter into the composition of at least some morphemes. Tone assignment happens at the lexical or postlexical levels. This distinguishes tonal languages from pitch-accent languages and intonational languages.

Intonational languages have no lexical tone assignment. Tones in pitch-accent languages are introduced post-lexically since they are not necessarily associated with particular morphemes (Pulleyblank 1986:20).[1]

Papiamentu and Saramaccan, two Atlantic Creoles, have tonal systems. Both have a bitonal system with the mora as the tone bearing unit (TBU), and tonal patterns that distinguish lexical categories. These exhibit downstepping and downdrift; features that are shared with West African languages and are absent, for example, in East Asian tonal languages (Yip 1995). Papiamentu and Saramaccan have partially restricted tonal systems, such as those of some Bantu languages (Voorhoeve 1968).

This paper provides evidence of strong typological similarities between the tonal systems of Papiamentu and Saramaccan with the systems of West African languages. These typological similarities constitute the basis for a proposal that there is a genetic affiliation between Papiamentu and Saramaccan with the Kwa and Bantu language families; an affiliation that reaches beyond the accidental

[1] In fact, the distinction between tone languages and pitch-accent languages has been interpreted in many studies as a difference between a restricted and an unrestricted distribution of tone (Hyman 1978, 1992; McCawley, 1978). However, tone languages have different degrees of restrictions regarding tone distribution.

lexical borrowing. Since Saramaccan has been classified as an English-based Creole, and Papiamentu as a Romance-based Creole, their similarities indicate that their substrata have a greater significance in Creole genesis than previously recognized.

Given that there is a strong correlation between typological and genetic relations (Greenberg 1974), Indo-European languages had no influence in Atlantic Creole tonogenesis. Changes in suprasegmental systems typically involve changes from tone --> pitch-accent --> stress-accent (Salmons 1992:272). There are shifts from tonal to stress systems, but no shift from stress to tonal systems (Salmons 1992). Therefore, both Saramaccan and Papiamentu's systems could not have emerged from the systems of their non-tonal lexifiers.

Papiamentu and Saramaccan's tonal systems emerged from their substrata: West African languages at a particular stage of development during the slave trade. Tonal features cannot be borrowed but integrated into the complex, coherent, and self-contained systems of these Creoles. Therefore, this paper explores only general typological features. Similarities regarding specifics of tonal behavior require further study; one that considers the effect of language change in the tonal systems of these Creoles. Also, I assume that tonal features which distinguish West African languages from other language families constitute reliable evidence of genetic affiliation *vis a vis* features identified as language universals that these Creoles could share with any tonal language.

In the following section (1), this paper presents evidence of similarities between Papiamentu, Saramaccan and West African languages. The parameters discussed include tone levels, tone bearing units, tone patterns associated with grammatical categories, tone spreading, and downstepping. It discusses phenomena that specifically distinguish restricted from non-restricted tone languages such as the type of tone bearing units (syllable or word/foot); fixed tonal patterns versus free tone assignment to grammatical categories; and tone spreading versus tone polarization or anticipation. Finally, Section 2 discusses the similarities and differences between these Creoles.Voorhoeve (1959 and 1961), Rountree (1972), Byrne (1987), and Ham (1999) provided some of our examples and descriptions of Saramaccan. Most of the Papiamentu examples were originally provided by Raúl Römer (1983 and 1991), Harris (1951), Bendix (1983), Rivera (1998), and Pickering and Rivera (2001). Other observations are based on data provided by these authors and on independent research.

1. Typological Features of Papiamentu and Saramaccan's Tone Systems

Papiamentu has lexically predetermined tones but also postlexical tone assignment. As a mixed system—a tone + stress language—Papiamentu combines features of tonal languages with those of stress languages (Rivera 1998). At the lexical level, it constitutes a bitonal system like that of many West African languages. Asian languages usually have a greater number of tones (up to five in some languages).

The Tone Bearing Unit (TBU) is the mora. Tonal systems of West African languages require that every unit carry a tone. For example, Papiamentu has contour tones only in long bimoraic stressed syllables (´ - High tone; ' - Stress; and Low tone unmarked) (Birmingham 1971:5):

(1) 'dúùna 'to give'

In Asian languages, contour tones can attach to syllables regardless of the number of morae. Contour tones are more numerous in Asian languages and level tones have a more important role in lexical distinctions for African languages (Chen 1992:61). For all practical purposes, morpheme and syllable are co-extensive in Asian languages (Chen 1992:60); and contour tones constitute an inseparable unit linked to a syllable.

On the other hand, Papiamentu has lexically predetermined tone patterns that respond to categorial distinctions (Kouwenberg and Murray 1994). For example, verbs carry tone patterns that distinguish these from nouns:

(2) 'biáha	'trip'	versus	'biahá 'to travel'
H-L			L-H

Tone in African languages plays an important role in distinguishing grammatical categories (Schuh 1978:251-254). Kwa languages such as Yoruba make extensive use of tone for lexical distinctions (Pulleyblank 1992:263). In Common Bantu there was mostly free tone assignment (Phillipson 1998:316). However, modern Bantu languages like Kimatuumbi and Kikuria exhibit different tonal patterns in the verb stem that respond to differences in tense-mood, and aspect in the verb (Odden 1989 and 1995:449). Although similar patterns are part of the tonal systems of some Asian languages, such as Tokyo Japanese (McCawley 1978:528), most do not rely on these for lexical distinctions, and have lexical free tone assignment.

Postlexically, Papiamentu assigns tonal alternations (polarization) to syllabic sequences within phrasal domains. Alternations result from the application of the Obligatory Contour Principle (OCP) in this tone-to-stress language. Salmons (1992:274) suggests that the application of the OCP operates differently in stress and in tonal languages. In tonal languages, it requires spreading while in stress languages it requires destressing or the creation of non-identical adjacent levels of stress. Indeed, alternating HLHL sequences are characteristic of tone-to stress languages, tonal languages shifting to a stress system. Some African languages, such as Ganda and Hausa, have systems similar to Papiamentu's (Hyman and Katamba 1993, Newman 1995). Nevertheless, tonal alternations are not the only tone shifting mechanism operating in Papiamentu, spreading is also a significant component in its system:

(3) Polarization:	[ún	sáku	blánku]$_{NP}$	→ é sáku blánku
	a	bag	white	

(4) Spreading: é sáku [á skér]$_{VP}$ → é sákú á skér

the bag PAST tear

Spreading applies only before phrase edges (4); and polarization applies freely in all other contexts.

Chen (1992:54) identifies three parameters that distinguish African from Asian tone languages: tone spreading or movement, metathesis or the melodic inversion of contour tones, and downstep and downdrift. Bao (1992:3) indicates that contour (Asian) as well as level tone (African) languages have tone spreading; however, spreading applies in a restricted fashion in Asian languages since, with few exceptions, every morpheme carries a lexically prespecified tone (Chen 1992:60). Spreading is not restricted in African languages and these even have long distance spreading (Cassimjee and Kisseberth 1992:26). As indicated, there are other postlexical phenomena in Saramaccan and Papiamentu such as downstepping and downdrift, which are characteristic of African languages but are practically absent from Asian languages.

Downstepping and downdrift constitute a gradual lowering of tones adjacent to a L tone. Pickering and Rivera (2001) have found evidence of downdrift in Papiamentu:

			225hz	193.42	200.45
(5) mi	tá	du'ná-	bó	a-	r!óz
I	PRES	give	you	rice	

In (5), the pitch readings indicate the effect of a L tone in the following H tone (! - lowering). Harris (1951), Römer (1991), and Bendix (1983) have also described downdrift effects in Papiamentu. These features show a clear typological connection between Papiamentu and Kwa and Bantu languages.

Saramaccan shares many of these features with Papiamentu. Saramaccan has a bitonal system and its TBU is the mora. Rountree (1972) indicates the presence of mid tones in this language (a tritonal system), but it is unclear at this point whether these are lexically distinctive. However, Ewe, the most likely substratum for Saramaccan (Ham 1999), also has a tritonal system. This is still short of four or five level tones typically found in Asian languages.

Saramaccan also has tone patterns associated with lexical categories, including patterns that distinguish verbs from other categories (Voorhoeve 1959, Byrne 1987:261):

(6) maaká 'to notice'	versus	maáka 'portent'	
L-L-H		L-H-L	

In these cases, each mora (vowel) bears a different tone since the mora is the TBU.

On the other hand, postlexically, Saramaccan has tone spreading and no polarization. Tone spreading in Saramaccan results from postlexical rules (Ham

1999). For example, a High (H) tone in a word can spread over to following words (E - lax mid front vowel):

(7) dí hánso mu̲jÉE → dí hánsó mú̲jÉE
the beautiful woman

In Saramaccan, the H tone of the determiner *dí* spreads to the first syllable of the word *mújÉE.* It has unbounded spreading, including any number of syllables between two H tones with no intervening phrasal edges. Other postlexical phenomena includes downstepping.

Rountree (1972:312) has described cases of downstep in Saramaccan triggered by a following L tone:

(8) álá njan´j!á̲n dE
there food is-V

The H tone in *ján* lowers because the following verb carries L tone.

There are five important parameters in the typological classification of these Creole systems: the number of tone levels, the TBU, the assignment of patterns to lexical categories, tone spreading, and downstepping. These establish fundamental distinctions between West African and Asian tonal languages. TBUs, the assignment of patterns to lexical categories, and tone spreading also distinguish restricted from non-restricted tone languages that identify stages in the shift from tone to stress systems, being the non-restricted type the one closer to a pure tonal system. Indeed, Papiamentu exhibits more features of restricted systems than Saramaccan. The following section describes the main differences between Papiamentu and Saramaccan in more detail.

2. Systems with one Lexically Prespecified Tone:

Non-restricted tone languages must have constituted Papiamentu's substrata; and Saramaccan's substrata must have been stable non-restricted tone languages. Indeed, most slaves in Curaçao and those from which the Saramacca descended were Ewe (Kwa) speakers, currently a stable non-restricted tone language (Singler 1996, Ham 1999). Papiamentu exhibits some features of restricted tone languages, while Saramaccan is essentially a non-restricted tone language. However, as tone languages, both share more parametric similarities than differences.

Saramaccan and Papiamentu share the following features:

(A) Tone is distinctive at the lexical level. There are two level tones: H and L.
(B) The mora is the TBU. Contour tones occur only in bimoraic sequences.
(C) Tonal patterns distinguish lexical categories.
(D) Phonological Words must have, at least, a syllable with H tone. However, these can have more than one H per word, resulting either from postlexical rules or lexical prespecification.

(C) Spreading and Polarization are strictly local (no gaps).
- — Spreading to morae unspecified for tone immediately following H tone.
- — Polarization: Morae unspecified for tone receives a polarizing tone with respect to following tone.

(D) There are floating tones, tone conservation, and downstepping/downdrift.

What makes Papiamentu closer to a restricted tone language and different from Saramaccan includes the following:

(A) A different set of lexical categories carry tone patterns in each language, being Saramaccan a language with freer tone assignment and a smaller number of categories subject to tone patterns.
(B) Polarization, a feature of tone-to-stress systems, applies in Papiamentu, not Saramaccan.
(C) Accent has an important role in Papiamentu, with stress at the lexical level while Saramaccan has only phrase level stress assignment.
(D) Saramaccan has lexically assigned falling and rising contour tones, while Papiamentu has falling contour tones but rising contour tones resulting from postlexical rules only. Papiamentu has a smaller set of lexically assigned contour and level tones.

Both are tone languages even if Papiamentu exhibits more changes. Tone spreading, free tone assignment, and other typically tonal phenomena indicate the operation of a tonal, not a stress system (Hyman 1978). Differences found in Papiamentu indicate a shift towards an accentual system: (a) stress and tonal systems in which stress and H tone are attracted to accented monosyllables; (b) polarization; and (c) fixed tonal patterns with a single H tone dominate as lexical properties.

3. Conclusions

Typologically, Saramaccan and Papiamentu are not only tonal languages, but tonal languages related to each other by their typological affiliation to West African languages. This proposition is strong evidence for a genetic affiliation between these languages which lexifiers belong to different branches of the Indo-European family. Both are descendants of the same parents, probably a Kwa language like Ewe, or closely related languages in West Africa.

References

Bao, Zhiming. 1992. Towards a Typology of Tone Sandhi. In L. A. Buszard-Welcher, J. Evans, D. Peterson, L. Wee, and W. Weigel (eds.) *Proceedings of the 18th Annual Meeting of the Berkeley Linguistics Society - Special Session on the Typology of Tone Languages*. Berkeley: University of California Press.

Bendix, Edward H. 1983. Sandhi Phenomena in Papiamentu, other Creoles, and African Languages. In Lawrence D. Carrington (ed.) *Studies in Caribbean Language*. St. Augustine, Trinidad: Society for Caribbean Linguistics.

Byrne, Francis. 1987. *Grammatical Relations in a Radical Creole: Verb Complementation in Saramaccan.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Cassimjee, Farida & Charles W. Kisseberth. 1992. Tonology of Depressor Consonants: Evidence from Mijikena and Nguni. In L. A. Buszard-Welcher, J. Evans, D. Peterson, L. Wee, and W. Weigel (eds.) *Proceedings of teh 18th Annual Meeting of the Berkeley Linguistics Society - Special Session on the Typology of Tone Languages*. Berkeley: University of California Press.

Chen, Matthew Y. 1992. Tone Rule Typology. In L. A. Buszard-Welcher, J. Evans, D. Peterson, L. Wee, and W. Weigel (eds.) *Proceedings of teh 18th Annual Meeting of the Berkeley Linguistics Society - Special Session on the Typology of Tone Languages*. Berkeley: University of California Press.

Greenberg, J. 1974. *Language Typology: A historical and analytic overview.* The Hague/Paris: Mouton.

Ham, William. 1999. Tone Sandhi in Saramaccan: A case of substrate transfer? *Journal of Pidgin and Creole Languages* **14**: 45-91.

Harris, C. C. 1951. *Papiamentu Phonology.* Unpublished doctoral dissertation, Cornell University, New York.

Hyman, L. M. 1978. Tone and/or accent. In D.J. Napoli (ed.), *Elements of tone, stress, and intonation.* Washington, D.C.: Georgetown University Press.

Hyman, L. M. 1992. Tone in phonology. In W. Bright (ed.) *International Encyclopedia of Linguistics Vol. 4.* New York: Oxford University Press.

Hyman, Larry M. & Francis X. Katamba. 1993. A New Approach to Tone in Luganda. *Language* **69.1**: 34-67.

Kouwenberg, Silvia & Eric Murray. 1994. *Papiamentu.* Languages of the World/materials 83. München/Newcastle: Lincom Europa.

McCawley, James D. 1978. What is a Tone Language? In Victoria A. Fromkin (ed.) *Tone: A Linguistic Survey.* New York: Academic Press.

Newman, Paul. 1995. Hausa Tonology: Complexities in an "Easy" Tone Language. In J. A. Goldsmith (ed.) *The Handbook of Phonological Theory.* Cambridge, Massachusetts: Blackwell.

Odden, D. 1989. Predictable Tone Systems in Bantu. In H. van der Hulst and N. Smith (eds.) *Autosegmental Studies on Pitch Accent Systems*. Dordrecht: Foris.

Odden, D. 1995. Tone: African languages. In J. A. Goldsmith (ed.) *The Handbook of Phonological Theory*. Cambridge, Massachusetts: Blackwell.

Phillipson, Gérard. 1998. Tone Reduction vs. Metrical Attraction in the Evolution of Eastern Bantu Tone Systems. In L. M. Hyman and C. W. Kisseberth (eds.) *Theoretical Aspects of Bantu Tone*. Stanford, CA: CSLI Publications.

Pickering, Lucy & Yolanda Rivera. 2001. Phonetic Correlates of Stress and Tone in a Mixed System. Ms. University of Puerto Rico-Río Piedras/University of Alabama-Tuscaloosa.

Pulleyblank, D. 1986. *Tone in lexical phonology*. Dordrecht, Holland: D. Reidel Publishing Company.

Rivera, Y. 1998. Tone and Stress in Papiamentu: the Contribution of a Constraint-Based Analysis to the Problem of Creole Genesis. *Journal of Pidgin and Creole Languages* **13.2**: 1-38.

Römer, R. G. 1983. Papiamentu tones. In E. Muller (ed.) *Papiamentu: Problems and possibilities, papers presented at the Conference on Papiamentu on the occasion of the 65th anniversary of Madero and Curiels Bank*. Zutphen: De Walburg Pers.

Römer, R.G. 1991. *Studies in Papiamentu Phonology (Caribbean Culture Studies)*. Amsterdam and Kingston: The University of Amsterdam.

Rountree, Catherine S. 1972. Saramaccan Tone in Relation to Intonation and Grammar. *Lingua* **29**: 308-325.

Salmons, Joe. 1992. Diachronic Typology and Tone-to-Stress Shift. *The Journal of IndoEuropean Studies* **20.3-4**: 269-281.

Schuh, Russell G. 1978. Tone Rules. In Victoria A. Fromkin (ed.) *Tone: A Linguistic Survey*. New York: Academic Press.

Singler, John V. 1996. Theories of Creole Genesis, Sociohistorical Considerations, and the Evaluation of Evidence: The Case of Haitian Creole and the Relexification Hypothesis. *Journal of Pidgin and Creole Languages* **11.2**: 185-230.

Voorhoeve, Jan. 1959. An Orthography for Saramaccan. *Word* **15**: 436-445.

Voorhoeve, Jan. 1961. Le ton et la grammaire dans le Saramaccan. *Word* **17**: 146-163.

Voorhoeve, Jan. 1968. Towards a Typology of Tone Systems. *Linguistics* **46**: 99-114.

Yip, Moira. 1995. Tone in East Asian Languages. In J.A. Goldsmith (ed.) *The Handbook of Phonological Theory*. Cambridge, Massachusetts: Blackwell.

P.O. Box 22742
San Juan, P.R. 00931

riveray@coqui.net
rivera@bama.ua.edu

Diachronic Aspects of Preferred Argument Structure in English and Broader Implications[1]

REIJIROU SHIBASAKI
University of California, Santa Barbara

0. Preferred Argument Structure: A Background History

Du Bois (1987) derived the theory of Preferred Argument Structure from his research on the ergative language, Sacapultec Maya. Because of no unitary category of subject, Du Bois refers to Dixon (1979) who is one of the initiators of labels for each core argument: A for the subject of a transitive verb (or a two-argument verb), S for the subject of an intransitive verb (or a one-argument verb), and O for the direct object of a transitive verb (or a two-argument verb). The findings from the dissection of Sacapultec Maya by these labels are summarized in (1) (from Du Bois 1987:829). From here on, Preferred Argument Structure is abbreviated as PAS.

(1) Dimensions and Constraints of Preferred Argument Structure

	Grammar	Pragmatics
Quantity	One Lexical Argument Constraint	One New Argument Constraint
Role	Non-Lexical A Constraint	Given A Constraint

Du Bois argues that PAS consists of two dimensions: grammatical and pragmatic dimensions, as in (1). The grammatical dimension has two constraints. One is the 'one lexical argument constraint', which derives from the fact that in Sacapultec narratives, only a small number of clauses have more than two core arguments as lexical NPs. In other words, any additional core arguments tend to appear as pronominals or zero forms in this language. The other constraint is the 'non-lexical A constraint', which means that if there is a lexical NP in a clause, it

[1] I would like to thank Mira Ariel, Susanna Cumming, John W. Du Bois, Carol Genetti, Sandra A. Thompson, Akiyo Maruyama, Joseph Park, Makiko Takekuro, and Toshiko Yamaguchi for invaluable comments for the preparation of this paper. Also, I would like to sincerely thank Elizabeth Closs Traugott for her comments on this paper at the conference. Needless to say, I am responsible for any remaining inconsistencies or mistakes.

tends to appear in S or O, but not in A.

The pragmatic dimension also has a pair of constraints. One is the 'one new argument constraint' which is set out to explain a strong tendency that clauses tend not to contain more than one new argument. The 'given A constraint' is the other constraint to explicate that new arguments have a tendency to appear in S or O, but not in A.

The arguments set forth by Du Bois (1987) have been attested in various languages accompanied with some cross-linguistic validity, but at times they have been challenged with some language-specific[2] or genre-specific behaviors of each label[3] in several languages. For example, it seems universally valid that A is strongly disfavored for encoding new information with a full lexical noun (e.g. see Kärkkäinen (1996) for English, Ashby & Bentivoglio (1994) for French and Spanish, Matsumoto (2000) for Japanese, and Arnold (1998) for a cross-linguistic study including English, Spanish and Mapudungun). On the other hand, language-specific aspects of PAS have been claimed, especially with regard to the role of S. Du Bois (1987) suggests that S and O are considered to behave in the same way to introduce new information in Sacapultec. Herring (1989:126ff) finds, however, that O does not behave like S in Tamil; in other words, S patterns with A in two of her texts, while S patterns with O in the other in Tamil. Durie (1988) also says about PAS in Acehnese (North Sumatra, Austronesian) that "the most salient distinction is between Actors, which rarely code new mentions, and Undergoers, which often do" (p.19) and "97% of new mentions are coded as Undergoers" (p.18).[4]

In addition to synchronic elaborations of PAS, some researchers have succeeded in uncovering the diachronic aspects of PAS. Building on their synchronic study in French and Spanish, for example, Ashby & Bentivoglio (to appear) also discuss diachronic aspects of PAS in French and Spanish, with

---

[2] Several new findings from Nepali are illustrated in Genetti & Crain (to appear). For example, inanimate referents are hard to be mentioned as pronominal, and pronouns never encode new information and are rarely, if ever, accessible or referential. See Du Bois, Kumpf, & Ashby (to appear) for other new works on Preferred Argument Structure in various languages.

[3] For example, O'Dowd (1990:382-83) illustrates that in her English data from paramedical training sessions, S and A are found to contain consistently lower percentage of new information than O and OBL. Kumpf (1993) also claims that "... the characterization of introduction as predominantly in S role is an artifact of the kind of data examined... namely Pear Story narratives" (cited in Kärkkäinen 1996:689). The S role is considered to be sensitive to the difference in genre and changes in discourse, for which Du Bois (1987:836) invented the term 'Information Pressure'.

[4] Preferred Argument Structure seems to be influenced by morpho-syntactic structures or semantic behaviors of each label in a given language. Acehnese is a 'Split-S' language (see Dixon (1979)). Therefore, S still indicates 'half-S' regardless of whether S is split in meaning into Actors and Undergoers. I am grateful to Susanna Cumming for this comment. As to different behaviors of S in her American conversational data from those in Sacapultec Maya in Du Bois (1987), Kärkkäinen (1996:697) attributes the difference to their morphological differences, saying that "... the two subject roles in English, A and S do not differ from each other morphologically, and both govern verb agreement in the same way".

several new speculations on the development of PAS. Resonant with other synchronic accounts of PAS, Ashby & Bentivoglio argue that S is ambivalent about introducing new information across time, and that even intransitive subjects of copulas (Se, in their term) also strongly disfavor new referents across time.[5]

Building on speculations from preceding studies, this study aims to determine whether PAS holds for early stages of English. As Ashby & Bentivoglio (to appear) concede in their concluding remarks, a simple comparison of PAS, in Old French and Modern French for example, is not sufficient for uncovering and understanding the nature of PAS. Therefore, this study aims to show the gradual transition of PAS in the history of English.

1. Data

For this study, the following texts, which may be considered to best reflect colloquial expressions of those times, are selected.

(1) Stage and Text:

Old English (OE):	*The Battle of Maldon* (c.991)
Middle English (ME):	*The Canterbury Tales* (c.1388-1400)
Early Modern English (EModE):	*The Tempest* (c.1612)
Present Day English (PDE):	*Lady Windermere's Fan* (c.1892)

The Battle of Maldon is the last text of the series of epics from *Beowulf.* The important thing is that the first conversational part of this text (lines 29-41) is considered to be 'the first literary use of dialect in English' (Robinson 1976:25-28).[6] From *The Canterbury Tales*, I chose *The Pardoner's Tale*. Because this study focuses on the first 200 clauses from each text to avoid any biases (see section 2), a text was necessary which has as a lot of conversation at the beginning, hopefully of a speaker/narrator's commitment to the story. For this, *The Pardoner's Tale* was one of the best. *The Tempest*[7] and *Lady Windermere's Fan* are both very popular plays of their respective periods. Moreover, the first

[5] In Ashby and Bentivoglio (to appear), there are many other new findings about the development of PAS. One of the most striking speculations concerns the 'pro-drop' phenomenon in Modern Spanish in terms of PAS (footnote 6). According to their analysis, A was already disfavored for the introduction of new information in Old French and Old Spanish (the ratios of lexical NPs in A are 26% and 16%, respectively), and most As were zero forms (56% for Old French and 76% for Old Spanish). A pro-drop phenomenon in Old Spanish was more dominant than in Old French (the overall ratio of zero forms are 28% in Old French and 39% in Old Spanish). This finding enables them to argue that "this trend apparently continues, with the pronoun having become obligatory in Modern French, but not in Modern Spanish." If other stages between Old French/Spanish and Modern French/Spanish provide support for their analysis, this speculation would be further solidified and shed new light on the 'pro-drop' phenomenon in Old French. For other diachronic studies, see papers cited in Ashby & Bentivoglio (to appear).

[6] Hiltunen (1997) suggests that it is possible to analyze *The Battle of Maldon* in terms of 'face' theory introduced by Goffman (1967).

[7] *The Tempest* may be different from the typical verse of Shakespeare's other plays. I am grateful to Elizabeth Closs Traugott for this comment.

parts of them are full of interaction by characters.

2. Methodology

For this study, I coded for four properties. First, the first 200 clauses are selected from each text to avoid any biases. Second, four categories are used for this study: A, S, O and Oblique. I will follow the basic definitions of A, S, and O given in the introduction. The definition of Oblique (hereafter, OBL) is an object of a preposition (see Thompson (1997) for the role of OBL in discourse).

The third coding property is the recency of mention of any lexical nouns and pronouns. There are several notions, terms and definitions relating to recency of mention (see Chapter 2 of Arnold (1998) for details). I shall use the terms, 'Old', 'Active', and 'New', and define them as follows: Old=had appeared in text previously, but not in the immediately previous clause; Active=appeared in the immediately previous clause; New=brand new to the text.

The fourth coding property relates to forms of reference. They are Null, Pronoun, and Lexical Noun (Null=zero forms (see below); Pronouns (hereafter, Pron) = *they*, *her*, *him*, *it*, ...etc.; Lexical (hereafter, Lex) = *hawk*, *messenger*, *Byrhtnoth*, ...etc.). Null forms are marked Ø in my examples, which means that a certain argument which is considered to be called for by the predicate's argument structure is not realized in the clause. These coding properties are illustrated in (3). Note that each category, A, S, O, and OBL are all underlined.

(3) *The Battle of Maldon* 25-28 (10C [c.991])[8]

25	*þa*	*Ø*	*stod*	*on* *stæð,*	*stiðlice*	*clypode*
	then	[messenger]	stood	on shore.DAT.SG	fiercely	called.out
		S/New/Null		OBL/New/Lex		

26	*wicinga*	*ar,*	*wurdum*	*mælde,*
	Viking.GEN.PL	messenger.NOM.SG	word.DAT.PL	spoke
		S/Act/Lex	OBL/New/Lex	

27	*se*	*on*	*beot*	*abead*	*brimliþendra*
	he	in	menace.ACC.SG	announced	seafarer.GEN.PL
	A/Active/Pron		OBL/New/Lex		

28	*ærænde*	*to þam*	*eorle*
	message.ACC.SG	to the	earl
	O/New/Lex		OBL/Old/Lex

'Then [Ø (messenger)] appeared on the opposite bank, a messenger of the Vikings called out fiercely, spoke in words, he threateningly delivered the Vikings' message to the earl (Byrhtnoth)...'

Two lines are added for each line of the original text. The italicized line is the original text, while the second and third lines are allotted for glossing and coding

[8] The glossing conventions are as follows: ACC=accusative; DAT=dative; GEN=genitive; NOM=nominative; PL= plural; SG=singular.

properties for PAS, respectively. The marker Ø in line 25 indicates that there is no syntactic argument in the slot, although it is called for by the argument structure of *mælde* 'spoke'. Abbreviated forms like S/Act/Lex in line 26 indicate, for example, that the word *ar* 'messenger' is regarded as S (= subject of an intransitive (or one-argument) verb, *mælde* 'spoke'), pragmatically Active (= appeared in the immediately previous clause, that is, Ø [= messenger]), and syntactically Lexical, not Pronoun or Null. Because of limitations of space, I will omit other examples.

3. Results

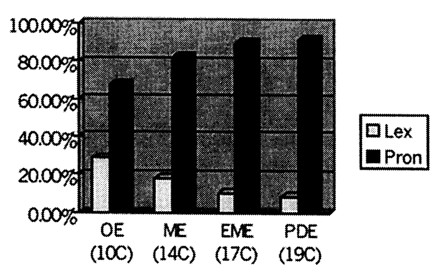
The analysis of the data yields several interesting findings about the diachrony of PAS in English. Some of them give support for Du Bois (1987) and other previous studies, while some seem to be new findings about PAS. Because of limitations of space, I will mainly focus on the proportional frequencies of lexical vs. pronominal mentions of each category, that is, A, S, O, and OBL over history.

3.1 Diachronic Aspects of A in English

The following are the findings about A. As explained in section 0, A is strongly disfavored cross-linguistically for encoding new information with a full lexical noun. The result of this analysis also strengthens this cross- linguistic generalization from a diachronic perspective.

(4) Findings about A based on 4 selected texts
  a. A has come to be disfavored for encoding (New/)Lex over history.
  b. Pronouns are strongly favored for A.
  c. The previous clause is the preferable place for the last mention over history.[9]

(5) Proportional Frequency of Pronominal vs. Lexical Mentions in A



[9] In OE through PDE, the most preferable position of reference for pronouns is the immediately previous clause (= Active, in this study). The ratios of Active over Old and New are: 74.2% (OE), 64.8% (ME), 60.3% (EModE), 68.8% (PDE). Cf. Givón (1983). The reason why Active is most pervasive in OE may be attributive to 'variation', a kind of appositional construction in OE verse.

Ashby & Bentivoglio (to appear) argue that A was already disfavored for new referents in Old French and Old Spanish (the ratios of lexical NPs in A are 26% and 16%, respectively; see footnote 5 for details). In OE, A behaves in the same way as in Old French (the ratio is 27.5%). Importantly, my data suggest that the proportional frequency of lexical nouns has gradually decreased over history, whereas that of pronouns has increased, as illustrated in (5) and (6a-d). This diachronic transition of information status in A supports the two constraints in (1): 'Non-Lexical A Constraint/Given A Constraint', with an implication that A strongly disfavors introducing new information beyond the realms of morpho-syntactic structures of ergative or accusative languages.

(6)

(a) A in OE (10C)

	Null	Pron	Lex	Total
Old	2	6	4	12
Active	1	44	4	49
New	0	2	14	16
Total	3	52	22	77

(b) A in ME (14C)

	Null	Pron	Lex	Total
Old	0	18	4	22
Active	0	50	2	52
New	0	6	10	16
Total	0	74	16	90

(b) A in EModE (17C)

	Null	Pron	Lex	Total
Old	0	10	2	12
Active	0	36	0	36
New	0	8	4	12
Total	0	54	6	60

(c) A in PDE (19C)

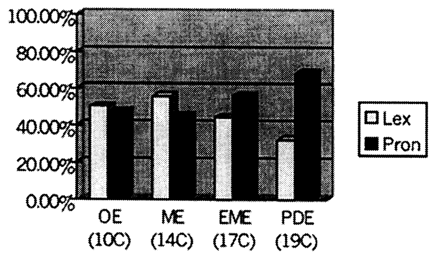
	Null	Pron	Lex	Total
Old	0	32	0	32
Active	0	56	2	58
New	0	2	6	8
Total	0	90	8	98

3.2 Diachronic Aspects of S in English

The findings about S tell us about the ambivalent status of introducing new referents from a diachronic perspective. As pointed out in Kärkkäinen (1996), S in her American English conversational discourse tends to pattern with A rather than O, though she notes that "... S is flexible and reacts to the changing discourse conditions more readily than A" (p.688). The texts used for this study are all written in British English, not in American English. Therefore, it may not be appropriate to simply compare the results of Kärkkäinen (1996) and those of this study; however, it seems possible to show how the S role has come to behave in PDE. In fact, the findings from Kärkkäinen (1996) relate well to those of this study (see below). The following are the diachronic findings of aspects of PAS in English. Note that the copula's complement (S-Comp) is included in S.

(7) Findings about S based on 4 selected texts
  a. (Active/)Pron and (New/)Lex were competing in OE and ME.
  b. After ME, (New/)Lex began to decrease, while (Act/)Pron became dominant in PDE.
  c. But still, S behaves differently from A.

(8) Proportional Frequency of Pronominal vs. Lexical Mentions in S



According to Ashby & Bentivoglio (to appear), S has been ambivalent about introducing new referents in Old/Modern French and Spanish. The findings from OE and ME support their finding about S. However, the ratio of introducing new referents in S began to decrease after ME on, as in (8) above, which goes against the diachronic findings for PAS in French and Spanish. It is true that the results from this study may not go beyond the realms of case study; however, the ratios of lexical nouns over pronouns have consistently become lower from ME through PDE.

(9)

(a) S in OE (10C)

	Null	Pron	Lex	Total
Old	0	2	8	10
Active	2	30	12	44
New	0	0	14	14
Total	2	32	34	68

(b) S in ME (14C)

	Null	Pron	Lex	Total
Old	0	12	8	20
Active	0	30	0	30
New	0	6	52	58
Total	0	48	60	108

(c) S in EModE (17C)

	Null	Pron	Lex	Total
Old	0	8	4	12
Active	0	30	0	30
New	0	8	32	40
Total	0	46	36	82

(d) S in PDE (19C)

	Null	Pron	Lex	Total
Old	0	18	2	20
Active	0	64	8	72
New	0	18	38	56
Total	0	100	48	148

Importantly, the low frequency of lexical nouns in S from ME through PDE can lead to Kärkkäinen's (1996) finding that S behaves like A in two of her three texts in American conversational discourse. This fact deserves consideration. S is sensitive to the difference in genre and changes in discourse in the sense of 'Information Pressure' (Du Bois 1987:836), as is often suggested in various studies (see note 3 and 4). In my opinion, however, S also seems to have changed its role, at least in the English texts I looked at, from favoring new referents to disfavoring them over time. This speculation needs further research, but the findings from this study support this view, as is summarized in (8) and (9a-d).

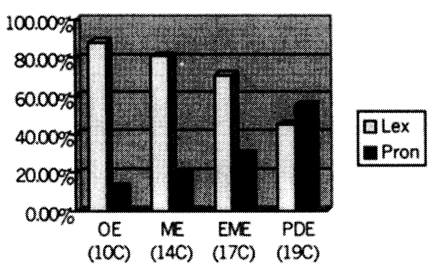
3.3 Diachronic Aspects of O in English

The findings about O are intriguing both diachronically and typologically. For example, Ashby & Bentivoglio (to appear) argue that O favors the introduction of new information in both Old and Modern French and Spanish. Likewise, synchronically, many studies reach a consensus on the skewed behavior of O introducing new information. However, the result of this study casts doubt on this almost unanimous view, as follows.

(10) Findings about O based on 4 selected texts
   a. In OE through EModE, O was preferred for the introduction of (New/)Lex.
   b. In PDE, O came to prefer (Act/)Pron to (New/)Lex.

As in (11), the proportional frequency of lexical nouns over pronouns has gradually but steadily decreased across time, and then reversed in PDE. It is attested in various languages that O strongly prefers new information to old/given information; however, it seems that the reversed situation has not yet been reported either typologically or diachronically.

(11) Proportional Frequency of Pronominal vs. Lexical Mentions in O



(12)

(a) O in OE (10C)

	Null	Pron	Lex	Total
Old	0	2	14	16
Active	1	6	24	31
New	0	8	80	88
Total	1	16	118	135

(b) O in ME (14C)

	Null	Pron	Lex	Total
Old	0	6	4	10
Active	0	14	2	16
New	0	2	88	90
Total	0	22	94	116

(c) O in EModE (17C)

	Null	Pron	Lex	Total
Old	0	6	6	12
Active	0	16	0	16
New	0	10	72	82
Total	0	32	78	110

(d) O in PDE (19C)

	Null	Pron	Lex	Total
Old	0	6	6	12
Active	0	46	6	52
New	0	2	32	34
Total	0	54	44	98

Because this investigation does not go beyond the realms of case study, I cannot make any broad generalizations about O. Yet, as discussed in footnote 9, pronouns have a strong tendency to refer to NPs in the immediately previous clause (=Active, in this study), and the number of pronouns has been increasing over history, as shown later. These two findings may support the diachronic transition of the O role; the pragmatic role of introducing new information may have been taken over to OBL, as shown in section 3.4. I will tentatively conclude here that, in English, pronouns frequently refer to NPs in the immediately previous clause, which enables the speaker/hearer or the writer/reader to easily perceive the information flow.[10]

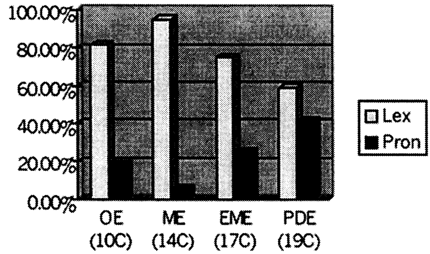
3.4 Diachronic Aspects of OBL in English

Thompson (1997:75) argues that the pragmatic role of OBL is rarely Given, Identifiable, and Tracking. In her English conversational data, the ratios of Given information in core and oblique labels are: 89% in A, 65% in S and O, and 35% in OBL (p.72). It is worth noting that new information is skewed in OBL in her data. The following are my findings about OBL.

(13) Findings about OBL based on 4 selected texts
  a. OE through PDE, OBL has been preferred for the introduction of (New/)Lex.
  b. In PDE, however, OBL comes to take (Act/)Pron though less preferable than (New/)Lex.

[10] As pointed out by Elizabeth Closs Traugott, pronouns function cataphorically as well as anaphorically in OE more often than not. Yet as shown in note 9, the proportional frequency of reference is anaphoric-oriented.

(14) Proportional Frequency of Pronominal vs. Lexical Mentions in OBL



(15)

(a) OBL in OE (10C)

	Null	Pron	Lex	Total
Old	0	8	42	50
Active	0	12	6	18
New	0	4	60	64
Total	0	24	108	132

(b) OBL in ME (14C)

	Null	Pron	Lex	Total
Old	0	2	6	8
Active	0	2	2	4
New	0	0	68	68
Total	0	4	76	80

(c) OBL in EModE (17C)

	Null	Pron	Lex	Total
Old	0	4	4	8
Active	0	12	0	12
New	0	8	68	76
Total	0	24	72	96

(d) OBL in PDE (19C)

	Null	Pron	Lex	Total
Old	0	4	2	6
Active	0	24	2	26
New	0	4	42	46
Total	0	32	46	78

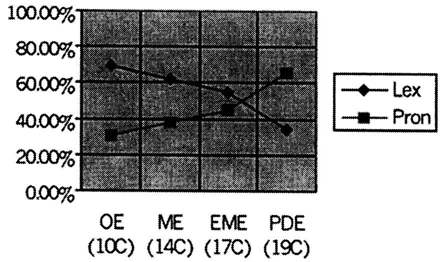
As shown in (14), the diachronic aspects of OBL may support the Thompson's findings for OBL. The important thing is, however, that in my data from ME on, OBL has been disfavoring the introduction of new information. Suppose that this tendency continues, the ratio of pronominal mentions in OBL would be dominant over that of lexical mentions. This speculation nicely leads to the findings in Kärkkäinen (1996:680ff). According to Kärkkäinen, OBL disfavors introducing new referents in (two of her) American conversational texts (the average ratio of lexical mentions is 22.3% in OBL).[11] My texts do not provide us with such a dramatic picture; however, the findings from Kärkkäinen (1996), Thompson (1997), and this study suggest that in contrast to previous research, there is possibility that even OBL may be influenced by genres and types of discourse (Information Pressure).

[11] It is noteworthy that the ratio of lexical mentions in OBL (22.3%) is less than that of O (31.6%) in Kärkkäinen (1996:680). Yet pragmatic dimensions of her data show a clear continuum of new information skewing: 5.1% in A, 7.9% in S, 47.6% in O, and 61.4% in OBL (*ibid.*:684).

4. Summary and Diachronic Implications

We have thus far investigated the paths each category has traveled through time. These paths are complex in that the information status of each category varies from stage to stage; however, the diachronic transition of PAS is not random, but is motivated. Because of limitations of space, I cannot present all findings from this study, but I will focus on two important findings which go beyond the previous views of PAS both cross-linguistically and diachronically.

(16) Transition of Pronouns vs. Lexical Nouns in PAS (a case study)



One is that all categories have come to favor pronominal mentions over history; in other words, both the number and proportional frequencies of pronouns have increased in all categories over history: 124/412 (30.1%) in OE, 148/394 (37.6%) in ME, 156/348 (44.8%) in EModE, 276/422 (65.4%) in PDE. These are summarized in (16). This study is just based on a kind of verse texts, not prose texts; therefore I will not generalize this diachronic transition to other genres. However, building on the findings from this study, I am led to conclude that information has come to be oriented toward the preceding discourse rather than the following discourse. Remember that as discussed in note 9, the most preferable position where pronouns refer to is the immediately previous clause. Thence, it can be considered that English has come to favor 1) lexically pronominal mentions and 2) pragmatically anaphoric functions. These are not absolute, but can be supported by the findings from this study.

The other important thing is that each category has grammaticalized the need to place the most appropriate information in it. As illustrated in section 3, each category shows its grammaticalized information status at each stage. They are summarized in (17). Note that the more to the left a category is, the less likely it is to be a lexical or new mention.

(17) Transition of Implicational Hierarchy

A (28.6%)	<	S (50%)	<	OBL (81.8%)	<	O (88.1%)	OE (10C)
A (17.8%)	<	S (55.6%)	<	O (81%)	<	OBL (95%)	ME (14C)
A (10%)	<	S (43.9%)	<	O (70.9%)	<	OBL (75%)	EME (17C)
A (8.2%)	<	S (32.4%)	<	O (44.9%)	<	OBL (60%)	PDE (19C)

Although the order of O and OBL in OE is reversed in ME, the hierarchy: A < S < O < OBL, based on the ratio of new and lexical information, seems to have been grammaticalized over history, at least after ME. This motivated hierarchy has diachronically moved in one direction, with the increasing number and proportional frequencies of pronouns, as in (16). All in all, information flow has been motivated grammatically and pragmatically in English, and diachronically as well.

References

Arnold, J. 1998. Reference form and discourse patterns. PhD dissertation, Stanford University.

Ashby, W. J. & Bentivoglio, P. 1993. Preferred argument structure in spoken French and Spanish. *Language Variation and Change* **5**: 61-76.

Ashby, W. J. & Bentivoglio, P. to appear. Preferred argument structure across time and space. In Du Bois, J. W., Kumpf, L. & Ashby, W. J. (eds.) *Preferred argument structure*. Amsterdam: John Benjamins.

Benson, L. D. et al. (eds.) 1987. *The Riverside Chaucer*. Oxford: Oxford UP.

Dixon, R. M. W. 1979. Ergativity. *Language* **55** (1): 59-138.

Du Bois, J. W. 1987. The discourse basis of ergativity. *Language* **63** (4): 805-855.

Durie, M. 1988. Preferred argument structure in an active language. *Lingua* **74**:1-25

Genetti, C. & Crain, L. D. to appear. Beyond preferred argument structure: Sentences, pronouns, and given referents in Nepali. In Du Bois, J. W., Kumpf, L. & Ashby, W. J. (eds.) *Preferred argument structure*. Amsterdam: John Benjamins.

Gibson, R. (ed.) 1993. *The Tempest*. Cambridge: Cambridge UP.

Givón, Talmy. 1983. Topic continuity in discourse: an introduction. In T. Givón, (ed.) *Topic continuity in discourse*. Amsterdam: John Benjamins.

Goffman, E. 1967. *Interaction Ritual*. Garden City, NY: Anchor Books.

Herring, S. 1989. Verbless presentation and the discourse basis of ergativity. *Papers from the 25th Annual Meting of the Chicago Linguistic Society*, 123-137. Chicago: Chicago Linguistic Society.

Hiltunen, Risto. 1997. Aspects of *The Battle of Maldon*. A lecture/handout at the Rikkyo Seminar on English Philology, May 20, 1997, at the Dept. of English, Rikkyo (St. Paul) University, Tokyo, Japan.

Kärkkäinen, E. 1996. Preferred argument structure and subject role in American English conversational discourse. *Journal of Pragmatics* **25** (5): 675-701.

Kumpf, L. 1993. Grammatical roles and participant introduction in native English and second language discourse. Ms., California State University, Long Beach.

Matsumoto, K. 2000. Intonation units, clause and preferred argument structure in conversational Japanese. *Language Science* **22**: 63-86.

O'Dowd, E. 1990. Discourse pressure, genre and grammatical alignment – after Du Bois. *Studies in Language* **14** (2): 365-403.

Raby, P. (ed.) 1995. *The Importance of Being Earnest and Other Plays* (by Oscar Wilde). Oxford: Oxford UP.

Robinson, F. C. 1976. Some aspects of the *Maldon* Poet's artistry. *Journal of English and Germanic Philology* **75**: 25-40.

Scragg, D. G. (ed.) 1981. *The Battle of Maldon*. Manchester: Manchester UP.

Thompson, S. A. 1997. Discourse motivations for the core-oblique distinction as a language universal. In Akio Kamio (ed.) *Directions in Functional Linguistics*. Amsterdam: John Benjamins.

Reijirou Shibasaki
Department of Linguistics
University of California, Santa Barbara
Santa Barbara, California 93106-3100
U.S.A.

reijirou@umail.ucsb.edu

Discourse-Related Effects on Speech Durations: A Challenge for Models of Speech Production[1]

CAROLINE L. SMITH
University of New Mexico

1. Discourse prosody

A substantial body of work has shown that prosodic characteristics of speech are affected by factors which can be conveniently referred to as "discourse-related". That is, the global organization of the spoken material has consequences for such properties as the pitch, durations and amplitude. Early research in this area focused on prosodic differences between boundaries that were perceived as smaller ("sentence") or larger ("paragraph"), and found differences in F0 and duration (Lehiste 1975, 1979; Kreiman 1982). Recent work has tended to concentrate on how the organization of the discourse is reflected in intonation (e.g., Grosz & Hirschberg 1992; Ayers 1994; Swerts & Geluykens 1994). Other studies have found evidence of an effect of discourse structure in several acoustic dimensions. For example, Herman (2000) compared the realization of the same English sentences produced either discourse-medially or discourse-finally, and found differences in F0, duration of the final pitch-accented syllable and RMS amplitude, even though she compared only those pairs of productions in which the intonational tones were phonologically identical, in order to ensure that any differences did not reflect distinctive variants of the sentence. In a study of direction-giving monologues in English, phrases at the beginning, middle and end of discourse segments were found to differ in speech rate, pause duration and several measures of F0 (Hirschberg & Nakatani 1996).

The studies just described looked at discourse structure by comparing phrases or sentences at different positions in a discourse. Other studies have looked at correlates of prosodic characteristics with different aspects of discourse structure. Studying spontaneous re-tellings of a short Dutch narrative, van Donzel (1999) found that both the magnitude of discourse boundaries and the information structure of the discourse contributed to determining intonational realizations and pause durations. Also in Dutch, Noordman *et al.* (1999) and den Ouden *et al.*

[1] This work was supported by NSF grant BCS-9983106. Many thanks to Lisa Hogan for analyzing the data and assisting with manuscript preparation.

(2000) studied speakers' reading of narrative texts, and found that F0 means, maxima and pause durations were affected by the hierarchical structure of the narrative as determined by theories of discourse structure.

The various studies cited above used different criteria for characterizing the status of an individual phrase or sentence in the larger structure of the discourse. A slightly different approach was used in the present experiment. Rather than attempting to analyze the overall organization of a discourse, the analysis focused on the relation between the topic of one sentence and of the sentence which follows. The scheme used by Nakajima and Allen (1993) for labeling topic boundaries in spontaneous instructional monologues was adapted to the needs of this study, which was based on a written instructional text. This topic-labeling scheme is adequate for describing the organization of instructional material, but is not intended to be a general theory of discourse organization. Its relative simplicity was a virtue for this study in part because it may ultimately be easier to generate an analysis of this type automatically. This would enable speech synthesis systems to analyze the topical organization of a text, then take it into account when determining what prosody to produce.

With the topic labeling providing an analysis of the text's organization, the experiment reported here investigated the relation between the structure of the text and the durational patterns produced when it was read aloud. Durations were chosen for study in part because much previous work has concentrated on F0. In addition, acoustic durations in English are known to be affected by segment identity and local context in a way that F0 is not, so it is interesting to see whether, like F0, durations are also affected by the overall organization of the discourse. The present paper focuses on the interactions between discourse organization and durations, and how these may reveal drawbacks in current models of speech production.

2. The Experiment

This experiment measured the effects of different types of topic transition on several durational properties (more details are reported in Smith and Hogan (2001)). In this experiment a male speaker of American English read a passage drawn from the manual for the computer drawing program Canvas (Deneba Systems 1997). The text was chosen because it offered a relatively well-defined topic structure for the analysis; moreover, such a text typifies the sort of material that speech synthesizers currently read aloud in, for example, Help systems.

The speaker was recorded reading the text aloud along with a set of "control" sentences ten times at intervals of approximately one week. The control sentences were constructed so that the final word of each sentence in the Canvas text—the "target" word—occurred in a sentence-medial position in the control sentence. For example, one sentence ending with *box* in the original text was as follows [italics not in original text]:

(1) You can search and replace character strings that you specify using the Find dialog *box*. (Deneba 1997:224)

The corresponding control sentence for the target word *box* is given in (2). In all control sentences, the target word was placed in the sentence so that three syllables preceded it and eight syllables followed it:

(2) The dialog *box* lists all settings in the program.

The Canvas text passage encompassed 60 sentences which ended in 38 different target words. Three durational measures were made on each sentence in the text:

- Sentence-final lengthening, measured as the increase in the duration of the target word when it occurred in sentence-final position in the text as compared to its duration sentence-medially in the control sentence
- Speech rate, measured as the number of syllables per second in the interpausal speech runs at the end of each sentence and the beginning of the following sentence
- Pause durations, measured as the length of time that elapsed between sentences.

In the analysis of the relation between these durational properties and the organization of topics in the text, the transition from one sentence of the text to the next was classified into one of four categories:

- ***Topic Shift***, if the following sentence introduced new material
- ***Topic Continuation***, if the following sentence continued the topic, advancing the narrative
- ***Elaboration***, if the following sentence provided additional detail about the preceding sentence
- ***Text Marker***, if the following sentence were an overt indicator of textual organization (an example of a Text Marker is *Note:*).

The classification was done by five linguists; in cases of disagreement, the labeling preferred by the majority was used. Below is a brief extract from the text (Deneba 1997:226) with the topic transitions labeled and the target words in italics:

(3) ...Otherwise, choosing Interactive will turn this feature *on*. (Topic Shift)

Spell checking *text* (Topic Continuation)
You can check the spelling of highlighted blocks of text, a selected text objects, or an entire *document*. ...

The analysis of the durational measures for each type of topic transition yielded these results:

- Topic Shifts occurred with significantly longer pauses than other types of transitions. Speech rate tended to be slower at a Topic Shift than at other transitions, but did not change at the transition from one sentence and the next.
- Sentence-final lengthening was similar in Topic Shifts, Continuations and Text Markers. However, at a Topic Continuation, speech rate increased significantly between the end of the first sentence and the beginning of the following sentence.
- Topic Elaborations had significantly less final lengthening than other transitions. Pauses were of similar duration for Elaborations, Continuations and Text Markers. Speech rate was faster for Elaborations and Text Markers, but did not change at the transition.

All three measures of duration were affected by the type of topic transition from one sentence to the next. Additionally, different types of transition were associated with different configurations of the three durational measures. Although these results apply to one speaker only (analysis of additional speakers is underway), the effects were substantial and statistically robust. For example, the amount of sentence-final lengthening at Topic Shifts was over twice what occurred at Elaborations. Therefore, these patterns ought to be incorporated into any model representing the factors which potentially contribute to the process of speech production.

But the explanation of these effects must be more complex than the usual explanations of segmental context and phrasal organization—the relationship between the topic of one sentence to the next causes differences in the durational properties of the end of the first sentence, the beginning of the second sentence, and the pause between them. The temporal extent of these differences means that the transition between sentences is not completely localized at the boundary between the sentences. This finding presents problems for models of speech production.

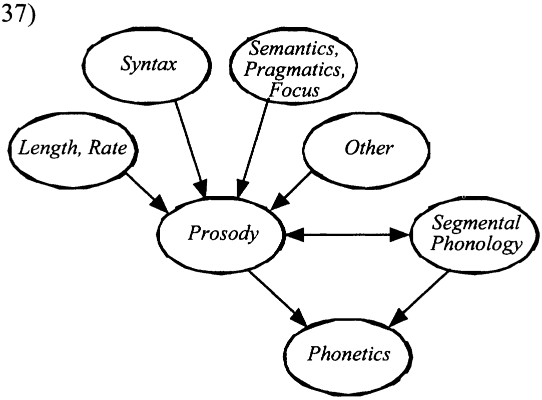
3. Models of speech production

One of the difficulties with the findings reported in section 2 is that the effects of topic organization relate to a subsequent sentence as well as the current one, so information outside the current prosodic domain must be accessed. In addition, these effects relate to the utterances' content and meaning, not their structure, and content is not usually considered to be the kind of information that is accessible to the prosody. The remainder of this paper considers the problems that these considerations raise for incorporating the results of this study into two representative models of speech production. The two models are based on extremely different representations of linguistic knowledge. One, referred to here as the "modular" model, assumes that grammar is divided into separate modules, each of which has access to only a subset of linguistic structure. Processing by different modules brings together the information needed to produce speech. The other model, referred to here as the "exemplar" model, does not differentiate among different aspects of linguistic structure. All the information needed to produce speech is combined in the representation, but complex processing is required to access it.

3.1 Conventional "modular" model

Current theories about the organization of speech production most often assume a model somewhat like that shown in (4), which is adapted from Shattuck-Hufnagel and Turk (1996). In this model, a variety of grammatical and extra-grammatical factors contribute to the determination of the prosody, but the model permits only the prosody and the segmental phonology to affect the phonetics (which presumably includes determination of durations).

(4) "Modular" model, adapted from Figure 5 in Shattuck-Hufnagel and Turk (1996:237)



The problem with this model is that topic organization would be able to affect durations only via the prosody. In other words, different types of topic transitions would have to be associated with differences in the prosodic structure which in turn would create differences in durations. The question becomes what prosodic unit(s) delimit the boundaries at which the topic transitions occur. The version of the prosodic hierarchy discussed by Shattuck-Hufnagel and Turk (1996) does not include any units larger than the intonational phrase, but many of the sentences in this experiment included more than one intonational phrase (IP). This is a problem if the IP is the largest unit available to associate with the boundaries at the transitions. However, if the Strict Layer Hypothesis must be modified so that

IPs can be nested within larger IPs, as Ladd (1996:244) argues, then a prosodic structure could be constructed so that each written sentence is coextensive with a (possibly nested) IP. This is one approach to the problem of relating sentences in a written text to prosodic units in speech. The assumption that topic transitions occur at sentence boundaries is appropriate in written text because a sentence contains a complete proposition. However, nesting IPs in order to make sentence boundaries coincide with IP boundaries would not, in itself, predict the observed differences at different types of transitions.

Two ways in which IPs might form the basis for an alignment of topic organization with prosodic structure are as follows: (i) IPs could be categorized in some way according to the type of topic relation that holds between adjacent phrases; or (ii) they could be organized into larger, superordinate units, and their position within the larger structure would correlate with topic organization. Both of these possible solutions attempt to introduce a reflex of structure in the semantic/discourse domain into prosodic structure. Proposal (i) would violate basic assumptions of the hierarchical model of prosody by distinguishing among sister units solely on the basis of explicit labels, rather than distinguishing them by their structure. Identifying stress feet as strong or weak might seem to be an accepted usage of labels on prosodic units, but this distinction is actually made on the basis of the position and/or syllabic content of each foot. The labels on stress feet are for convenience, whereas the proposed labels on IPs designating their transition types would encode crucial differences.

Although Shattuck-Hufnagel and Turk (1996) do not mention it, the phonological utterance is the top level of the prosodic hierarchy in many versions of prosodic theory. Under the right circumstances, an utterance can include more than one sentence, so proposal (ii) might seem to pose little difficulty. Sentences which are closely related could be connected into a single phonological utterance, which would predict minimal marking of the transition between them. A less closely-related pair of sentences could be in separate utterances, in essence placing a larger boundary between them, which predicts greater marking of the transition. But even this proposal, in the spirit of the prosodic hierarchy, does not capture the complexity observed in the durational marking of transitions between sentences.

First of all, even in the very simple transition-labeling scheme used in this study, there were four types of transition, each with its own durational characteristics. In order to distinguish among four types, it would be necessary to posit more different structures than just the distinction between a sentence which ended at an utterance boundary and one not at an utterance boundary. A four-way distinction would require expanding the prosodic hierarchy. More significantly, statistical tests showed that the various durational measures were very rarely correlated. Each appears to vary independently of the others so they do not combine to make a consistently "bigger" boundary at a Topic Shift, for example, than at a Topic Continuation. Rather, these are different types of boundary. For this reason, the topic effects could not be handled by treating Topic Shifts as the delimiter of a prosodic unit "topic", with the other transition types delimiting smaller units included within a "topic". Such a structure would fail to represent the differences among Topic Continuations, Elaborations and Text Markers. A hierarchical prosodic structure offers no tidy way to represent this.

I conclude that in order to represent the types of effects observed in the present study, the "modular" model would have to be significantly altered, and weakened. Either some kind of provision would have to be added for labeling prosodic units according to their semantic/pragmatic/discourse function, or the organization of the model would have to be restructured so that these components have direct access to the phonetics, nullifying the crucial role of the prosody in mediating the many factors that have been shown to play a role in production.

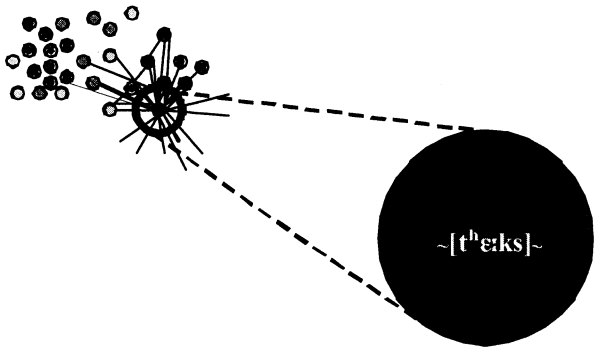
3.2 Exemplar models

The "modular" models described in section 3.1 compartmentalize different aspects of linguistic structure; the model's predictive power comes, in part, from constraints that the structure imposes on how each component can reference properties under the control of another component. Exemplar models take an alternative view, in which all aspects of linguistic structure have the potential to interact. Exemplar models are somewhat akin to the connectionist models used in psychology to model neural processes (Rumelhart *et al.* 1986). In the linguistic literature, exemplar models have been discussed chiefly as an account of speech perception (Johnson 1997). Bybee (2001) and Pierrehumbert (2001, in press) may be the first to explore how these models can represent speech production.

Different authors have proposed variants of exemplar models for language processing. The description here is a somewhat simplified account based principally on Bybee (2001). The core idea of an exemplar model applied to language is that language users have knowledge of specific instances of linguistic units recorded in memory. These specific instances are the exemplars. Different versions of the theory have different proposals about the size of the unit(s) that are stored; here, for simplicity, I will assume that the stored unit is the word. In this case, each exemplar records a particular pronunciation and usage of a word. In (5), the large circle represents one exemplar of the word *text*, depicting a pronunciation in which the vowel is lengthened but the final [t] is not pronounced. (In this and the next figure, a phonetic transcription is enclosed in square brackets; the ~ ~ notation is being used to symbolize an exemplar, which includes information about meaning and context, not just the pronunciation.)

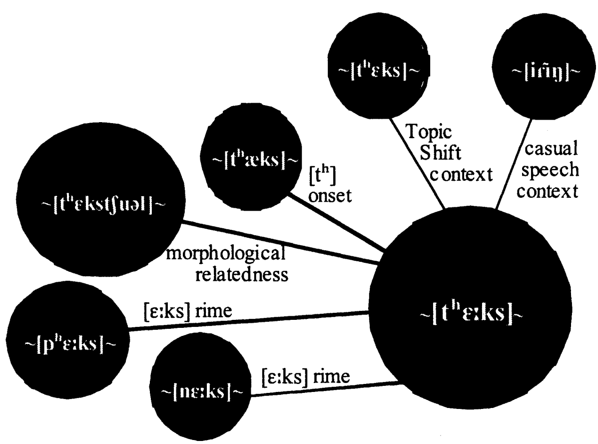
An exemplar is activated each time it is used, with activation decaying over time. Recently-used exemplars are therefore more activated than those not used for some time, and more frequently-used exemplars are more activated than those used less often. Stronger activation is represented in (5) as darker shading.

(5) A group of exemplars, showing some of the connections to one exemplar of the word *text* pronounced with a lengthened vowel



Each exemplar is linked to other exemplars with which it shares some property. The shared property may relate to any aspect of the exemplars—phonological, semantic, contextual, to name a few. A subset of the links to one exemplar is shown in (5) and expanded in (6). This exemplar, with its lengthened vowel, would be linked not only to other exemplars of the same word, but to exemplars of phonologically, morphologically or semantically related words. It would also be linked to other exemplars that occurred in a similar context. For example, if this exemplar was in a Topic Shift context, it would be linked to other exemplars that occurred at Topic Shifts.

(6) Some examples of shared properties which could result in connections being formed among exemplars of the same and different words.



This linking of words occurring in similar contexts leads to a possible scenario for an exemplar model account of the patterns associated with topic transition types. If a word occurs with a lengthened vowel at the end of a sentence preceding a Topic Shift, then other exemplars that are linked to it by virtue of having similarly lengthened vowels, will also be activated. If Topic Shifts and lengthened vowels co-occur again and again, the pattern of activation is reinforced, so that eventually the language user forms a "schema" among exemplars with lengthened vowels occurring in Topic Shift context. According to Bybee (2001:39), schemas are "emergent generalizations over complex representations." That is, the representation stores a great deal of information about the exemplars that the language user has experienced, and the patterns of activation in the links among these exemplars result in the emergence of generalizations about the sets of exemplars connected by these links. These generalizations can then result in the associated properties spreading to new exemplars. Variable patterns such as those due to topic organization can emerge from the process of producing speech, even if they do not always occur, or occur to varying degrees. Variability in the input will be reflected by variability in the output. Both are expected.

As described here, exemplar models are very flexible; essentially any property that is shared by a group of exemplars has the potential to generalize as a schema, if it occurs with sufficient frequency. This is both the strength and weakness of these models. Their flexibility means that they can account for all kinds of patterns of co-occurrence among different kinds of linguistic properties. At the same time, without limits on possible schemas, the model makes no predictions about which are more likely, beyond the fact that more frequently occurring patterns are more strongly activated and more likely to generalize.

Exemplar models predict that generalizations about a specific context will originate with words which the language user has experienced in that context. Only after repeated usage of words associating a property and context can a pattern be generalized as a schema. This process seems like an unlikely path for the generalization of effects such as those associated with topic transitions. These effects depend on the relation between two entire sentences, not specific words, and are very unlikely to occur more than once with the same words. Without repetition, they would not be able to generalize, since the connection between a property and a context is associated with specific exemplars. It is not clear how a schema could develop from a pattern that is not tied to specific linguistic units. The generalization would have to come from repeated co-occurrence of activation in two sets of connections: those connecting exemplars sharing particular topic contexts and those sharing the associated durational properties. Modification or extension of the model would be required in order to allow a schema to develop without repetition of individual exemplars exhibiting a pattern.

The flexibility of the exemplar model means that it could incorporate these effects probably with less modification than the modular model, but this same flexibility means that it could potentially generalize other patterns which co-occur accidentally. The generality of this model is both its strength and its weakness.

4. Summary

In recordings of one American English speaker reading an instructional text aloud, durational differences were observed which appear to depend on the relation between the topics of consecutive sentences. These differences, along with the findings of other similar research, pose a problem for current models of speech production, because they suggest that the structure of the semantic/pragmatic information in a discourse can have measurable effects on the acoustic durations. While it has long been known that durations are influenced by many factors (see, for example, Klatt 1976), neither of the models of speech production as discussed here can immediately incorporate the effects of topic organization. The ideal model would be more constrained than an exemplar model, but more flexible in allowing interaction among different components of the grammar than the modular framework assumed in much current research. Such a model, which remains to be developed, would be valuable in helping to understand how the physical dimensions of speech reflect the different linguistic dimensions of the message.

References

Ayers, Gayle. 1994. Discourse functions of pitch range in spontaneous and read speech. *OSU Working Papers in Linguistics* **44:** 1-49.

Bybee, Joan. 2001. *Phonology and Language Use*. Cambridge: CUP.

Den Ouden, Hanny, Leo Noordman, & Jacques Terken. 2000. Prosodic correlates of text structure. *Proceedings of the 10th Annual Meeting of the Society for Text and Discourse*, 40-41.

Deneba Systems. 1997. *Canvas 5 User's Guide*. Miami: Deneba Systems.

Grosz, Barbara & Julia Hirschberg. 1992. Some intonational characteristics of discourse structure. *Proceedings of the 2nd ICSLP*, Banff, Alberta, 429-432.

Herman, Rebecca. 2000. Phonetic markers of global discourse structures in English. *Journal of Phonetics* **28:** 466-493.

Hirschberg, Julia & Christine Nakatani. 1996. A prosodic analysis of discourse segments in direction-giving monologues. *Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics*, 286-293.

Johnson, Keith. 1997. Speech perception without speaker normalization. In K. Johnson & J. Mullennix (eds.) *Talker Variability in Speech Processing*, 145-165. San Diego: Academic Press.

Klatt, Dennis H. 1976. Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America* **59**: 1208-1221.

Kreiman, Jody. 1982. Perception of sentence and paragraph boundaries in natural conversation. *Journal of Phonetics* **10:** 163-175.

Ladd, D. Robert. 1996. *Intonational Phonology*. Cambridge: CUP.

Lehiste, Ilse. 1975. The phonetic structure of paragraphs. In A. Cohen & S. G. Nooteboom (eds.) *Structure and Process in Speech Perception*. Proceedings of the Symposium on Dynamic Aspects of Speech Perception, 195-203. New York: Springer-Verlag.

Lehiste, Ilse. 1979. Perception of sentence and paragraph boundaries. In B. Lindblom & S. Ohman (eds.) *Frontiers of Speech Communication Research*, 191-201. London: Academic Press.

Nakajima, Shin'ya & James Allen. 1993. A study on prosody and discourse structure in cooperative dialogues. *Phonetica* **50:** 197-210.

Noordman, Leo, Ingrid Dassen, Marc Swerts, & Jacques Terken. 1999. Prosodic markers of text structure. In K. van Hoek, A. Kibrik, & L. Noordman (eds.) *Discourse Studies in Cognitive Linguistics*. Selected Papers from the 5th International Cognitive Linguistics Conference, 133-148. Amsterdam: John Benjamins.

Pierrehumbert, Janet. 2001. Exemplar dynamics: word frequency, lenition and contrast. In J. Bybee & P. Hopper (eds.) *Frequency Effects and Emergent Grammar*, 137-157. Amsterdam: John Benjamins.

Pierrehumbert, Janet. In press. Word-specific phonetics. In C. Gussenhoven & N. Warner (eds.) *Laboratory Phonology VII*. Berlin: Mouton de Gruyter. [from http://www.ling.nwu.edu/~jbp/publications.html]

Rumelhart, David, James McClelland & the PDP Research Group. 1986. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge, MA: MIT Press.

Shattuck-Hufnagel, Stephanie & Alice Turk. 1996. A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* **25:** 193-247.

Smith, Caroline & Lisa Hogan. 2001. Variation in final lengthening as a function of topic structure. *Proceedings of Eurospeech 2001 Scandinavia*, 955-958.

Swerts, Marc & Ronald Geluykens. 1994. Prosody as a marker of information flow in spoken discourse. *Language and Speech* **37:** 21-43.

Van Donzel, Monique. 1999. *Prosodic Aspects of Information Structure in Discourse*. The Hague: Thesus.

Caroline L. Smith
Department of Linguistics
University of New Mexico
Humanities 526
Albuquerque, NM 87131-1196

caroline@unm.edu

Syntactic Projectability and Co-Participant Completion in Japanese Conversation

POLLY SZATROWSKI
University of Minnesota

0. Introduction

In this paper, I investigate the role of syntax in projecting possible ends of utterances. Research on turn-taking and co-participant completion has concluded that co-participants use many concomitant features of talk-in-interaction in addition to grammatical structure to project the ends of utterances. My purpose is not to refute the contribution of these concomitant features, but rather to refine further the role of grammatical structure in syntactic projectability.

The notion of projectability has been a key concept for determining the place where turn transition is relevant in conversation analytic research. According to Sacks, Schegloff, & Jefferson (1974:702), speakers construct turns using sentences, clauses, phrases, and lexical constructions which "allow a projection of the unit-type under way, and what, roughly, it will take for an instance of that unit-type to be completed." I will use the term syntactic projectability to refer to the ability to project a syntactic category that may come later in an utterance. I define co-construction as the creation of a relative clause, clause, simple or complex sentence by 2 or more participants. I will mark co-participant completions, that is, the completion of one speaker's utterance by another participant, with an arrow in the examples in this paper.

Research by Fox *et al.* (1996) and Hayashi (1999) contrast the tightly knit clause structure in English to loosely connected elements in Japanese, and use this as a basis for claiming that Japanese lacks "early projection" strategies. I hypothesize that co-construction will be more likely when there is higher syntactic projectability and restriction on what comes later in a sentence, and will demonstrate that there are syntactic practices in Japanese which allow for early projection.

1. Previous research on co-construction

In this section, I will review research related to co-construction by Lerner (1991), Mizutani (1993), Fox *et al.* (1996), and Hayashi (1997). Then I will introduce Minami's model for Japanese syntax which I will use in Section 3 to analyze syntactic projectability in Japanese.

Lerner (1991:444) defines a "compound turn constructional unit (TCU)" as "[a]ny turn unit which in the course of its construction projects a [preliminary component + final component] format" as in (1). K projects what comes next from D's 'if' clause, and completes the compound TCU.

(1) 1D: so if one person said he couldn't invest	Preliminary Component
→ 2K: then I'd have to wait	Final Component
(Lerner 1991:445)	

Mizutani (1993:7) suggests that co-construction is possible in Japanese because one can project the end of an utterance from modal adverbs and conjunctions. Modal adverbs such as *doo mo* 'somehow', *doo yara* 'somehow', *tyotto* 'a bit', *nan to naku* 'somehow' link with sentence endings, and clauses which follow conjunctive expressions such as *no de* 'because', *kara* 'because', *mono desu kara* 'because' vary depending on the conjunctive expression.

In a study of repair and projectability in Japanese and English, Fox *et al.* (1996:208-214) found that the elements in Japanese do not always form a coherent syntactic structure because S, O, and V are not always expressed, and the elements expressed "seem to be more independent from one another" than in English. English requires an overt subject as the "beginning" of its "tightly knit clause structure," and syntactic projection starts earlier in an utterance in English because the beginnings of TCUs project possible organizations for what is to follow. They conclude that Japanese participants are end-oriented, and "engage in syntactic practices which do not make easy 'early projection' strategies." Rather they use "wait and see" strategies, and projection is done bit-by-bit. "[T]he beginnings of TCUs in Japanese do not tend to have elements that syntactically project the possible organization of what is to follow."

Hayashi (1999) claimed that co-participant completion of Lerner's two-part compound TCU format is rare in Japanese. For example, after a *–tara* 'if' clause the co-participant rarely produces the equivalent of the 'then' clause in Japanese, rather the co-participant adds the final 1 or 2 words of the 'then' clause. In (2), the first speaker H continues after the *–tara* 'if' clause with the final component of the compound TCU and the co-participant M completes the final component with the final verb[1].

[1] TRANSCRIPTION NOTATION:

//	the part of the utterance after the // is overlapped by the next utterance
(0.5)	numbers inside ( ) give the length of a pause in tenths of seconds
:	colon indicates that the previous syllable is lengthened, number of colons reflects the amount of lengthening
?	rising intonation (not necessarily a question)
,	short pause, or continuing intonation
→	co-participant completion

The Japanese romanization follows that of Jorden with Noda (1987). Romanization of cited examples has been adjusted to maintain consistency.

(2) 1H *Okurahoma tte:: eego de hatuon-sitara:* Preliminary Component
Oklahoma QT English in pronounce-COND
If (you) pronounce *Oklahoma* in English,
2H *zenzen nihongo no Okurahoma to //ti-* Final Component
totally Japanese of Oklahoma from dif-
lit., totally totally from the Japanese *Okurahoma*, //(it)'s dif-
→ 3M *tyau nen na*
different FP FP
(it)'s different, isn't it.
(Hayashi 1999:480; boxes and annotation mine)

Hayashi concluded that English clauses have a "tightly-knit" structure, and a great deal of information is given about how to end a clause in the first part of the clause. In contrast, Japanese has a looser syntactic structure and "projection is done more bit-by-bit" than in English, and co-participants use "wait and see" strategies (Hayashi 1999:495).

It is interesting to note that the initial element of the final component (a simple sentence) in 4H in (2) is *zenzen* 'totally,' a modal adverb which strongly restricts the predicate. *Zenzen* 'totally' projects a negative predicate, in this case *tigau* 'it is different' at the end of the sentence. Previous research by Japanese grammarians can help explain this projectability.

Minami (1964, 1974, 1993, 1997) proposed a model for the syntactic structure of sentences in Japanese which is based on results from extensive analysis of the occurrence of non-predicate components and predicate elements in the internal structure of subordinate clauses (Table 1). This model unifies and confirms results of research by many Japanese grammarians from this unique perspective.

Minami's model is like an onion. When you cut through an onion, the outer layer of the onion and the next layer is not connected. However, if you cut through to the middle, you will reach a point where the layers will connect back up on the other side. Like an onion with 4 layers, a Japanese sentence begins with non-predicate components, which are ordered in layers D, C, B, A, and these layers are followed by the predicate elements ordered A, B, C, D, in the reverse order. (3) is a made-up sentence with components/elements in all 4 levels.

GLOSS ABBREVIATIONS: CAUS= causative, COHORT=cohortative, COND=conditional, COP=copula, DO=direct object, EVID=evidential, FP= final particle, GER=gerund, IO=indirect object, NEG=negative, NOM=nominalizer, PASS=passive, PF=perfective, POT=potential, PROG=progressive, QP= question particle, QT= quotative, SUB=subject, TENT=tentative, TOP=topic, V=verb

(3) *Nee, doo yara kinoo kanozyo ga Ken ni kaimono o sa-se -ta rasii ne.*
hey somehow yesterday she SUB Ken IO shopping DO do-CAUS-PF EVID FP
|*Ken ni kaimono o* **A** *sa-se*|
|*kinoo kanozyo ga* **B** *-ta*|
|*doo yara* **C** *rasii*|
|*Nee,* **D** *ne.*|
'Hey, somehow yesterday she made Ken do the shopping apparently, huh.'

The different levels are connected from the inside. The innermost level is the A level, the core semantic meaning of the sentence *Ken ni kaimono o sa-se* 'make Ken do shopping'. The next level out is the B level, in which temporal adverbs such as *kinoo* 'yesterday' and the subject *kanozyo* 'she' connect with the perfective ending of the verb. Then, in the C level the modal adverb *doo yara* 'somehow' connects with the evidential *rasii* 'seems, evidently'. The outer D level is the most interpersonal, e.g., the attention request *Nee* 'Hey' connects with the final particle *ne* 'huh'. In this paper, I will focus primarily on levels A, B, and C.

2. Analysis

In my analysis I will show how Minami's model can be used to analyze syntactic projectability in co-constructions of simple and complex sentences in Japanese. The data for this study consist of 35 examples of co-construction collected from natural conversations including everyday talk, company meetings, etc.

In the 20 examples of co-construction of simple sentences/clauses in my data, the co-participant used an A level component more often than a B or C level component to project the end of the speaker's utterance: A(13)>B(4)>C(3).

Examples in which an A level component was used to project the completion included i) NP *o* [direct object] + V, ii) NP *ni* [indirect object] + V as in (4), and iii) NP *made* 'until'/*e* 'to' + motion V as in (5).

(4) 1R *T-san wa nanka d(e)isizyon mitai no o itumo hito ni*
Mr. T TOP like decision like NOM DO always people IO
Mr. Tanaka like decision-like things always to people(the indirect object)
→ 2H *yudaneru*
entrust
(he) leaves it up to (them) (Ono & Yoshida 1996:118)

(5) 131K *zyaa, dokka sotti no oyazi no kaisya no hoo made,*
then somewhere there COP father of company of alternative up to
Then somewhere, as far as my father's company there
→ 132T *ittyaou.*
(I) guess (I)'ll go. (Sakuma, Sugito, & Hanzawa 1997: Data p.30)

There were 4 examples where the co-participant completed the speaker's utterance after a B level component; iv) NP *ga* [subject] + predicate as in (6)[2] and v) B level modal adverb + predicate as in (2) and (7) (first →).

(6) 1Y *Nanka* *Papa no iikata* *ga* *nee,*
somehow Papa of way of talking SUB you know
Somehow Papa's [your] way of talking, you know
→ 2A *moo* *iyami //da* *kara.*
really mean COP so
is really mean so. (Jones 1990:22-3)

Hata (1991) describes *tinzyutu hukusi* 'modal adverbs' as forms that "supplement and emphasize the modal meaning (negation, supposition, hypotheticality) of the predicate. *Tinzyutu* ... generally refers to the speaker's psychological state" (p.22; translation mine). Modal adverbs heighten projectability because they restrict the predicate. They have co-occurrence relations with negation (*kessite* ... *nai*[NEG] 'it never happens'), assertion/supposition (*kitto* ... *daroo* [TENT] 'for sure it must ...'), etc.

(7) 33S *Tada tan ni*
Just simply
→ 34A *dokka* *no syatyoo-san* *to ka* *sa, tada aisatu ni kite* *sa,*
someplace of chief or FP just greeting for come-GER FP
someplace's chief or you know, just comes for a greeting you know, and
→ 35S *sositara, so, misete,*
then right show-GER
then, right, (we/you) show (it) and (Kuwabara 1996:10)

Examples in which a C level component was used to project the completion included vi) NP *wa* [topic] + predicate as in (8)[3], and vii) C level modal adverb (*moshikasitara* 'perhaps') + evidential (*ka mo sirenai* 'maybe') as in (9).

(8) 1H *Soo yuu* *tyuuto hanpa* *na* *kanzi* *de*
that kind halfway COP feeling COP-GER
With that kind of halfway feeling
2H *koo yattyatta* *tte iu no ga* *boku wa*
this do completely-PF QT say NOM SUB I TOP
(his) ending up doing (things) like this, I, for one
→ 3S *ikenai* *to* *omou*
go-POT-NEG QT think
think *it's wrong*. (Ono & Yoshida 1996:125)

[2] The overall rarity of *NP ga* subjects in Japanese conversation suggests that subjects may not be the best thing on which to base a comparison of projection strategies.
[3] (8) is from a critical discussion about a married man who was flirting with a Japanese woman. The B level *ga* (SUB) component (dotted line) also projects the verb *ikenai* 'is wrong'.

(9) 100H *Soo suru to, kore, "siraseru" tte iu no mo, nanka, mosikasitara,*
then this inform QT say NOM also somehow perhaps
then, this, "informing" too, somehow, perhaps
101M *N.*
Uh huh.
102H *nanka, waketa hoo ga yokatta no.*
somehow divide-PF alternative SUB good-PF FP
somehow, it's that it would have been better to divide (it)
→ 103M *ka mo sirenai n da kedo //ne:.*
maybe NOM COP but FP
it's that maybe (it is so) but, you know. (Kuwabara 1995)

When an outer level component is used to project the end of a simple sentence/clause, the first speaker may add more components/elements before the co-participant completion. That is, the first speaker may add inner level components before the final projected one is reached as in (2) and (9).

Although Hayashi (1996) claimed that co-construction of compound TCUs is rare in Japanese, I found 15 examples in my data. Following Minami (1964:85, 1997:31), I will treat A level *–te* gerund clauses as predicate rather than sentence modifiers, and divide B level clauses into 3 types, B3, B2 and B1 (Table 1). The distribution of co-constructions of complex sentences in my data was: B3(0)<B2(2)<B1(5) <C(8).

According to Minami (1964:83-86, 1997:29-32), some B level subordinate clauses restrict while others do not restrict the predicate. For example, imperative and desiderative verb forms cannot occur after a *node* 'because' clause (Nagano 1952), and the past *-ta* form cannot occur after *-ba* 'if' or *-tara* 'if' clauses (simple condition). In contrast, sequential clauses ending in the *–te* (gerund) or verb stem do not restrict the predicate.

There were 7 examples of co-participant completion of a complex sentence beginning with a B level clause; 5 after a sequential *-te2* (GER) clause as in (7) (second →), and 2 after a *-tara* 'if, when' clause as in (10).

(10) 3H *modotte kuru no ka to omotTARA*
return-GER come NOM QP QT think-COND
when (I) thought (he) would come back
→ 4T *netyatta mitai*
end up falling asleep-PF seems
(it) seems that (he) ended up falling asleep.
(Ono & Yoshida 1996:123)

There were 8 examples of co-construction in which the co-participant completed a complex sentence beginning with a C level clause; 6 after a *kara* 'because' clause as in (11), and 2 after a *kedo(mo)* 'but' clause as in (12). Both (11) and (12) are from invitation conversations. The inviter completes a sentence beginning with the invitee's dispreferred response.

(11) 49R *dakara, osamaru <u>tokoro ni, hora, osame//rarenai KARA</u> nee.*
so put away place in you know put away-POT-NEG so you know
so, (I) can't put (things) away in the places they belong so, you know
→ 50N *A:a. Zyaa, <u>toobun wa isogasii wake da</u>.*
oh then a while TOP busy case COP
O:h. Then, (it)'s that (you) will be busy for a while, at least.
(Szatrowski 1993: Data p.18)

(12) 35A (0.5) *<u>ii KEDO</u> sa:,*
okay but FP
(it)'s okay but, you know,
→ 36B *<u>Boroboro?</u>*
(you)'re tired? (Szatrowski 1993: Data p.15)

3. Conclusion

I hypothesized that there would be more cases of co-construction when the first speaker used a component with higher syntactic projectability and more restriction on what would occur later in the utterance. However, my data suggested that the opposite was true. I summarize the results from my analysis of examples of co-construction of simple and complex sentences in my data using a continuum of semantic and syntactic projectability and restriction in co-constructions. The arrows indicate increasing frequency in my data.

In co-constructions of simple sentences, there were fewer examples with C, more with B, and the most with A level components. In addition, most of the B and A level components used for projection were core arguments with semantic connection with the verb, that is semantic projectability seemed stronger.

(13)

SIMPLE SENTENCE:	**A** (13)	←	**B** (4)	←	**C** (3)
	+ semantic projectability	<————————			– semantic projectability
	– syntactic projectability	<————————			+ syntactic projectability
	– restriction	<————————			+ restriction

In complex sentences, I found more examples with C than B level clauses. Again, co-participant completion seemed to be related more to semantic or pragmatic projectability.

(14)

COMPLEX SENTENCE:	**B**$_3$ (0)	→	**B**$_2$ (2)	→	**B**$_1$(5) →	**C** (8)
	– semantic projectability		————————>			+ semantic projectability
	+ syntactic projectability		————————>			– syntactic projectability
	+ restriction		————————>			– restriction

In conclusion, in co-constructions of both simple sentences as well as complex sentences in Japanese, semantic/pragmatic projectability seemed to be a stronger factor than syntactic projectability. This supports Fox *et al.*'s (1996) claim that syntactic projection is not very strong in Japanese. However, contrary to Fox *et al.*'s claim that the beginnings of TCUs in Japanese do not tend to syntactically project the possible organization of what is to come, I found that there were initial components from which the co-participant could project later elements in the sentence. For example, outer B and C level components such as modal adverbs not only project the syntactic category of the element to follow, but also have strong co-occurrence restrictions on the actual lexical item that can fill that slot.

In head-initial languages like English, the negative head precedes the verb phrase, and because it is not dependent on the negative polarity item for licensing, it does not project the negative polarity item. In contrast, in head-final languages like Japanese, the negative polarity item precedes the negative head, and because it is dependent on the negative head for licensing, it projects its licensor (e.g., negation). This difference in the position of the head in relation to its complement in Japanese and English allows for higher projectability of the actual item in Japanese. It is important to note that while initial core arguments in English may project the syntactic category, they do not necessarily project the actual lexical item to fill that slot. As in Japanese, there is a need to investigate further the role of semantic/pragmatic projectability in English co-constructions.

Finally, Fox *et al.* (1996) and Hayashi (1999) claimed that Japanese co-participants have to "wait to see." However, as the onion metaphor suggests, with co-constructions that begin with outer level components, they may also have to "wait to say," because highly projectable final elements may only come after less projectable inner level components and elements.

References

Fox, Barbara A., Makoto Hayashi, & Robert Jasperson. 1996. Resources and repair: a cross-linguistic study of syntax and repair. In E. Ochs, E. A. Schegloff & S. A. Thompson (eds.) *Interaction and grammar*, 185-137. Cambridge: Cambridge University Press.

Hata, Iku. 1991. Dai-itibu hukusiron no keehu (Chapter 1: Lineage of adverb theory). *Nihongo kyooiku sidoo sankoosyo* 19, 1-46. Tokyo: National Language Research Institute, Ookurasyoo.

Hayashi, Makoto. 1999. Where grammar and interaction meet: A study of co-participant completion in Japanese conversation. *Human Studies* 22:375-499.

Jones, Kimberly. 1990. *Conflict in Japanese conversation*. PhD dissertation, University of Michigan.

Jorden, Eleanor Harz with Mari Noda. 1987. *Japanese: the spoken language*. New Haven: Yale University Press.

Kuwabara, Kazuko. 1995. *Teean no koozoo to sutoratezii no koosatu* (Investigation of the structure of proposals and strategies). Japan Women's University, MA Thesis.

Kuwabara, Kazuko. 1996. Nihongo no "teean" no danwa no koozoo bunseki (Structural analysis of Japanese proposal discourse). *Nihon Zyosi Daigaku Daigakuin Bungaku Kenkyuuka Kiyoo* 2:1-12.

Lerner, Gene H. 1991. On the syntax of sentences-in-progress. *Language in Society* 20:441-458.

Minami, Fujio. 1964. Hukubun (Complex sentences). *Kooza gendaigo 6 Koogo bunpoo no mondaiten*, 71-89. Tokyo: Meezi Syoin.

Minami, Fujio. 1974. *Gendai nihongo no koozoo* (The structure of modern Japanese). Tokyo: Taisyuukan Syoten.

Minami, Fujio. 1993. *Gendai nihongo bunpoo no rinkaku* (Outline of the grammar of modern Japanese). Tokyo: Taisyuukan Syoten.

Minami, Fujio. 1997. *Gendai nihongo kenkyuu* (Studies on modern Japanese). Tokyo: Sanseedoo.

Mizutani, Nobuko. 1993. Kyoowa kara taiwa e (From co-constructed talk to dialogic talk). *Nihongogaku* 12:4:4-10.

Nagano, Masaru. 1952. "Kara" to "no de" to wa doo tigau ka (How do *kara* 'because' and *no de* 'because' differ?). *Kokugo to Kokugogaku* 334.

Ono, Tsuyoshi & Eri Yoshida. 1996. A study of co-construction in Japanese: We don't "finish each other's sentences." In N. Akatsuka, S. Iwasaki & S. Strauss (eds.) *Japanese/Korean linguistics* 5, 115-130. Stanford: CSLI.

Sacks, Harvey, Emanuel Schegloff, & Gail Jefferson. 1974. A simplest systematics for the organization of turn-taking in conversation. *Language* 50:4:696-735.

Sakuma, Mayumi, Seiju Sugito, & Kanichi Hanzawa. 1997. *Nihongo no bunsyoo/danwa* (Japanese discourse/conversation). Tokyo: Oohuu.

Szatrowski, Polly. 1993. *Nihongo no danwa no koozoo bunseki- kanyuu no sutoratezii no koosatu-* (Structure of Japanese conversation- invitation strategies). Tokyo: Kurosio.

Szatrowski, Polly. 2000. Kyoodoo hatuwa ni okeru sankasya no tatiba to gengo/higengo koodoo no kanren ni tuite (Relation between participant status and verbal/nonverbal behavior in co-construction). *Nihongo Kagaku* (Japanese Linguistics) 7:4:44-69.

Szatrowski, Polly. (to appear) When do Japanese speakers co-construct utterances? In M. Noda & P. Wetzel (eds.) *Language in culture*. Columbus, OH: Ohio State University Press.

University of Minnesota- ILES
215 Nolte Center
Minneapolis, MN 55455

szatr001@tc.umn.edu

Table 1: Components and Elements of Subordinate Clauses (Minami 1993:96-97, Minami 1964:85; bold lines and roman numerals on top added by Szatrowski)[4]

SYNTACTIC ELEMENTS	CLAUSES	A -*nagara* 'while'	-*tutu* 'while'	-*te*$_1$ GER (manner)	repetition of V-stem	B -*te*$_2$ 'and' (sequence)	*to* 'when'	-*nagara* 'although'	*node* 'because'	*noni* 'although'	-*ba* 'if'	-*tara* 'if'	*nara* 'if'	-*te*$_3$ 'and' (reason)	V-stem	-*zuni* 'without'	-*naide* 'without'	C *ga* 'but'	*kara* 'because'	*keredo* 'but'	*si* 'and'	-*te*$_4$ GER
PREDICATE COMPONENTS	-*mai*, *daroo* (TENT), -*u*/-*yoo* (COHORT)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	+	+	+	–
	-*ta*/*da* tense/aspect	–	–	–	–	–	–	–	+	+	–	–	+	–	–	–	–	+	+	+	+	–
	-*nai* (NEG)	–	–	–	–	+	+	+	+	+	+	+	+	+	–	–	–	+	+	+	+	+
	-*masu* (DISTAL)	–	–	–	–	+	+		+	+	+	+	+	+	–	(+)	–	+	+	+	+	+
	Polite form	+	+	+	–	+	+	+	+	–	+	+	+	+	+	+	+	+	+	+	+	+
	Giving/receiving Verb	+	+	+	–	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	-(*r*)*areru* (PASS)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	-(*s*)*aseru* (CAUS)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	Verb	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
NON-PREDICATE COMPONENTS i	NP *o* (DO)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
ii	NP *ni* (IO)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
iii	NP + case particle	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	State adverb	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	Degree adverb	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	A level clause	+	+	+	(+)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
iv	NP *ga* (SUBJECT)	–	–	–	–	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	Place modifier	–	–	–	–	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	Time modifier	–	–	–	–	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
v	*tonikaku* 'anyway', *yahari* 'afterall'	–	–	–	–	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	Evaluative modifier	–	–	–	–	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	B level clause	–	–	–	–	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
vi	NP *wa* (TOPIC)	–	–	–	–	–	–	–	(+)	–	–	–	–	–	–	–	–	+	+	+	+	+
vii	*tabun* 'probably', *masaka* 'no way'	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	+	+	+	+
	C level clause	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	+	+	+	+
						B$_1$	B$_3$	B$_3$	B$_3$	B$_3$	B$_2$	B$_2$	B$_2$	B$_3$	B$_1$	B$_1$	B$_1$					

[4] I have indicated the 3 levels A, B, and C with bold lines in this table. Minami divides sentences into the non-predicate components and predicate elements given across the top of the table. He determined the level for each of the different components and elements based on their possible occurrence in the clauses given on the left. A '+' sign indicates that a component/ element can occur, and a '–' sign that it cannot.

Against Richness of the Base: Evidence from Nganasan*

OLGA VAYSMAN
Massachusetts Institute of Technology

0. Introduction

Since Optimality Theory is a highly output-oriented grammatical theory, the strongest hypothesis is that all systematic, language-particular patterns are the result of output constraints, and that there is no other place from which such patterns can derive. In particular, input is not a level of derivation that can be constrained. This principle is known as Richness of the Base hypothesis, and it states that there are no constraints on the input structure of words, and that all linguistic constraints are statements on the surface structure only. In other words, Richness of the Base attributes all systematic phonological patterns to constraint rankings, not to difference in inputs.

For example, the fact that no English words end in [h] cannot derive from a restriction on English lexicon forbidding [h]-final morphemes. Rather, it must be the case that English grammar **forces** all its outputs to obey the prohibition on final [h]. It means that even if there were an h-final lexical entry in English, which would provide an [h]-final input, the corresponding output of the English grammar would never be [h]-final. Therefore, the absence of [h]-final words in English must be explained within OT by a grammar – a ranking of constraints – with the property that no matter what the input, the output of the grammar will not be [h]-final. In particular, the OT analysis of English must consider a set of inputs to the grammar (the base) that is as rich as possible: the base consists of all universally possible inputs, including those that are [h]-final, those that contain lateral fricatives, those that have lexically marked tone and so on.

In this paper, I consider some consonant gradation facts from a Uralic Samoyedic language Nganasan, and argue that (at least the strict interpretation of) the Richness of the Base hypothesis runs into problems when we deal with full range of relevant data from this language, namely isolated words, compounds, and borrowings.

* Many thanks are due to the audiences of BLS 28 and NAPhC 2 for their comments, suggestions, questions and objections. Thanks to Cheryl Zoll and Michael Kenstowicz for support and discussion.

1. Principles of Consonant Gradation[1]

Let us begin by defining what consonant gradation is, and how it works in this particular language. The grade alternation, or gradation, of consonants is a phenomenon of alternative appearance of two grades, traditionally called strong and weak, depending on some phonological or morphophonological environment. For Nganasan, consonant grade alternations are alternations between voiceless (strong in traditional terminology) and voiced ("weak") obstruents. The reflexes of gradation are given in table in (1) below:

(1) Gradation reflexes (cf. Helimsky 1998, Tereshenko 1968a)[2]

strong grade (voiceless)	h	t	k	s	ç	c
weak grade (voiced)	b	ð	g	ɟ	ɟ	ɟ

As we can see from the table, this alternation is basically obstruent voicing alternation. For the purposes of this paper, I will disregard other aspects of the alternations in question, such as place and manner of articulation of the reflexes, and concentrate on the alternation between the grades, so I will refer to the voiceless and voiced series of the consonants.

1.1. Voicing of Intervocalic Obstruents

We will now see how principles of Consonant Gradation are responsible for the distribution of obstruent voicing in the language. Intervocalically, obstruents are voiced if they are foot-initial, and voiceless if they are inside a foot. The data in (2) below, which use the 3rd person possessive singular suffix, illustrate this distribution. In the first word, *nï-tï* 'his/her/its wife', the suffix-initial consonant is inside a foot, and it shows up as voiceless [t]. The same is true for the words for 'his/her/its salmon', 'his/her/its caviar' and 'his/her/its tear' and others: the suffix-initial consonant surfaces as voiceless [t] when it is foot-internal.

Contrast this, for example, with the word *jütü-ðü* 'his/her/its hand', where the stem itself constitutes a foot. In this word, the suffix-initial consonant is foot-initial, and it appears in its "weak" grade, i.e. as voiced [ð]. The same generalization can be seen in words *suuðəə-ðu* 'his/her/its lung', *məku-ðu* 'his/her/its spine', the word for *taa-ðu* 'his/her/its deer', and other words where the consonant of the suffix starts a foot.

[1] Nganasan, also known as Tawgy or Tawgy Samoyed, is a Uralic Samoyedic language, spoken by the north most ethnic group in Eurasia in Taimyr peninsula. About 1300 speakers (1989). The data in this talk is partly taken from source grammars (Helimsky 1998, Tereshenko 1979, Prokofjev 1937) and subsequently checked with native speakers, and partly comes from field work on the language in March 2000 and October 2000. Most of the discrepancies between grammars and my field work are noted.

[2] Note on Transcription: IPA transcription for consonants, so palatal consonants will be transcribed as [c], [ɟ], [ç], etc., Prenasalized consonants are transcribed with a nasal homorganic superscript nasal, e.g. [ⁿt] or [ᵑg]. Vowels: symbols traditional for Finnic linguistics: [i], [e], [ü] are front vowels; and [a], [u], [o], [ə] and [ï] are non-front.

(2) Intervocalic [t]/[ð] gradation, 3rd person possessive suffix –tV/-ðV

(nï-**tï**) 'his/her/its wife'	(tïri)(mi-**ti**) 'his/her/its caviar'
(ɟütü)-(ðü) 'his/her/its hand'	(taa)-(ðu) 'his/her/its deer'
(suu)(ðəə)-(ðu) 'his/her/its lung'	(cimi)-(ði) 'his/her/its tooth'
(məku)-(ðu) 'his/her/its spine'	(ɲinï)-(ðï) 'his/her/its older brother'
(bïnï)-(ðu) 'his/her/its rope'	(baku)(nu-**tu**) 'his/her/its salmon'
(kəə)(lï-**tï**) 'his/her/its tear'	(kəi)-(ði) 'his/her/its side'
(sela)-(ðu) 'his/her/its fat'	(sürü)-(ðü) 'his/her/its winter'
(ɟalï)-(ðï) 'his/her/its day'	(ɲoru)(mu-**tu**) 'his/her/its copper'
(ŋuhu)-(ðu) 'his/her/its mitten'	(ci^{n}te)-(ði) 'his/her/its hill'
(ɟama)(ða-**tu**) 'his/her/its animal'	(moo)(ɟa-**tu**) 'his/her/its stub'

The distribution of voicing is the same whether the gradating consonant starts a suffix or is inside a suffix. In (3) there is another example of the same voicing distribution with a similative suffix, where the gradating consonant ([k]/[g]) is in the second syllable of the suffix. The generalization about voicing distribution still holds: when the obstruent of the suffix is foot-initial, it is voiced, and when it is foot-internal, it is voiceless.

(3) Intervocalic [k]/[g] gradation, similative suffix -rəkï/-rəgï

(nï-rə)(**gï**) 'similar to a woman'	(ɟütü)-(rə**kï**) 'similar to a hand'
(bïnï)-(rə**kï**) 'similar to a rope'	(suu)(ðəə)-(rə**kï**) 'similar to a lung'
(taa)-(rə**kï**) 'similar to a deer'	(baku)(nu-rə)(**gï**) 'similar to salmon'
(tïri)(mi-rə)(**gï**) 'similar to caviar'	(ɲinï)-(rə**kï**) 'similar to a brother'
(kəə)(lï-rə)(**gï**) 'similar to a tear'	(kəi)-(rə**kï**) 'similar to a side'
(ɟama)(ða-rə)(**gï**) 'similar to an animal'	(ŋuhu)-(rə**kï**) 'similar to a mitten'

The data in (4) below gives us yet another example of intervocalic obstruent voicing distribution in verbs rather than nouns that we saw before, with a participial suffix with [s]/ [ɟ] alternation. Despite the morphological (nouns vs. verbs) difference, the generalization about consonant gradation is exactly the same: when an intervocalic [s] is foot initial, it voices, and when it is foot-internal it stays voiceless.

(4) Intervocalic [s]/[ɟ] gradation, participial suffix -ɟï/-sï

(bïtï)-(ɟi) 'drink'	(buaʔ)(tə-sï) 'look up'
(ɟilə)-(ɟi) 'lift'	(bə)(ðuaʔ)(tə-sa) 'grow'
(ɟorə)-(ɟa) 'cry'	(ho)(ðəʔ)(tə-sa) 'write'
(hotə)-(ɟa) 'write out'	(ɟorə)(lə-sa) 'start crying'
(bï)(ðïr) (nä^{n}tï)-(ɟi) 'be thirsty, want to drink'	(bï)(ðïp)(tï-sï) 'give to drink'

1.2. Voicing of postconsonantal and coda obstruents

We will now examine the distribution of voicing in obstruents that are not intervocalic. Nganasan obstruents after another consonant are always voiceless. In

(5), you see the same suffix as in (2), the 3rd person possessive suffix, but its obstruent always surfaces as voiceless [t] after consonant-final stems below, regardless of whether this obstruent is foot-initial or foot-internal:

(5) Postconsonantal (always voiceless), 3rd person possessive suffix –tV/-ðV

(tər-**tu**) 'his/her/its hair'	(kam-**tu**) 'his/her/its blood'
(ka)(ðar)-(**tu**) 'his/her/its light'	(maʔ-**tu**) 'his/her/its house'
(bïʔ-**tu**) 'his/her/its water'	(sïr-**tu**) 'his/her/its ice'
(he)(ⁿɟir)-(**tu**) 'his/her/its shaman's drum'	(ni)(luʔ)-(**tu**) 'his/her/its life'
(huəʔ)-(**tu**) 'his/her/its fur overcoat'	(so)(ŋïl)-(**tu**) 'his/her/its pillow'
(ðö)(ruʔ)-(**tu**) 'his/her/its cry'	(baŋ-**tu**) 'his/her/its dog'

The same is true of coda obstruents: they are always voiceless in the language. In addition to appearing in their "strong" (i.e. voiceless) grade, coda obstruents in the language are neutralized to glottal stop[3]. The underlying place of articulation, however, is clear in forms where the obstruent is intervocalic, like the Accusative singular forms below:

(6) Coda Obstruents (always voiceless, place neutralization to glottal stop)

bïʔ (cf. Acc. Sg. *bïðïm* with suffix *–m* and epenthesis)	'water'
huəʔ (cf. Acc. Sg. *hiəɟim* with suffix *–m* and epenthesis)	'fur overcoat'
maʔ (cf. Acc. Sg. *maɟəm* with suffix *–m* and epenthesis)	'house'
niluʔ (cf. Acc. Sg. *niluɟim* with suffix *–m* and epenthesis)	'life'
ðöruʔ (cf. Acc. Sg. *ðörugum* with suffix *–m* and epenthesis)	'cry'
keguʔ (cf. Acc. Sg. *kekuðum* with suffix *–m* and epenthesis)	'fog'
sïnüʔ (cf. Acc. Sg. *sïnügüm* with suffix *–m* and epenthesis)	'memory'
hiaʔ (cf. Acc. Sg. *hiaɟim* with suffix *–m* and epenthesis)	'oak'

The two non-intervocalic positions, therefore, are never contrastive as far as obstruent voicing is concerned. The obstruents in postconsonantal and coda positions are voiceless.

1.3. Obstruent Voicing in Word-Initial Position

The only position where obstruent voicing is contrastive is word-initial position. Word-initial voicing of the obstruents is not predictable from any principle of Consonant Gradation (or any other principle, as far as I can see). The data in (7) illustrates this point: a word can start with either a voiced or a voiceless obstruent, even though all the word-initial obstruents are also, obviously, foot-initial. With that we have now seen the distribution of voicing of obstruents in all positions; the table in (8) summarizes the distribution.

[3] Interestingly enough, the only obstruents that do not get neutralized to glottal stop in the coda are labials. They are also invariably voiceless in the coda position and therefore comprise the alternation [$h^{(w)}$] ~ [b] ~ [p].

(7) Word-initial (contrastive voicing)

ɟühü 'sledge'	**ɟ**äte 'stone'	**k**ita 'cup'	**k**asu 'bark'
cehï 'nail'	**ç**iəɟə 'tongue'	**g**ətə 'swan'	**g**ula 'crow'
satu 'clay'	**s**ïr 'ice'	**t**ər 'hair'	**t**urku 'lake'
baŋ 'dog'	**b**asa 'iron'	**ð**öruʔ 'cry'	**ð**ajpu 'dry wood'
h^{u}aa 'tree'	**h**oðür 'letter'		

(8) Summary of voicing of obstruents

Word Position / Prosodic Position	Intervocalic	Post consonantal	Coda	Word-Initial
Foot-initial	vcd	vls	——	vcd/vls
Foot-internal	vls	vls	vls	——

As table in (8) illustrates, obstruent voicing in all but one word/prosodic position is predictable. It seems reasonable to analyze the pattern of Nganasan consonant gradation as intervocalic lenition that is blocked in case an intervocalic obstruent is foot-internal. Obstruent voicing in the language is in complementary distribution everywhere but in word-initial position. Importantly, that position is **the only one that has contrastive voicing of obstruents**.

2. No Suffix-Root Asymmetry

It is also important that the lenition is crucially not a suffix versus root asymmetry, where consonants in suffixes would gradate (lenite) and the ones in the roots would not. In (9) we have an illustration of the same type of consonant alternations we saw before, but within roots. These are the examples which show the alternation between Nominative singular and plural.

(9) Root-internal Grade Alternations – Nominative plural suffix

	Nom.sg.	Nom.pl.	Gloss
a.	(kuhu)	(ku)(buʔ)	'skin, hide'
	(basa)	(ba)(ɟaʔ)	'iron'
	(ke^{n}te)	(ke)(^{n}deʔ)	'sledge'
	(ŋuta)	(ŋu)(ðaʔ)	'berry'
	(məku)	(mə)(gaʔ)	'back'
	(ɟakə)	(ɟa)(güʔ)	'twin'
b.	(ka)(ðar)	(kata)(rəʔ)	'light'
	(he)(^{ɲ}ɟir)	(he^{n}sï)(rəʔ)	'shaman's drum'
	(bïʔ)	(bï)(ðïʔ)	'water'
	(huəʔ)	(hiə)(ɟiʔ)	'fur overcoat'
	(ci)(ɟar)	(cisa)(rəʔ)	'benefit'
	(hä)(ɟir)	(häsi)(rəʔ)	'fish-hook'
	(ho)(^{n}ɟir)	(ho^{n}çi)(rəʔ)	'edge, side'

The only difference between Nominative singular and plural in examples in (9a) is that in the plural, a glottal stop is added, thus closing the final syllable. The stems in (9b) have the final syllable closed in the singular (stems of this type are consonant-final stems). When a plural ending, the glottal stop, is added, the resulting complex coda is broken by an epenthetic vowel. The epenthesis changes the foot structure: singular *kaðar*, for example, shows up as *katarəʔ* in the plural.

Another example of the same type is in (10). It is an alternation between a noun and an adjective derived from it. The suffix that is added to a noun to create an adjective is a single vowel that takes on the features of the preceding vowel because of vowel harmony. When we add this derivational suffix, the prosodic structure is, again, changed. With vowel-final stems, we get a final CVV foot, and with consonant-final stems, it adds an extra syllable.

(10) Root-internal Grade Alternations – Denominal Adjective suffix –V

Noun	**Adjective**	**Gloss**
(basa)	(ba)(ɟaa)	'iron (noun) – iron (adj)'
(ka)(ðar)	(kata)(ra)	'light (noun) – light (adj)'
(bïʔ)	(bï)(ðïï)	'water – watery, wet'
(ci)(ɟar)	(cisa)(rə)	'benefit – beneficial'
(satu)	(sa)(ðuu)	'clay (noun) – clay (adj)'
(ɟäte)	(ɟä)(ðeə)	'stone (noun) – stone (adj)'
(ni)(luʔ)	(nilu)(ɟu)	'life – lively'
(maʔ)	(masa)	'house – domesticated'

The result is that the distribution of voicing of root obstruents is exactly as the distribution of voicing of suffix obstruents: they lenite if they are foot initial and intervocalic, and do not lenite otherwise. **The local conclusion from these alternations** is that Intervocalic Lenition applies the same way in roots as it does in suffixes: it is restricted by metrical structure, but not by morphological boundaries or morpheme identity (root versus suffix).

3. Two Basic Models of Complementary Distribution

Our observation so far is that the only position in the language where we do find **contrastive voicing is word-initial** position, in every other position the obstruent voicing is in complementary distribution. Logically, complementary distribution can be modeled in two distinct ways, regardless of the framework:

Under the possible first model, principles of consonant gradation require that 1) postconsonantal, coda and foot-internal consonants are voiceless (Fortition), and 2) intervocalic obstruents are voiced (Lenition). In this case, Consonant Gradation would override whatever underlying specifications obstruents have, leaving only word-initial consonants unaffected. Under this approach, there is no need to restrict underlying representations. Whatever the underlying voicing of the obstruents, they surface as voiced or voiceless according to the principles of consonant gradation, and only word-initial obstruents will surface with their underlying voicing, because no constraint requires otherwise.

Under another possible analysis, on the other hand, the consonant gradation principle only requires that intervocalic obstruents are voiced (this principle will, of course, be violated if the obstruent is not foot-initial). Consonant gradation adds [+voice] to intervocalic foot-initial obstruents, and does not affect obstruent voicing in any other position. Underspecified obstruents surface as voiceless, and specified word-initial obstruents surface with their underlying voicing. Notice that if this analysis is the right one, crucially there has to be a constraint on underlying representations, namely that **only word-initial obstruents have underlying specification for voicing.**

4. Compounds

Let us now consider another set of data, compounds, that helps us to decide which of the two possible basic analyses is correct. Now, why is this data so important? When a compound is formed, an obstruent that is word-initial (prespecified for voicing) in the lexicon becomes word-internal, which puts it in the right position to be affected by Consonant Gradation. If the first analysis is correct, and consonant gradation overrides any voicing prespecifications, the first obstruent of the second part of the compound should surface as either voiced or voiceless depending on its position in prosody (i.e. voiced if it is both foot-initial and intervocalic, and voiceless otherwise).

If the second analysis is correct, and the principles of consonant gradation operate only on obstruents that are not specified for voicing underlyingly, the initial obstruent of the second root of the compound should surface with its underlying voicing specification, regardless of its position in prosody and whether or not it is intervocalic.

It is important for our task to make sure that the initial obstruent of the second root can be in different prosodic positions, i.e. that the whole compound is one phonological word. Indeed, it seems to be the case that foot boundaries do not have to coincide with edges of the roots in the language: the footing is continuous throughout the compound. The evidence for analyzing a compound as one phonological word comes from two sources: from consonant gradation itself and from the pattern of stress assignment.

4.1. Compounds as One Phonological Word

The first piece of crucial data showing that compounds should be analyzed as one phonological word is in (11). In the left column, the feet are assigned continuously from left to right, not respecting the morpheme boundaries. This footing gives us the correct reflex of the gradating consonants. The indicators are not, of course, root-initial consonants but the medial consonants of the second roots which we know adhere to the principles of consonant gradation. Thus, the word *kobaɟa* 'earring' (which is a compound of the word *ko* that means 'ear' and the word *basa* which means 'iron') shows the "weak", i.e. lenited grade of the consonant *s*, indicates that this consonant is foot-initial. In the right column of (11) we see that if we were to start the footing anew with the second root, we would make the wrong predictions for the voicing of this consonant: the compound would surface as *kobasa*, with the non-lenited strong *s*. Therefore, reflexes of gradation of consonants inside the second root of the compounds establish that a compound is one phonological word.

(11) Compounds (one phonological word, footing continuous)

(cïïp)(sinə)(ðəba)(ɟa) 'bracelet' (cïïpsinəðə 'wrist' + basa 'iron')	*(cïïp)(sinə)(ðə)(ba**s**a)
(hïa)(ɟəki)(ða) 'glove' (hïaɟə'thumb' + kita 'cup')	*(hïa)(ɟə)(ki**t**a)
(koba)(ɟa) 'earring' (ko 'ear' + basa 'iron')	*(ko)(ba**s**a)
(ɟama)(ðaŋo)(ga) 'insect' (ɟamaða 'animal' + ŋoka 'numerous')	*(ɟama)(ða)(ŋo**k**a)
(bïʔɟü)(bü) 'boat' (bïʔ 'water' + ɟühü 'sledge')	*(bïʔ)(ɟü**h**ü)

This conclusion is supported by stress pattern of compounds compared with the stress pattern of words when they are not compounded. The examples in (12) illustrate the stress assignment pattern of the language. Primary stress is marked with double underlining, and secondary with single underlining. The first syllable of the words always receives primary stress, and 3rd, 5th and so on vocalic moras receive secondary stress.

(12) Stress pattern in words in isolation (non-compounds) [4]
(cïïp)(sinə)(ðə) 'wrist'
(basa) 'iron'
(hïa)(ɟə) 'thumb'
(kita) 'cup'
(ɟama)(ða) 'animal'
(ŋoka) 'numerous'

When a compound is formed, the stress is assigned exactly as if a compound is a monomorphemic word, i.e. there is only one primary stress, and, even more

4 The stress pattern presented here (trochaic with main stress on the first foot) is the pattern of the Vadey dialect of Nganasan, and is historically original. The other dialect of the language, Avam, has a more complicated pattern of stress assignment: main stress is on the penultimate mora, with some retractions onto the antepenult from a schwa or [ï], provided that the antepenultimate vowel is not itself a schwa. Secondary stress is trochaic, with feet assigned from left to right. That the stress assignment is continuous throughout the compound, however, is true of both dialects.

importantly, the initial vowel of the second root of the compounds does not have to be stressed at all. It receives stress only if the first root of the compound happens to contain an even number of moras.

Again, in the right-hand column we see that if we were to assign stress to the two parts of the compound separately, we would get the wrong predictions for stress, as well as for the reflexes of the gradation.

(13) Stress pattern in compounds (continuous stress assignment throughout)

(cïïp)(sinə)(ðəba)(ɟa) 'bracelet' (cïïpsinəðə 'wrist' + basa 'iron')	*(cïïp)(sinə)(ðə)(basa)
(hïa)(ɟəki)(ða) 'glove' (hïaɟə'thumb' + kita 'cup')	*(hïa)(ɟə)(kita)
(koba)(ɟa) 'earring' (ko 'ear' + basa 'iron')	*(ko)(basa)
(ɟama)(ðaŋo)(ga) 'insect' (ɟamaða 'animal' + ŋoka 'numerous')	*(ɟama)(ða)(ŋoka)
(bïʔɟü)(bü) 'boat' (bïʔ 'water' + ɟühü 'sledge')	*(bïʔ)(ɟühü)

Both of these sets of data indicate that a compound is treated as one phonological word, and that we can look at compounds to see how obstruents that are word-initial in the lexicon behave inside phonological words on the surface. The first consonant of the second root of the compound, therefore, can be found both foot-initially and foot-internally.

The compounds in (14) suggest that the voicing of the consonant that we are interested in does not depend on any principles of consonant gradation. It surfaces with underlying voicing regardless of its prosodic position. In the word for 'alcohol' which is a compound of a word *bïtï* 'drink' and the word *kosü* meaning 'dry', the second root of the compound surfaces as voiceless [k] rather than voiced [g], even though it is both intervocalic and foot-initial. In the word *bïʔɟübü* 'boat' (compound of *bïʔ* 'water' and *ɟühü* 'sledge'), the first obstruent of the second root appears after a consonant (glottal stop), and it is foot-internal. Yet, it does **not** get devoiced, and surfaces as palatal voiced stop [ɟ]. It does not get devoiced, even though it is not only foot-internal but also postconsonantal. And, of course, both of these positions are the positions for voiceless, or "strong" grade of obstruents. Apparently, therefore, there is no fortition condition that overrides the underlying voicing specification of this consonant.

(14) Compounds (underlying voicing preserved)

(ɟütü)(basa) 'ring'	
(ɟütü 'hand' + basa 'iron')	
(hᵘaa)(ɟütü) 'branch'	
(hᵘaa 'tree' + ɟütü 'hand')	
(bïʔɟü)(bü) 'boat'	*(bïʔsü)(bü)
(bïʔ 'water' + ɟühü 'sledge')	
(basa)(cimi) 'shovel'	*(basa)(ɟimi)
(basa 'iron' + cimi 'tooth')	
(bïtï)(kosü) 'alcohol'	*(bïtï)(**g**osü)
(bïtï 'drink' + kosü 'dry')	
(hiᵑhi)(təə)(bu) 'dawn'	*(hiᵑhi)(ðəə)(bu)
(hiᵑhi 'night' + təəbu 'tail')	
(koʎ)(süo)(kou)(ðaj)(kaa) 'snake'	*(koʎ)(süo)(**g**ou)(ðaj)(kaa)
(koʎsüo 'worm'+ kouðajkaa 'long')	
(ŋoᵑhu)(ʔaɟep)(tï) 'badmouth, gossip'	*(ŋoᵑhu)(ʔasep)(tï)
(ŋoᵑhuʔa 'bad'+ ɟeptï 'lip')	

The bottom line is that the language **can** have a voiced obstruent foot-internally even when it is intervocalic, but only if this voicing is also present in the input. However, there are no words, outside the compounds, that have specification for voicing anywhere but word-initially. These facts suggest that consonant gradation operates only on obstruents that are underspecified for voicing underlyingly, and **cannot override voicing specifications**. The second analysis of the phenomenon is correct. The following **generalization** about underlying representations becomes apparent: all and only word-initial consonants in the language are specified for voicing underlyingly.

5. Borrowings

What does this pattern tell us about the Richness of the Base Hypothesis? It is clear that we have a generalization about underlying representations, but we also have to make sure that the language **cannot** have inputs without the first obstruent specified for voicing (and inputs with voicing specifications anywhere but on the word-initial obstruent), not just that the language **does not** have such inputs. In other words, we have to show that this principle is an active **constraint on underlying representation**, rather than merely a coincidence or historical epiphenomenon. To do that, we will now consider borrowings from Russian, which are a potential source of voicing specifications word-internally.

Examples of the borrowings are in (15) below. In the Russian word for "weather" (which means "bad weather" in Nganasan), there are two voiced obstruents, one of which, [g], is foot-internal. It is borrowed into Nganasan as *pokoða*, with the foot-internal obstruent as voiceless [k]. The voicing of the word-initial obstruent of the source word is preserved. The same is true for all the obstruents in the borrowings: obstruent voicing of the source language is preserved always and only word-initially, and the voicing of all word-internal obstruents depends on principles internal to Nganasan, in no way reflecting the voicing in Russian words.

(15) Borrowings (contrastive voicing only word-initially)

ðoroga 'paved road' from Russian /doroga/ 'road'	
pokoða 'bad weather' from Russian /pogoda/ 'weather'	*pogoða
hapriga 'factory' from Russian /fabrika/ 'factory'	*ha**b**ri**k**a
cemnoða 'darkness' from Russian /tʲemnota/ 'darkness'	*cemno**t**a
bolka 'Volga' (proper name) from Russian /volga/ 'Volga'	*bol**g**a
kəɲigülï 'holidays' from Russian /kanʲikuli/ 'holidays'	* kəɲi**k**ülï
səltaʔ 'soldier, police' from Russian /soldat/ 'soldier'	* səl**d**aʔ
çiroða 'orphan' from Russian /sʲirota/ 'orphan'	*çiro**t**a
səhaga 'little dog, show-dog' from Russian /sobaka/ 'dog'	* səba**k**a

We already know from the compound data that the language does not change the underlying voiced obstruents into voiced on the surface, but rather only regulates the voicing of obstruents that are underspecified for voicing in the input, leaving the specified obstruents with their underlying voicing. The voicing of word-internal obstruents in the borrowings, however, is predictable and not contrastive.

It follows, therefore, that the borrowings' underlying representations (and crucially **not** the surface representations) are adjusted to "keep" the obstruent voicing of the source language only word-initially, and to lack voicing prespecifications in all other positions.

6. Conclusion

The natural conclusion is that the language restricts its vocabulary. Nganasan inputs **have to be constrained** in at least one way: all and only word-initial obstruents must be specified for voicing underlyingly. Sets of data from compounding and borrowings, taken together, show us that the constraint on contrastive voicing is an active requirement on underlying representations. This data and analysis provide a clear counterexample to the Richness of the Base hypothesis.

References

Anttila, Arto 1995. Deriving variation from grammar: A study of Finnish genitives. ROA-63.

Collinder, Björn 1965. *An Introduction to Uralic Languages*. University of California Press, Berkeley and Los Angeles.

Collinder, Björn 1959. *Comparative Grammar of the Uralic Languages*. Stokholm: Almqvist & Wiksell.

Comrie, Bernard 1981. *The Languages of the Soviet Union.* Cambridge University Press.

Helimsky, Eugene 1998. Nganasan. In Daniel Abondolo (ed.) *The Uralic languages*. London, New York: Routledge.

Inkelas, Sharon and Young-mee Yu Cho 1993. Inalterability as Prespecification. *Language*, **69**(3): 529-575.

Kager, Rene 1995. On affix allomorphy and syllable counting. ROA-88.

Prince, Alan 1983. Relating to the grid. *LI,* **4**:19-100.

Prince, Alan and Paul Smolensky 1993. Optimality Theory: Constraint Interaction in Generative Grammar. Ms., Rutgers University, New Brunswick and University of Colorado, Boulder.

Prokofjev, G. 1937. [Прокофьев Г.Н.] Нганасанский (тавгийский) диалект. – Языки и письменность народов СССР, ч. 1, Москва.

Setälä, Emil Nestor 1912. Uber Art, Umfang und Alter des Stufenwechsels im Finnisch-Ugrischen und Samojedischen. – Finnisch-Ugrische Forschungen XII Anz., 1-128.

Smolensky, Paul 1995. "On the internal structure of Constraint Component of UG". Handout of a talk given at UCLA, April 7, 1995 ROA 86-0000.

Smolensky, Paul 1996. The initial state and 'Richness of the Base' in Optimality Theory. Technical Report JHU-CogSci-96-4, John Hopkins Univ. ROA-154.

Tereshenko, N. 1962. [Терещенко Н.М.] О выражении позессивных отношений в самодийских языках. – Вопросы финно-угорского языкознания, Москва-Ленинград, с. 81-93.

Tereshenko, N. 1968a. [Терещенко Н.М.] Нганасанский язык. – Языки народов СССР, т.3, Москва.

Tereshenko, N. 1968b. [Терещенко Н.М.] К генезису частей речи (на материале самодийских языков). – Вопросы частей речи (на материале языков различных типов), Ленинград.

Tereshenko, N. 1979. [Терещенко Н.М.] Нганасанский язык. Ленинград.

Olga Vaysman
Dept. of Linguistics and Philosophy
Massachusetts Institute of Technology
77 Massachusetts Ave.
Cambridge, MA 02139

olga@mit.edu

Transitivity and Change of State Verbs

SAUNDRA K. WRIGHT
California State University, Chico

0. Overview

It has long been established that there is an important relationship between syntax and semantics in verb classification. In Fillmore's 1970 paper "The Grammar of *Hitting* and *Breaking*," he demonstrates that a verb's syntactic behavior can be predicted by its semantic meaning; crucially, verbs can be grouped together into semantically identifiable classes. This approach to verb classification is thoroughly explored in Levin's 1993 seminal work on English verb classes where she classifies over 3,000 verbs into almost 50 separate semantic classes.

In this paper, I focus on a specific class of verbs known as *change of state verbs*. These are verbs that involve a change in the internal composition of an entity undergoing a particular event. However, in recent work, change of state verbs have been subdivided into two different groups: *externally caused change of state verbs* include members such as *break, cool,* and *freeze*; *internally caused change of state verbs* include members such as *bloom, decay,* and *erode* (Levin 1993, Levin and Rappaport Hovav 1995). Further examples are given in (1):

(1) a. Externally Caused Change of State Verbs: *bake, boil, break, cool, crack, dry, freeze, lengthen, melt, open, shatter, straighten, widen, . . .*

b. Internally Caused Change of State Verbs: *bloom, blossom, corrode, decay, erode, ferment, germinate, molt, rot, rust, sprout, stagnate, wilt, wither, . . .*

Externally and internally caused change of state verbs can be distinguished in both a semantic and a syntactic sense. Semantically, they differ in that the externally caused change of state verbs imply the existence of an external causer with immediate control over the eventuality; meanwhile, internally caused change of state events involve properties inherent to the entities undergoing the events that are responsible for bringing about the eventuality. Syntactically, these verbs show differences in terms of their argument expression options. Externally caused change of state verbs participate in the causative/inchoative alternation. This is a transitivity alternation where a verb can be used both intransitively and in a related transitive variant, as shown in (2):

(2) a. The window broke.
    b. Bill broke the window.

However, it has been shown that while the internally caused change of state verbs occur in the intransitive variant, they fail to occur in related transitive forms (Levin and Rappaport Hovav 1995). An example is given in (3):

(3) a. The roses bloomed.
    b. *John bloomed the roses.

Nonetheless, it is apparent that some internally caused change of state verbs are found in transitive causative constructions, as demonstrated by examples like those in (4):

(4) a. Salt air rusted the metal pipes.
    b. Early summer heat wilted the petunias.

These examples raise an important question about the distribution of internally caused change of state verbs in transitive constructions: How can this variation be accounted for? In other words, why do we see examples like those in (4) but fail to see examples like (3b)? In this paper I propose an answer to that question. I argue that semantic and pragmatic properties—as opposed to syntactic properties—are responsible for the variability found with transitive uses of change of state verbs. In particular, a variety of factors combine together to determine the argument expression options associated with an individual verb. These factors include *controllability* (the degree to which an event can be externally manipulated), *causer type* (whether the event is human driven or nonhuman driven), and *subject-modification* (whether the causer is in a modified or unmodified form). Finally, I argue that it is possible to predict how acceptable a verb sounds in a transitive construction by a statistical model that combines and weights these different factors.

1. Transitive Uses of Change of State Verbs

The causative/inchoative alternation has often been used to characterize change of state verbs (Fillmore 1970). As I discuss above, externally caused change of state verbs clearly participate in this alternation (5a), but according to Levin and Rappaport Hovav (1995), internally caused change of state verbs do not (5b):

(5) a. The vase shattered. / Lynn shattered the vase.
    b. The cactus bloomed. / *The gardener bloomed the cactus.

Nonetheless, recent work by McKoon and Macfarland (2000) and Wright (2001) show that internally caused change of state verbs do occur in transitive constructions, as indicated by a number of attested corpus examples:

(6) a. Early summer heat blossomed fruit trees across the valley. (LN 1999)
b. Salt air and other pollutants can decay prints. (LN 1982)
c. Raindrops selectively erode clay particles. (BNC B1E)
d. The onset of temperatures of 100 degrees or more, on top of the drought, has withered crops. (NYT 1986)

However, even though these verbs are found in transitive constructions, it is clear that they do not pattern identically with the externally caused change of state verbs and in fact differ from these verbs in a number of important ways.

As I discuss in detail in Wright (2001), internally caused change of state verbs differ from externally caused change of state verbs in four main respects—frequency, causer type, verbal interpretation, and acceptability ratings. First, random searches of corpus data indicate that internally caused change of state verbs occur significantly less often in transitive constructions than externally caused change of state verbs ($t = 8.18$, $df = 26$, $p < .01$). Secondly, when internally caused change of state verbs are used transitively, they are more likely to involve a nature-related causer; meanwhile externally caused change of state verbs are more likely to involve a human causer (Chi-square = 340, $df = 2$, $p = 0$). Third, transitive uses of internally caused change of state verbs are more likely to involve a metaphorical interpretation than those with externally caused change of state verbs (Chi-square = 50.51, $df = 1$, $p = 0$). Finally, results from survey data show that internally caused change of state verbs are rated as less acceptable than externally caused change of state verbs in transitive constructions ($t = 9.41$, $df = 83$, $p < .0001$). These findings clearly demonstrate important distinctions between internally and externally caused change of state verbs and again raise questions about the behavior of change of state verbs with regard to transitivity: How can differences in transitivity be explained?

2. Analysis

In this paper I argue that semantic and pragmatic properties are responsible for the frequency and range of acceptability ratings associated with transitive uses of change of state verbs. In particular, I look at three factors that play a role in predicting transitive behavior: (1) causer type, (2) controllability, and (3) selectional restrictions/subject-modification. Finally, I show that these factors interact together to predict the acceptability scores associated with transitive uses of change of state verbs. Specifically, I demonstrate that a multiple regression model, incorporating these three factors, can predict acceptability ratings of transitive constructions remarkably well.

2.1 Causer Type

Causer type is one factor that can help to account for differences in transitive behavior across change of state verbs. There is a fairly extensive literature investigating what comprises the "prototypical" transitive event, and it has consistently been noted that the prototypical transitive evént involves a human causer deliberately performing an action that brings about a change of state. This is in contrast to less prototypical transitive events, which are those that typically involve nonhuman causers. (See Lakoff (1977), Hopper & Thompson (1980), DeLancey (1984, 1985, 1987), and Croft (1991), among others, for further discussion.)

In general, human-driven events are judged as more acceptable than nonhuman-driven events in transitive constructions. For example, sentences like those in (7), containing human causers, are rated by subjects as being completely acceptable—regardless if they are presented within or outside of any particular context:

(7) a. Nick broke the plate.
b. Jody cracked the teacup.

These high acceptability scores are related to the fact that these constructions have properties that correlate well with prototypical transitivity; in other words, they involve a human causer deliberately performing an action on an inanimate entity.

In English, however, it has often been noted that it is possible to have a nonhuman causer in subject position. Examples are given in (8):

(8) a. The bat broke the window.
b. The wind cracked the teacup.

While sentences likes those in (8) are generally considered to be grammatical, they tend to be judged as sounding somewhat unnatural and especially bizarre when presented out of any discourse context. In large part, this is due to the fact that these sentences do not correlate well with prototypical transitivity. In fact, causative verbs generally sound peculiar with nonhuman subjects because we typically try to associate transitive constructions with some notion of agentivity.

In these respects, we see a clear difference between externally and internally caused change of state verbs. Externally caused change of state events correlate well with the properties associated with prototypical transitive events. They often involve some notion of agentivity and involve human-driven events that bring about a change of state in some inanimate entity. However, internally caused change of state events are not typically associated with agentivity; in fact, they tend to involve nature-driven events that bring about a change of state in some biological entity. I suggest that these differences account in part for the distinctions between externally and internally caused change of state verbs that are found in transitive constructions. Change of state verbs that are more likely to occur transitively are those that are more likely to occur with human causers (i.e., the externally caused change of state verbs). Meanwhile, change of state verbs that are less likely to occur transitively are those that are less likely to occur with human causers (i.e., the internally caused change of state verbs). Moreover, differences in causer type can also help to account for the variation found within the change of state class, as well as across the change of state classes.

This link between agentivity and transitivity can be represented according to a continuum of event types, where potential agentivity corresponds positively with transitivity use. As depicted in Figure 1, events like *break* are often human driven; correspondingly, they often occur in transitive constructions. Events like *freeze* are somewhat less likely to be human driven, and they are somewhat less likely to occur in transitive constructions, and so on.

Figure 1 Transitivity and Agentivity with Change of State Events

```
-Transitive                                                    +Transitive
-Agent                                                              +Agent
  |----------------|----------------|-----------------|---------------|---------------|
                 bloom          germinate            freeze          break
```

Thus, I'd argue that an important difference between internally and externally caused change of state verbs has to do with agentivity and prototypical transitivity. Externally caused change of state verbs are more likely to occur transitively because they contain properties (i.e., human causers) that tend to be associated with prototypical transitive events.

2.2 Controllability

A second factor to consider is controllability. In a more general discussion of event types, Smith (1970) notes that events that must be externally manipulated by an external causer must occur transitively (e.g., *cut*), while events that cannot be externally manipulated by an external causer must occur intransitively (e.g., *laugh*). This distinction seems to be especially relevant for change of state verbs and might again best be represented in terms of a continuum of event types. In particular, I suggest that change of state verbs occur along a scale of controllability where their position on the scale is dependent upon the degree to which they can be manipulated by an outside source. Most externally caused change of state verbs can clearly be manipulated by an external causer; these events involve animate beings acting upon inanimate objects and involve events which can be manipulated externally. However, most internally caused change of state verbs are more questionable with regard to control; in fact, they can only questionably be manipulated by an external causer. Essentially, these events lead to some cognitive confusion. Again, they involve nature-related causers acting upon biological entities. They do involve external factors; in a blooming-event, for example, external factors such as sun, water, and soil nutrients are present in bringing about the event. However, the real question is to what extent do these external factors control the event and to what extent do they serve to facilitate an event which is largely under its own control? The answer to this question is a bit uncertain, and I'd argue that this type of vagueness accounts for the degraded acceptability ratings found with these verbs in transitive constructions. Externally caused change of state events can clearly be manipulated by an outside source; thus, they occur toward the *cut*-end of the scale of controllability. Internally caused change of state verbs are less clear in this respect; minimally, they occur closer toward the internal locus of control. This scale is depicted in Figure 2:

Figure 2 Locus of Control with Verbal Events

```
+ Internal Locus                                              + External Locus
of Control                                                          of Control
    |----------------|---------------|------------------|------------------|---------------|
  laugh           bloom           erode              freeze              break            cut
```

Thus, controllability also appears to play an important role in accounting for differences in transitive behavior across change of state verbs: externally caused change of state verbs are more likely to occur transitively because they tend to involve an external locus of control.

2.3 Selectional Restrictions/Subject-Modification

One last factor that I'd like to address is the role that selectional restrictions and subject-modification play in predicting verb behavior. Internally and externally caused change of state verbs also differ in the types and range of causers that tend to be associated with their events. Specifically, I suggest that internally caused change of state verbs select for a narrow range of causer types—so narrow, in fact, that causer type is actually inherent to the meaning associated with the verbs themselves and, therefore, doesn't need to be overtly specified to be understood as playing a role in the event. Meanwhile, externally caused change of state verbs select for a wider range of causer types; causer type is not inherent to the meaning associated with these verbs and, thus, must be overtly specified to be understood as being part of the actual event.

Evidence for these claims comes from further empirical results. In a survey task, I asked ten subjects to list three typical causers associated with a variety of different internally caused change of state verbs. For example, they received a verb like *bloom* and were asked to list three typical causers associated with a blooming-event. Overall, I found that there was considerable agreement in the listing of causer type where on average 8.5 distinct causers were listed across all subjects for each verb. This suggests that the range of causer types truly is restricted for verbs in this class. In fact, as shown by the data in (9), we have an expectation that these "prototypical" causers play a role in these events: if they are excluded or denied, the sentences really take on a somewhat surprising or unexpected interpretation:

(9) a. The roses bloomed, but the warm weather didn't cause it.
b. The plants wilted, but the heat didn't cause it.
c. The vegetables rotted, but the moisture didn't cause it.

For the externally caused change of state verbs, meanwhile, we see something quite different. In a similar survey task, subjects were asked to list three typical causers associated with a variety of different externally caused change of state verbs. For example, they received a verb like *break* and were asked to list three typical causers associated with a breaking-event. Here I found there was much greater variation in the listing of causer type; on average, 14.8 typical causers were listed for each verb. Moreover, even if the most typical causers are excluded from the event, as shown in (10), the sentences don't take on that same element of surprise:

(10) a. The car window broke, but I didn't cause it.
b. The car window broke, but the hail didn't cause it.
c. The car window broke, but the collision didn't cause it.

Furthermore, these verbs exhibit differences in acceptability ratings depending on their subject type. Specifically, sentences with internally caused change of state verbs are rated higher when the causer in subject position is "modified"—that is, being emphasized as playing a stronger than normal role in the event ($t = 2.54$, $df = 26$, $p < .01$). For example, sentences like (11a) and (12a) sounded worse to subjects than sentences like (11b) and (12b):

(11) a. ?Last July, sunlight wilted the begonias.
b. Last July, the intense sunlight wilted the begonias.

(12) a. ?This past summer, moisture rotted the tomatoes.
b. This past summer, extremely moist conditions rotted the tomatoes.

I argue that these differences in acceptability scores are related to selectional restrictions. For the internally caused change of state verbs, causer type is largely inherent to the meaning associated with the verbs themselves; thus, overtly specifying these causers is unnecessary—if not completely redundant—unless they are being highlighted (i.e., modified) as playing a stronger than normal role in the event.

However, for the externally caused change of state verbs, I did not find the same differences in judgments. Externally caused change of state verbs did not sound any better to subjects when the causer in subject position was in its bare form or when it was being modified ($t = .27$, $df = 26$, $p = .78$). Examples are given below:

(13) a. Over time, pressure shattered the rock.
     b. Over time, intense pressure shattered the rock.

(14) a. This morning, the sun melted the ice.
     b. This morning, the intensely hot sun melted the ice.

These facts are supported by the hypothesis that externally caused change of state verbs do not select for any particular causer type; the subjects sound equally as good in their modified and unmodified forms. Thus, I suggest that another distinction between externally and internally caused change of state verbs is causer type and, specifically, the degree to which causer type is selected for by the verb.

2.4 Multiple Regression Model

As I have demonstrated, a variety of factors contribute to predicting transitive behavior across change of state verbs. I conclude by showing that it is possible to predict how good a verb sounds in a transitive construction by a formula that weights these different types of factors. Interaction among different factors in predicting acceptability ratings can be demonstrated statistically as input into a multiple regression model, where the model can predict actual acceptability ratings remarkably well.

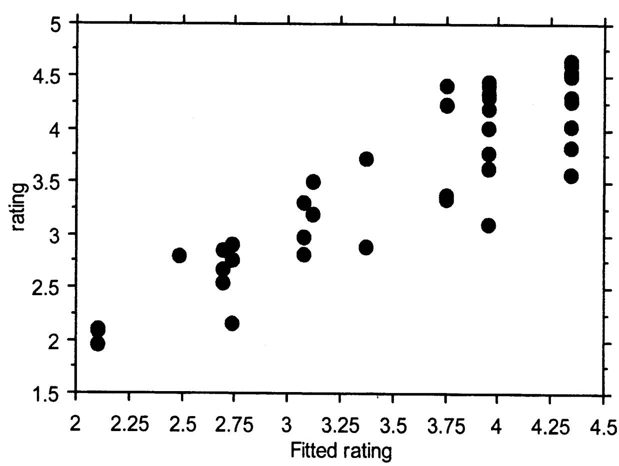
To carry out this examination, I presented subjects with a variety of transitive sentences involving change of state verbs and then asked them to judge the acceptability of these sentences on a five-point scale, where a rating of "1" indicated the sentence was not acceptable and a rating of "5" indicated that the sentence was completely acceptable. The sentences for this task were manipulated according to the factors discussed above (i.e., agentivity, controllability, and subject-modification).

Each sentence was then assigned values for the contribution of these factors to determining its acceptability. The sentences were coded in the following way. Agentivity received a binary assignment; I assigned "1" to sentences containing human subjects and "0" to those with nonhuman subjects. Subject-modification also received a binary assignment; here I assigned "1" for modified subjects and "0" for unmodified subjects. For controllability, I assigned one of five possible values to each sentence; this was based on the notion discussed earlier that controllability is a scalar property. Events that could not easily be externally controlled received a score of "1" for controllability; events that could easily be externally controlled received a score of "5" for controllability. Events that were more variable with regard to controllability received intermediate scores from "2" – "4". These scores were assigned for each verbal token with the assistance of two independent raters.

Separate multiple regression models were fit to establish the contribution of these factors to the direction and strength of subjects' acceptability ratings for both the internally and externally caused change of state verbs. The results demonstrate that for the internally caused change of state verbs, all three factors significantly contribute to the model. A model incorporating these three factors returns $r^2 = .80$, $p < .0001$. This represents the proportion of variability (80%) in the dependent variable (acceptability ratings) that can be explained by the set of independent variables (agentivity, controllability, and subject-modification). Although all three factors contributed significantly to predicting acceptability ratings, the weightings of the factors did differ. Controllability was the strongest factor overall ($t = 10.7$, $p < .0001$). This suggests that locus of control played the strongest role in accounting for acceptability judgments, and according to this model, the likeliness that an event could be externally manipulated correlated positively with acceptability ratings. Subject-modification was the least significant factor ($t = 3.3$, $p < .01$), although it still played a strong role; subject-modification also corresponded positively with acceptability ratings. Finally, agentivity was a strong predictor as well ($t = -4.8$, $p < .0001$). However, agentivity was a negative contributor for internally caused change of state verbs; in other words, it predicted lower scores rather than higher scores.

By establishing a fitted line to the data, we can compare the predicted acceptability ratings with the actual acceptability ratings. Figure 3 shows a graphical display of this model.

Figure 3 Acceptability Ratings: Internally Caused Change of State Verbs

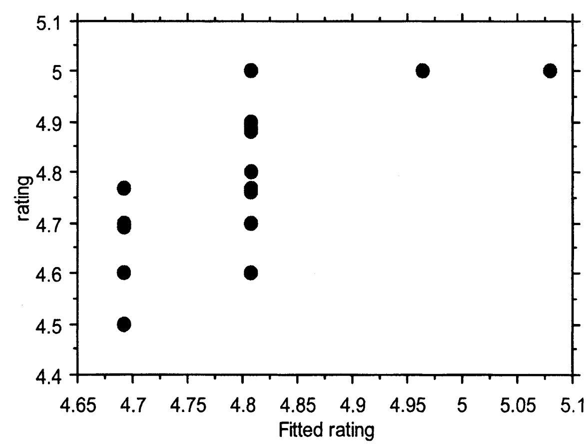


The actual acceptability ratings are given on the y-axis; the fitted ratings are given on the x-axis. Points toward the top of the graph indicate high acceptability scores; points toward the right of the graph indicate corresponding high predicted acceptability scores. Likewise, points toward the bottom of the graph indicate low acceptability scores, and points toward the left of the graph indicate corresponding low predicted acceptability scores. The linear relationship shown in the graph demonstrates that a multiple regression model, incorporating agentivity, controllability, and subject-modification, predicts acceptability ratings remarkably well.

For the externally caused change of state verbs, only two of the factors significantly contributed to the model. A model incorporating these three factors returns $r^2 = .65$, $p < .0001$. Here agentivity was the strongest factor overall ($t = 8.4$, $p < .0001$) and made a positive contribution to the model; that is, sentences with agentive causers received higher acceptability ratings. Controllability was a slightly less significant predictor ($t = 3.6$, $p < .001$), and subject-modification was not a predictor at all ($t = .3$, $p = .77$). The fact that subject-modification failed to predict acceptability ratings for externally caused change of state verbs is not surprising, given the analysis discussed in Section 2.3. Since externally caused change of state verbs have a wider range of causer types (i.e., causers which aren't implicit in the actual events), the model correctly predicts that subject-modification should have no effect in judging acceptability scores.

By establishing a fitted line to the data, we can compare the predicted acceptability ratings with the actual acceptability ratings. Figure 4 shows a graphical display of this model:

Figure 4 Acceptability Ratings: Externally Caused Change of State Verbs



Again, the actual acceptability ratings are given on the y-axis, and the predicted ratings are given on the x-axis. The linear relationship shown in the graph demonstrates that a multiple regression model, incorporating agentivity and controllability, predicts acceptability ratings fairly well. (Subject-modification is included, although it makes no significant contribution to the model.) Notably, the correlation between actual ratings and fitted ratings shown here is not as strong as the relationship depicted in the graph in Figure 3. This, however, is not surprising given the fact that there is far less variation in this data overall. (In fact, there was considerable overlap at many points in the graph.) Nonetheless, we now have clear evidence that these various semantic and pragmatic factors play an important role in accounting for the differences in acceptability ratings associated with internally and externally caused change of state verbs: when these factors are inputted into a multiple regression model, they can significantly predict the differences in acceptability rating scores.

3. Conclusion

Change of state verbs can be split into two main types. Internally caused change of state verbs include members such as *bloom, blossom,* and *wilt*; externally caused change of state verbs include members such as *break, cool,* and *crack*. In this paper I argue that both internally and externally caused change of state verbs are found in intransitive and related transitive constructions; in other words, all change of state verbs participate in the causative/inchoative alternation. However, these verbs differ in terms of the frequency and acceptability with which they are used transitively.

In this paper I show that a variety of semantic and pragmatic factors can help to account for the differences found across transitive uses of change of state verbs: causer type, controllability, and subject-modification. Finally, I demonstrate that the predictive power of these factors can be described statistically as input into a multiple regression model where the model predicts the actual acceptability ratings remarkably well.

References

Croft, William. 1991. *Syntactic Categories and Grammatical Relations.* Chicago, IL: University of Chicago Press.

DeLancey, Scott. 1984. Notes on Agentivity and Causation. *Studies in Language* **8**: 181-213.

DeLancey, Scott. 1985. Agentivity and Syntax. In *Papers from the Parasession on Causatives and Agentivity at the 21st Regional Meeting*, 1-12. Chicago Linguistics Society, Chicago, IL.

DeLancey, Scott. 1987. Transitivity in Grammar and Cognition. In R. S. Tomlin (ed.) *Coherence and Grounding in Discourse*, 53-68. Philadelphia, PA: John Benjamins.

Fillmore, Charles J. 1970. The Grammar of *Hitting* and *Breaking*. In R. Jacobs & P. Rosenbaum (eds.) *Readings in English Transformational Grammar*, 120-133. Waltham, MA: Ginn.

Hopper, Paul and Sandra A. Thompson. 1980. Transitivity in Grammar and Discourse. *Language* **56**: 251-299.

Lakoff, George. 1977. Linguistic Gestalts. In *Papers from the 13th Regional Meeting* 225-235. Chicago Linguistics Society, Chicago, IL.

Levin, Beth. 1993. *English Verb Classes and Alternations: A Preliminary Investigation.* Chicago, IL: The University of Chicago Press.

Levin, Beth and Malka Rappaport Hovav. 1995. *Unaccusativity: At the Syntax-Lexical Semantics Interface.* Cambridge, MA: MIT Press.

McKoon, Gail and Talke Macfarland. 2000. Externally and Internally Caused Change of State Verbs. *Language* **76**: 833-858.

Smith, Carlotta S. 1970. Jespersen's "Move and Change" Class and Causative Verbs in English. In M. A. Jazayery, E. C. Polomé, & W. Winter (eds.) *Linguistic and Literary Studies in Honor of Archibald A. Hill. Vol. 2: Descriptive Linguistics*, 101-109. The Hague: Mouton.

Wright, Saundra K. 2001. Internally Caused and Externally Caused Change of State Verbs. PhD dissertation, Northwestern University, Evanston, IL.

Corpora

British National Corpus (BNC)
Brown Corpus (Brown)
Collins Cobuild (Cobuild)
Lexis-Nexis, Academic Universe (LN)
New York Times (NYT)
Oxford English Dictionary (OED)

Dr. Saundra K. Wright
Department of English
California State University, Chico
Chico, CA 95929-0830

skwright@csuchico.edu

A Lexical Approach to English Floating Quantifiers[1]

EUN-JUNG YOO
Seoul National University

0. Introduction

This paper investigates how the syntactic and semantic characteristics of floating quantifiers (FQ) in English can be explained within the framework of the Head-Driven Phrase Structure Grammar. Previous analyses treat FQs either basically as NP quantifiers (Postal 1974, Maling 1976, Sportiche 1988, McCawley 1998), or as VP quantifiers that are syntactically and semantically distinct from NP quantifiers (Dowty & Brodie 1984). In contrast, this paper proposes an analysis in which FQs are base-generated VP modifiers as in Dowty & Brodie, but their logical (or semantic) contributions are made analogous to that of quantificational determiners, through a precise lexical description of FQs that specifies the information on "storage" and "retrieval" of the quantifier meaning. The proposed analysis is "lexical" in its nature in the sense that the syntactic and semantic properties concerning distribution and scope interaction are encoded as part of the lexical information that the FQs have.

1. Properties of English FQ Constructions

In English, only the quantifiers *all, each*, and *both* can "float." This contrasts to FQ constructions in languages like Japanese and Korean, which permit floating of a wider range of quantifiers including numeral quantifiers. Examples of English FQs are provided in (1) below.

(1) a. The children have all read the books.
    b. The students have each arrived.
    c. John's brothers have both read the book.

Floating quantifiers show characteristics that differ from normal quantificational determiners. First, FQs in English are subject-oriented. In (2), for

[1] An earlier version of this paper was presented at the 15th Pacific Asia Conference on Language, Information, and Computation in February 2001. I would like to thank both the BLS 28 audience and PACLIC 15 participants for their insightful comments. All remaining errors are my own.

example, the quantifier can be construed with the subject NP, but not with the object NP.

(2) The children have all introduced their friends to us.

Thus (2) is not interpreted as 'The children have introduced all of their friends to us'. While some languages such as French, Japanese, and Korean allow quantifiers to "launch" from objects, no language is known to permit FQs hosted by adjuncts. This should be predicted in a proper analysis of FQs.

Second, FQs appear in front of a VP or AP.

(3) a. The computers all will have been moved to the new office.
b. The computers will all have been moved to the new office.
c. The computers will have all been moved to the new office.
d. The computers will have been all moved to the new office.
e. *The computers will have been moved all to the new office.
f. *The computers will have been moved to the new office all.

(4) a. We were all fast asleep. (Quirk *et al.* 1985:382)
b. The children are all healthy.

Third, FQs are hosted by (more or less) definite plural NPs. The examples in (5) and (6), from Dowty & Brodie, illustrate the kinds of NPs the FQ can semantically depend on.

(5) a. John, Mary and Susan all left.
b. John and Mary both left.
c. The students all left.
d. (?) Five students have all turned in their exams.
e. Five contestants, who were selected as finalists by the judge yesterday, will all perform again tomorrow.

(6) a. *John, Mary or Susan all left.
b. *Few students all left.
c. *No students all left.
d. *At least five students all left.

NPs like *five students* usually have an interpretation equivalent to 'at least five students'. However, as Ladusaw (1982) and Dowty & Brodie have suggested, they may also have an interpretation corresponding to 'exactly five particular students'. Example (5e) supports the view that in certain contexts, NPs like *five students* may have a definite interpretation. Thus when indefinite NPs have definite interpretations, they can be the semantic target a FQs.

Finally, unlike ordinary quantificational NPs, FQs do not exhibit scope ambiguity with respect to other scopal elements such as negation adverbs and modals (Dowty & Brodie 1984:77).[2]

(7) a. The students all didn't leave.
a′. $[\forall x|\ student'(x)]\ not'(leave'(x))$
b. The students didn't all leave.
b′. $not'([\forall x|\ student'(x)]\ leave'(x))$

(8) a. The contestants all can win.
a′. $[\forall x|\ contestant'(x)]\ can'(win'(x))$
b. The contestants can all win.
b′. $can'([\forall x|\ contestant'(x)]\ win'(x))$

With FQs, scope ambiguity is not observed in raising verb constructions either (McCawley 1998:631).

(9) a. His conclusions all appear to be incorrect.
b. His conclusions appear to all be incorrect.

(10) All his conclusions appear to be incorrect.

While (10) is ambiguous with regard to which takes wide scope, *appear* or *all*, in (9), the surface order between the two decides relative scope.

2. Previous Analyses

2.1. Derivational Approaches

Within transformational grammar, it is generally assumed that there is syntactic dependency between the host NP and the FQ. Thus the following FQ sentences in (11) are presumed to be related to the ones in (12) via movement.

(11) a. The children all left.
b. John believes the reporters both to have left.

(12) a. All (of) the children left.
b. John believes both (of) the reporters to have left.

There have been largely two approaches with respect to the syntactic dependency. The first is to posit rightward movement of the quantifier, deriving (11) from (12), as in Postal (1974), Mailing (1976), and McCawley (1998). As Haegeman & Guéron (1999) note, however, such a downward derivation increases the overall

[2] Actually, the sentence (7a) also has an additional reading in which negation takes wide scope. Following Dowty & Brodie's suggestion, however, I will take this additional reading to be an outcome of metalinguistic negation.

complexity of the grammar since the generalization can no longer hold that the moved constituent c-commands its trace. Another approach, which is proposed in Sportiche (1988), is to posit leftward movement of the NP, under the assumption that the phrase Q NP is in the [SPEC, VP] in the deep structure. While Sportiche mainly deals with French FQ constructions, the possibility of extending his analysis to English examples is also discussed. The sentences in (12) are derived, if the whole Q NP sequence is moved to [SPEC, IP], whereas the ones in (11) are generated when only the NP is moved to [SPEC, IP]. The latter option that derives (11) is described schematically in (13).

(13) $NP^{\wedge}$ ... [$_{Xn}$ Q [$_{NP*}$ *e*] XP]

In (13), Q is adjacent to NP*, an NP trace, and there is an anaphoric relation between $NP^{\wedge}$, the overt antecedent of Q, and the trace. Thus the anaphor-like behavior between the host NP and Q is captured via movement.

However, as Sportiche himself notes, such an analysis raises a question for the examples where Q appears before the first auxiliary verb:

(14) a. The carpets will ^have ^been ^being ^dusted for two hours.
b. The carpets all will have been being dusted for two hours.

In (14a), at least one empty NP position can be postulated in front of each verb, so that the occurrence of Q in ^-marked positions is explained by (13). On the other hand, in (14b), the occurrence of *all* in front of the finite auxiliary verb is problematic, because there is no position available for the empty NP* due to the presence of the subject NP in [SPEC, I]. Sportiche claims that this problem can be solved if the trace is assumed to be in [SPEC, I] and the subject $NP^{\wedge}$ is topicalized. However, the assumption that the subject NP in (14b) is in the topic position is problematic, because topicalization of an object is still possible with an FQ in the pre-Infl position.

(15) a. An office this large, the students all will desire.
b. To the opera, the students all have been.

Given that multiple topicalization is not allowed in English, examples like (15) show that the subject NP is not in the topic position in (14b).

Another problem with Sportiche's approach (and other derivational approaches) is that it cannot explain why English permits quantifier floating only for a few quantifiers (i.e., *all, each,* and *both*). It cannot be said that quantifier floating is allowed only for universal quantifiers, since *every* does not float. Given that some languages permit a wider range of quantifier floating (and that the characterization of the permitted range is not clear), this approach would need to introduce an ad hoc constraint to the grammar in order to prevent the derivation in (13) from applying to non-FQs in examples like (16).

(16) a. *The students three will leave.
b. *The children most will leave.

Furthermore, as noted in Sportiche, this analysis provides no explanation as to why FQs do not appear before a *wh*-trace.

(17) a. How angry do you think they all were __?
b. *How angry do you think they were all __?

Since nothing prohibits the XP in (13) from being a *wh*-trace, Sportiche's analysis cannot account for the contrast in (17).

2.2. Dowty & Brodie's Non-Derivational Analysis

Dowty & Brodie (1984) propose that FQs are base-generated as VP modifiers. Thus in their analysis, FQs are introduced by the following syntactic rule:

(18) <1, $[_{VP}$ Q VP], Q′(VP′)>.

Their analysis focuses on providing formal semantics for FQ constructions. According to them, FQs, as VP quantifiers, belong to the semantic type distinct from that of determiners. More specifically, FQs map VP-denotations into argument NP denotations. Such a semantic function of FQs accounts for why quantifiers float only from arguments, but not from adjuncts. Furthermore, since the semantic rule for VP quantifiers depends on the non-empty intersection of all the sets in the NP denotation, it will require that the host NPs should be the class of definite plurals.

Dowty & Brodie's analysis naturally explains the properties discussed in section 2, without positing any transformation process. However, since it relies heavily on rule-to-rule semantic interpretations, it is hard to be incorporated into a syntactic framework that does not assume a compositional model-theoretic semantics that accompanies each expression. In what follows, I will present an alternative non-derivational analysis of FQ constructions, in which both syntactic and semantic aspects of these sentences can be accounted for via interactions with each other within a feature structure.

3. FQs and Quantifier Retrieval

3.1. Lexical Representations of FQs

In presenting an analysis of FQs, we must first consider what kind of logical forms can be assigned to the sentences with FQs. Since the host NPs are plural NPs, how to deal with their semantic representation should be decided. Following Link (1983), I assume that collective and mass entities are contained in the model as individuals.

Link (1983) and Dowty (1986) argue that the addition of *all* to plural NPs invokes universal quantificational force, and has an effect dubbed as the Maximizing Effect, requiring the predicate in question to be true of every member of the group. Incorporating this generalization, the logical form of (19a) can be expressed by (19b), using an informal restricted quantificational logic notation:

(19) a. The students all sneezed.
b. [the y| students′(y)]([∀x| constituent-of′(x,y)] (sneezed′(x)))

In (19b), 'const(ituent)-of' is a function that resembles Link's (1983) relation 'atomic-part-of', and identifies each member x of the group y which is contextually salient.

On the other hand, in the case of a collective predicate, it is semantically abnormal that the predicate holds for each member of the given group. Accordingly, the group entity itself should be predicated.

(20) a. The students all gathered.
b. [the y| students′(y)]([∃x | group′(x) & ([∀z | constituent-of′(z,y)] (constituent-of'(z,x)))] (gathered′(x)))

The logical representation (20b) may look complicated. However, such complexity seems to be unavoidable in order to take into account the group reading in examples like (21).

(21) Most of the 20 students gathered.

In the latter, the logical form can be described as '[the y| students′(y) & |y|= 20]([∃x | group′(x) & ([most z | constituent-of′(z,y)](constituent-of′(z,x)))] (gathered′(x)))'.[3]

As (19b) and (20b) indicate, the logical representations of FQ sentences vary depending on the distributive/collective interpretation of the predicate. Furthermore, when the FQ *all* is used with a collective predicate, an existential quantifier arises in the interpretation. We take this fact to support a lexical approach to the FQ construction because, in such approach, two different lexical descriptions of the FQ may result in different logical interpretations.

Now let us consider how the informal logical representation discussed so far can be expressed in our theory. Our representation of quantifier scope is based on that of Pollard & Sag (1994), in which the semantic contribution of a word or phrase is represented as the value of the CONT(ENT) attribute in the feature structure. In the CONT value of a type *psoa (parameterized-state-of-affairs)*, the quantifier in the QUANT(IFIER)S list is taken to have scope over the

[3] I thank an anonymous abstract reviewer of the HPSG 2001 for pointing this out to me.

NUC(LEUS) value. Accordingly, the CONT of (19a) can be described by (22). When there is more than one quantifier in the QUANTS, the preceding member is taken to have scope over the preceded one(s).

(22)
$$\left[\begin{array}{ll} \text{QUANTS} & \left\langle \boxed{2}\left[\begin{array}{l} \text{DET } \textit{the} \\ \text{IND } \boxed{6} \\ \text{RESTR } \left\{\left[\begin{array}{l} \textit{students} \\ \text{INST } \boxed{6} \end{array}\right]\right\} \end{array}\right], \boxed{4}\left[\begin{array}{l} \text{DET } \textit{forall} \\ \text{IND } \boxed{8} \\ \text{RESTR } \left\{\left[\begin{array}{l} \textit{const-of} \\ \text{CONSTITUENT } \boxed{8} \\ \text{SUM } \boxed{6} \end{array}\right]\right\} \end{array}\right] \right\rangle \\ \text{NUC} & \boxed{3}\left[\begin{array}{l} \textit{sneeze} \\ \text{SNEEZER } \boxed{8} \end{array}\right] \end{array}\right]$$

In (22), the numeral tag $\boxed{6}$ indicates a plural entity, while $\boxed{8}$ expresses a constituent member of the plural entity. Likewise, based on (20b), the CONT of (20a) can be described as a structure whose QUANTS contain a definite quantifier followed by an existential quantifier, and whose NUC(LEUS) consist of the psoa *gather*.

In Pollard & Sag's 1994 theory of quantifier scope, the meaning of a quantifier 'starts out in storage' in the QS(TORE) and is "inherited" into a larger phrase in the structure, and then "retrieved" to take scope over a certain phrase or sentence. This theory is revised and extended in Pollard & Yoo (1998) in order to account for scope phenomena in raising verb constructions and unbounded dependency constructions. It contains a set of new assumptions for the account of quantifier scope: i) the QSTORE feature is relocated as a LOCAL attribute, ii) a new feature POOL is introduced as an additional LOCAL attribute, iii) "ordinary" lexical heads "collect" all the QSTORE values of their "selected arguments," iv) QSTORE values are inherited only from the semantic daughter of a phrase, and v) quantifier retrieval is possible either at a lexical head or a phrase. Among the QSTORE, POOL, and RET(RIEVED) values, the following constraint holds:

(23) For a sign, the RETRIEVED value is a list whose set of elements forms a subset S of the POOL value; and the QSTORE value is the relative complement of the set S.

As in Pollard & Sag, the elements in the RETRIEVED also appear in the QUANTS value to take their scope.

When considering how to represent the FQ *all* in the lexicon, the most natural assumption that we can make is that *all* introduces a quantifier in its POOL, just like quantificational determiners. In Pollard & Yoo, the words that give rise to a quantifier meaning are classified as *quant(ifier)-word*, and their POOL and QSTORE values are lexically specified. These words are distinguished from ordinary lexical heads in that their QSTORE values are not the union of all the QSTORE values of their arguments.

What I will further propose in this paper is that the RETRIEVED value of certain quantifier-introducing words should be lexically specified as well. Thus, I claim that in addition to ordinary cases where a quantifier is retrieved at some structural node, obeying a set of constraints on the features POOL, QS(TORE), RET(RIEVED), and QUANTS, it is necessary to specify the RET (and thus QUANTS) values lexically for some quantifier-introducing words.[4]

3.2. Analysis and Explanation

Based on the foregoing discussion, I propose that the quantifier scoping in English FQ sentences can be accounted for by specifying retrieved quantifiers in the lexical entry of the FQ *all*. As discussed in 3.1, the quantifiers arising in the FQ sentences vary depending on the semantic type of the predicates. Accordingly, two lexical entries are provided for the FQ *all*. The first entry in (24) is for the sentences with a distributive predicate.

(24) all_1 (for distributive predicates)

$$\left[\begin{array}{ll}
\text{MOD} & \text{VP}\vee\text{AP[SUBJ}\langle\text{NP [QS }\{\boxed{2}\}]\rangle]\text{: }\boxed{3} \\
\text{CONT} & \left[\begin{array}{l}\text{QUANTS }\boxed{5}\\ \text{NUC }\boxed{3}\text{ [ARG }\boxed{8}]\end{array}\right]\\
\text{POOL} & \{\boxed{2}\}\cup\{\boxed{4}\}\\
\text{QS} & \{\ \}\\
\text{RET} & \boxed{5}\ \langle \boxed{2}\left[\begin{array}{l}\text{DET } \textit{the}\\ \text{IND }\boxed{6}\end{array}\right], \boxed{4}\left[\begin{array}{ll}\text{DET} & \textit{forall}\\ \text{IND} & \boxed{8}\\ \text{RESTR} & \left\{\left[\begin{array}{l}\textit{const-of}\\ \text{CONSTITUENT }\boxed{8}\\ \text{SUM }\boxed{6}\end{array}\right]\right\}\end{array}\right]\rangle
\end{array}\right]$$

In (24), the QUANTS list, which has the same value $\boxed{5}$ as the RET, contains two quantifiers, $\boxed{2}$ and $\boxed{4}$. The quantifier $\boxed{2}$ indicates a definite quantifier arising from the subject NP, and $\boxed{4}$ corresponds to '[∀x| constituent-of′(x,y)].' When the FQ appears in a sentence, the quantifiers will scope over the NUC(LEUS) value, which is the CONT of the VP. Accordingly, (19a) will have the CONT in (22).

The second lexical entry of the FQ *all* for collective predicates is described in (25).

---

[4] Lexical specification of quantifier retrieval is independently motivated, because there are other examples like (i) where a quantifier word needs to contain a nonempty QUANTS value.

(i) Every kid's favorite toy broke.

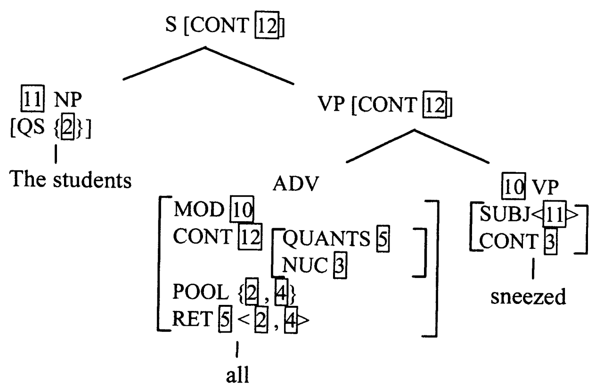
See Yoo (2001) for the detailed discussion of how the narrow scope reading '[the x| toy′(x) & ([∀y| kid′(y)] poss′(y,x))] broke′(x)' can be accounted for by the lexical specification of retrieved quantifiers in the lexical entry of the possessive determiner *'s*.

(25) *all*$_2$ (for collective predicates)

$$\begin{bmatrix} \text{MOD} & \text{VP}\vee\text{AP[SUBJ}\langle\text{NP [QS }\{\boxed{2}\}]\rangle]: \boxed{3} \\ \text{CONT} & \begin{bmatrix}\text{QUANTS } \boxed{5} \\ \text{NUC } \boxed{3} \text{ [ARG } \boxed{8}] \end{bmatrix} \\ \text{POOL} & \{\boxed{2}\} \cup \{\boxed{4}\} \\ \text{QS } \{\} & \\ \text{RET } \boxed{5} & \langle \boxed{2}\begin{bmatrix}\text{DET } \textit{the} \\ \text{IND } \boxed{6}\end{bmatrix} \boxed{4}\begin{bmatrix} \text{DET } \textit{exists} \\ \text{IND } \boxed{8} \\ \text{RESTR } \left\{ \begin{bmatrix}\textit{group} \\ \text{INST } \boxed{8}\end{bmatrix}, \begin{bmatrix} \text{QUANTS} \langle [\forall\boxed{7} \mid \textit{const-of}(\boxed{7},\boxed{6})]\rangle \\ \text{NUC} \begin{bmatrix}\textit{const-of} \\ \text{CONSTITUENT } \boxed{7} \\ \text{SUM } \boxed{8}\end{bmatrix}\end{bmatrix} \right\} \end{bmatrix} \rangle \end{bmatrix}$$

I assume that the choice between (24) and (25) is made on semantic grounds. Thus, to analyze collective reading examples like (20a) with the entry (24) will yield semantic anomaly, because it is not the case that each constituent member of the group 'gathers'. Given the lexical entry in (24), the sentence in (19a) can be assigned the structure in (26). In (26), the numeral tags $\boxed{2}$ and $\boxed{4}$ indicate the elements in the QUANTS in (22) and the tag $\boxed{3}$ corresponds to the NUC in (22).

(26)



Now let us consider how the lexical entries in (24) and (25) provide explanations for the properties of FQs discussed in section 2. First of all, the fact that only three words, *all, each*, and *both*, are used as FQs is easily explained, since only these three will have a lexical entry that looks like (24) or (25). Other quantifiers in English will only have the entry of determiners.

Second, the subject-oriented property of FQs is accounted for, since the index of the quantifier $\boxed{8}$ is related to that of the VP or AP subject, i.e. $\boxed{6}$, via the relation *constituent-of*. This analysis also can be extended to account for FQs in the languages allowing both subjects and objects to have associated FQs. This is because the FQ may access the arguments of the verb that it modifies via the

ARG(UMENT)-ST(RUCTURE) feature (or the SUBJ and COMPS features) of the verb.

Third, since *all* selects a VP or AP via its MOD feature, it appears as a modifier of a VP or AP in the structure. In other words, *all*, as an adjunct daughter, combines with a VP or AP, constituting a head-adjunct phrase. Yet it remains to be answered why the modifier does not appear in the post-VP positions, as (3e-f) illustrate. There seem to be two ways to account for this. One solution is to posit a Linear Precedence (LP) rule that holds between "light" modifiers like FQs and a modified VP/AP. (See Abeille & Godard (1999:92) for the same type of LP rule proposed for French.) In the LP rule, we can specify that a "light" non-head daughter must precede the head daughter. Alternatively, if we adopt Bouma *et al.*'s (2001) proposal that post verbal adjuncts are added to the DEPENDENTS list by a lexical rule, then we can specify that certain adverbs such as *all, both, each, only, never, certainly*, and *probably* are not added by this lexical rule. The only way these adverbs are introduced in a sentence is via a head-adjunct structure in which the head comes after the adjunct.

Moreover, as Sag & Fodor (1994) argue, if FQs are treated as VP-adjoined modifiers, then the ungrammaticality of (17b) can be elegantly explained by employing a traceless approach to extraction available in HPSG. Examples like (17b) are ruled out, simply because there is no VP/AP that FQs can combine with.

Next, the generalization that FQs are hosted by (contextually) definite NPs is accounted for, because, according to (24) and (25), the SUBJ element of the VP that *all* modifies has a QS member whose DET is *the*. Therefore, if the host NP is an indefinite NP like *some students*, the DET value of the NP's QS is *exists*, which conflicts with the constraint imposed on the descriptions (24) and (25).

Furthermore, the use of the relation *constituent-of* predicts that a singular entity cannot be the host of a FQ. Since there is no plural (or group) entity from which its constituents can be extracted, the following examples are correctly ruled out:

(27) a. *The student has all/each arrived.
b. *Each student had all arrived.

Finally, in (24) and (25), the universal quantifier associated with the FQ is required to be retrieved lexically and take scope over the VP that it modifies. Thus the quantifier associated with *all* does not inherit into a larger phrase or the lexical head that selects the phrase. Consequently, when the modified VP contains a modal or negation, *all* has wide scope over such scopal elements. This explains the interpretations of (7a) and (8a). Likewise, when *all* is lower than the modal or negation element in the structure, as in (7b) and (8b), it takes narrow scope.

So far our discussion has been focused on the quantifier *all*. The analysis proposed for *all* can be easily extended to the discussion of *each* or *both*. As Dowty & Brodie note, unlike *all*, *each* and *both* are restricted to individual-level NPs and do not have group reference.

(28) a. *Each (of the) student(s) gathered in the stadium.
b. *Both students are a happy couple.

This implies that, between the two types of logical forms, (19b) and (20b), only the (19b) type is available. The distinction between *each* and *both* will, of course, come from the fact that *both* is limited to an NP whose cardinality is 2. Since the FQ *each* does not occur with collective predicates, its lexical entry will look like the first entry of *all*.

The present analysis can also account for the examples like (29)-(30) where the host NP contains a quantifier with a more complex RESTR value.

(29) The three students have all played tennis.

(30) John's students all came to the party.

In (29), we take the quantifier in the QS(TORE) value of the subject NP to be '[the y| students′(y) & |y| = 3],' in which |y| indicates the cardinality of y. As (24) imposes, this quantifier, indicated by $\boxed{2}$ in (24), cannot be retrieved or inherited into a larger VP. Instead, it is retrieved lexically, together with the universal quantifier arising from *all*. Consequently, in (29), there will be two quantifiers in the QUANTS list of the sentence, i.e., '[the y| students′(y) & |y| = 3]' and '[∀x| constituent-of′(x,y)]'.

Likewise, in (30), the quantifier in the RET and QUANTS list of *all* is the definite quantifier associated with the subject NP and the universal quantifier arising from *all*. Thus the two quantifiers in the QUANTS list of the sentence will be '[the y| students′(y) & possess′(john, y)]' and '[∀x| constituent-of′(x,y)]].' Accordingly, complicated examples such as (29-30) can also be handled by the lexical entries in (24) and (25).

4. Concluding Remarks

In this paper, a constraint-based, lexical approach to English floating quantifier constructions has been presented. Drawing on Dowty & Brodie's assumption that FQs are base-generated as VP modifiers, I have focused on providing an appropriate CONTENT value for the sentences containing FQs. In representing the semantic contribution of the FQs, I have proposed that both "quantifier storage" and "quantifier retrieval" take place lexically at their sites. Accordingly, a FQ, which functions as a semantic head of a VP, carries all the necessary semantic information for the VP, including the quantifier meaning. With such a mechanism of lexical specification of quantifier retrieval, I have shown that various properties of FQs, with respect to syntactic distributions, the types of host NPs, and scope interaction with adjacent elements, can be accounted for by a precise description of the lexical entry of the FQs.

References

Abeille, A. and D. Godard. 1999. A lexical approach to quantifier floating in French. In G. Webelhuth *et al.* (eds.) *Lexical and Constructional Aspects of Linguistic Explanation*, 81-96. Stanford, CA: CSLI.

Bouma, G. and R. Malouf. & I.A. Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory* **19**: 1-65.

Dowty, D. 1986. Collective predicates, distributive predicates, and *all*. *ESCOL* **86**: 97-115.

Dowty, D. and B. Brodie. 1984. The semantics of "floated" quantifiers in a transformationless grammar. In *WCCFL* 3, 75-90. Stanford, CA: CSLI.

Haegemann, L. and J. Guéron. 1999. *English Grammar*. Blackwell Publishers.

Ladusaw, W. 1982. Semantic constraints on the English partitive construction. In *WCCFL* 1, 231-242.

Link, G. 1983. The logical analysis of plurals and mass terms: a lattice-theoretic approach. In R. Bauerle *et al. Meaning, Use, and Interpretation of Language*. Berlin and New York: Walter de Gruyter.

Maling, J. M. 1976. Notes on quantifier postposing. *Linguistic Inquiry* **7**: 708-718.

McCawley, J. D. 1998. *The Syntactic Phenomena of English*. Chicago: The University of Chicago Press.

Pollard, C. and I. A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: The University of Chicago Press.

Pollard, C. and E. J. Yoo. 1998. A unified theory of scope for quantifiers and *wh*-phrases. *Journal of Linguistics* **34**: 415-445.

Postal, P. M. 1974. *On Raising*. Cambridge: MIT Press.

Quirk, R., S. Greenbaum, G. Leech, & J. Startvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

Sag, I. A. and J. D. Fodor. 1994. Extraction without traces. In *WCCFL* 13. Stanford, CA: CSLI.

Sportiche, D. 1988. A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* **19**: 425-449.

Yoo, E.-J. 2001. English floating quantifiers and lexical specification of quantifier retrieval. *Language and Information* **5**:1-15. Seoul, Korea.

Eun-Jung Yoo
Department of English Language and Literature
Seoul National University
Seoul, 151-742, Korea

ejyoo@snu.ac.kr

Vowel Acquisition in Hungarian: A First Look at Developmental Data

KRISZTINA ZAJDÓ
University of Washington, Seattle

1. Vowels: The orphans of speech research

Vowel acquisition in children is a poorly researched area of speech development. Studies of phonological acquisition and theories of phonological development assumed that vowels develop early. Until recently, vowels were no more than the "poor relations of consonants" (Ball and Gibbon 2002:xi). Phonological assessments of children with typical and atypical speech development routinely ignored to mention the status of vowel production.

However, recent investigations have revealed that the path from the emergence of quasi-vowels in babbling to more adult-like vowel production capabilities of older children requires an awareness of the underlying vowel system of the target language as well as articulatory learning. While the age at which vowels are fully acquired is unknown, data suggest that the major steps of vowel development occur during the first six years of life.

2. On the necessity of studying vowel development

One serious limitation of prior research is that, while developmental changes in children's vowel perception, discrimination and representation have been given considerable attention in several languages, studies on vowel acquisition have focused almost exclusively on English. Because the vowel systems of diverse languages vary along several dimensions, there is a clear need for investigating vowel development in children from various language communities.

So as to expand our knowledge about normal tendencies in the development of children's speech sound (and specifically vowel) production in general and to contribute to remediation techniques of vowel disorders in children and adults, the careful formulation of a general theory of vowel acquisition appears to be in order. Developing a theory of vowel acquisition should not, however, be viewed merely as an end-product of summarizing our current knowledge of the matter. Rather, this process should, in a step-like fashion, guide us further in exploring issues that explain both the approaches of the individual as well as the general tendencies in the acquisition of any one particular or several vowel systems. In order to account for distinct patterns that may occur while acquiring all the available (phonologically diverse) vowel systems of the world's languages, a theory that can equally accommodate the different paths that might be taken while developing various vowel systems is required. Such a theory must, unquestionably, be based on cross-linguistic developmental data.

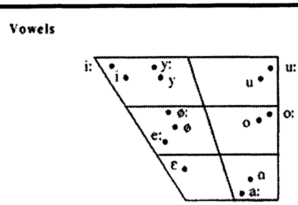
3. The present study

To facilitate the process of generating a wide knowledge base of vowel developmental patterns in languages other than English, this project is aimed at determining developmental tendencies in vowel acquisition in 2 to 4-year-old monolingual children, acquiring the standard dialect of Hungarian as their first language. It is hoped that the results of this inquiry will constitute a much-needed step towards furthering our knowledge in the right direction, so as to contribute to the emergence of a general theory of vowel acquisition, both in children and adults.

3.1. The vowel system of standard Hungarian

The vowel inventory of the standard dialect of Hungarian contains fourteen monophthong vowels that are perceived by native speakers as seven vowel-pairs. The vowels are differentiated along four dimensions: (1) front and back tongue-positions; (2) high, half-closed, half-open or open tongue-height (only the vowel /a:/ belongs to this latter category); (3) rounded or unrounded lip-position; (4) short or long duration.

(1) The vowel system of the standard dialect of Hungarian (Szende, 1999).



i	vis	*visz*	'carry'	i:	vi:z	*víz*	'water'
y	yt	*üt*	'hit'	y:	ty:z	*tűz*	'fire'
u	u	*ujj*	'finger'	u:	u:t	*út*	'road'
ɛ	ɛz	*ez*	'this'	e:	e:l	*él*	'live'
ø	øl	*öl*	'kill'	ø:	lø:	*lő*	'shoot'
o	ok	*ok*	'cause'	o:	to:	*tó*	'lake'
ɒ	hɒt	*hat*	'six'	a:	va:r	*vár*	'castle'

3.1.1. Vowel-pairs differentiated primarily by quantitative features

Five of the pairs (the high vowels /i/ - /i:/, /y/ - /y:/ and /u/ - /u:/, and the mid vowels /ø/- /ø:/ and /o/ - /o:/) are differentiated primarily by their duration, being relatively similar in other aspects. The average ratio of short and long vowels is quite pronounced in the adult language: it has been documented to be around 1:2 in spontaneous speech (e.g. Kassai 1979, Tarnóczy 1974).

3.1.2. Vowel pairs differentiated primarily by qualitative features

The members of the vowel-pairs /ɛ/ - /e:/ and /ɑ/- /a:/ are differentiated primarily by qualitative features though they are both qualitatively and quantitatively different. The vowel /ɛ/ is described as short half-open nonlabial front, whereas the vowel /e:/ is a long half-closed nonlabial front vowel. The /ɑ/ sound is a low back vowel, whereas the central vowel /a:/ is the only vowel produced with the lowest jaw-position. The /ɑ/ - /a:/ pair also differs in the lip-position required during their production: while the former is slightly labial, the latter is nonlabial.

3.2. Research methods

Speech data were gathered from 3 age groups of children, at the ages of 2;0, 3;0 and 4;0, years. At the time of testing, subjects were within a 15-day interval of their designated age. Each group included 8 subjects of each gender to allow for cross-gender comparison, accounting for 48 child participant.

Children and their parent(s) were recorded in two 30-45 minute sessions (both sessions were divided into two sub-sessions) occurring within a 14-day time frame. Participants were given 28 puppets (7 at each sub-session) with pre-assigned $C_1V_1(:)C_1V_1(:)$ structured names. To allow for examination of the effects of consonantal environment, all 14 Hungarian vowels were included in two tokens, each one with a different consonant. Target tokens are listed in (2).

(2) Target tokens

Session 1A *Nonsense tokens*	Session 1B *Meaningful words*	Session 2A *Nonsense tokens*	Session 2B *Meaningful words*
/gɑgɑ/	/bɑbɑ/	/ka:ka:/	/pa:pa:/
/bɛbɛ/	/pɛpɛ/	/de:de:/	/le:le:/
/titi/	/pipi/	/mi:mi:/	/pi:pi:/
/toto/	/lolo/	/no:no:/	/lo:lo:/
/pøpø/	/gøgø/	/bø:bø:/	/kø:kø:/
/dudu/	/bubu/	/tu:tu:/	/pu:pu:/
/mymy/	/nyny/	/ly:ly:/	/my:my:/

Parents were asked to involve the child in a free-play situation while modeling and thereby "teaching" the child the puppet names. Elicitation strategies were oriented to imitated and spontaneous naming. Children were expected to readily learn the puppet names and produce them, not only in immediately imitated forms but also by deferred imitation (that is, in spontaneous speech.) Parents were asked to provide the child with as many occasions as possible to produce the required tokens. Caregivers were instructed to make every attempt to have their child produce each token at least five times, if at all possible.

Both sessions were divided into two sub-sessions with a break between them. An attempt was made to generate tokens that contain early acquired speech sounds, in order to keep the task relatively easy even for the youngest children. Meaningful tokens were both disyllabic words and one-syllable morphemes in a reduplicated format.

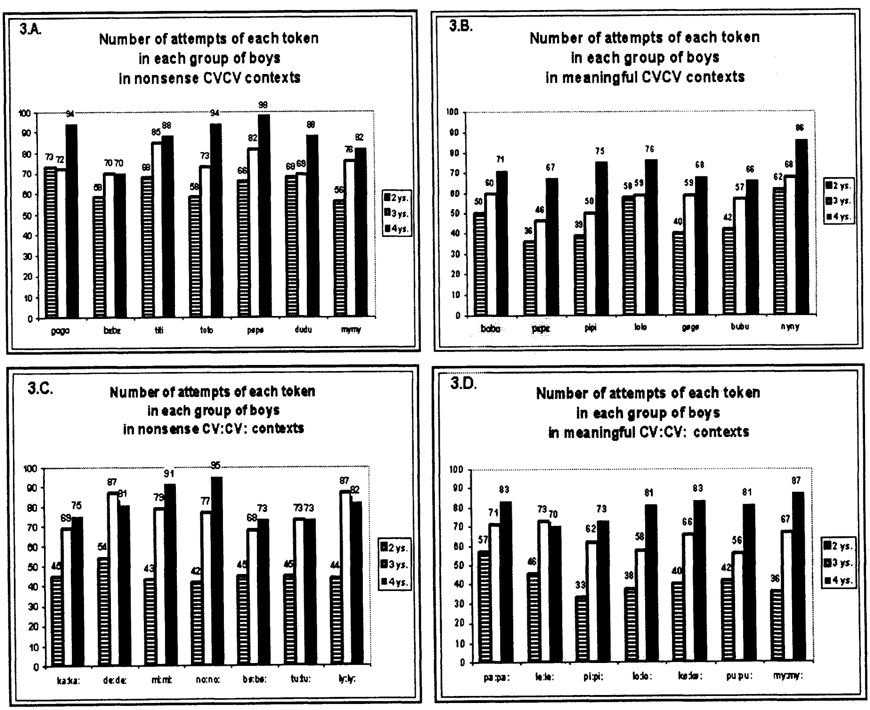
A monolingual native Hungarian speaker analyzed the tokens perceptually. Correct production of a disyllable represented an utterance in which both vowels were perceived to be identical to those in the target token; no consideration was given to consonants. It is argued that acceptable production of both vowels (the first in stressed and the second in unstressed position) is a good indicator of the acquisition of a particular vowel.

3.3 Results

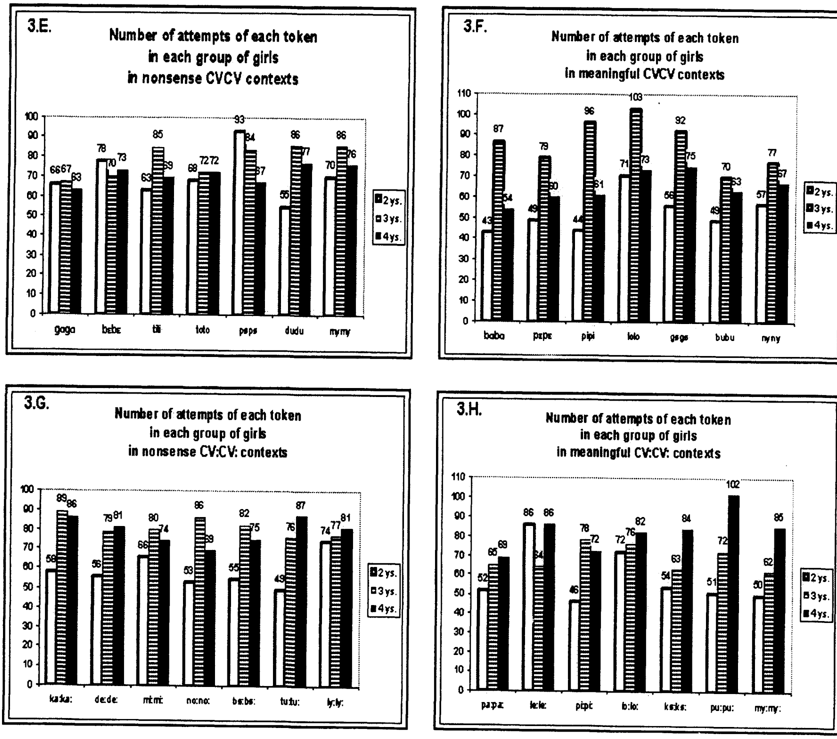
3.3.1. Number of attempts

The data reflect several tendencies of speech development. As expected, in general the youngest children attempted to produce the targeted tokens the least times.

(3a-d) Data for boys displayed by vowel duration and nonsense vs. meaningful targets



(3e-h) Data for girls displayed by vowel duration and nonsense vs. meaningful targets

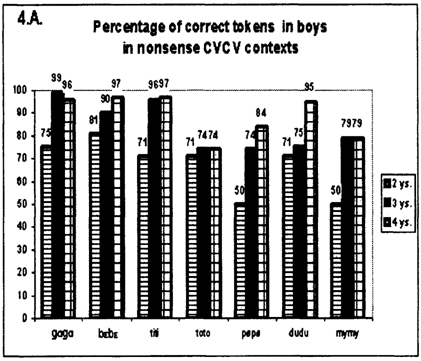


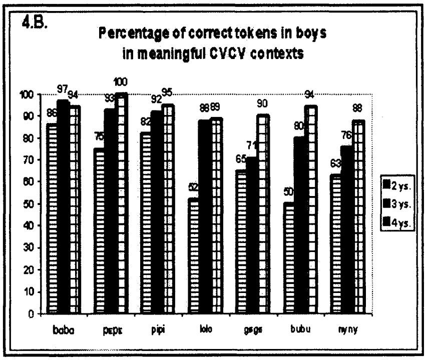
Specifically, there appears to be a major difference in the number of times targets with short vs. long vowels were attempted. While disyllables with short vowels were produced more frequently, puppet names containing long vowels were uttered less. This pattern of "selective production" is well documented in the literature (e.g. Stoel-Gammon and Cooper 1984, Vihman 1976). Ferguson and Farewell (1975) refer to this strategy as "great selectivity … in picking the words [a child] attempts to say" (p. 433). That is, some children only attempt to replicate sound sequences with those speech sounds that they perceive to be part of their own sound repertory. In other words, sounds that are not yet acquired are less likely to be attempted. This acquisition pattern appears to be strongly present in children at 2;0 years; to a lesser extent, it also operates in 3;0 years old children. By the age of four, it seems that the avoidance of producing tokens with long vowels disappears. This developmental pattern probably reflects a maturation process of the speech mechanism that (due to physiological, neurological and speech motor development) becomes increasingly suitable for producing vowels with longer duration. Thus one conclusion that we may draw from the results is that, in the early stages of speech development, the production of long vowels appears to be more challenging than that of the short ones. However, by the age of 4;0 years, children's production level of the two vowel sets appears to be comparable. An avoidance pattern may also provide an explanation for the lower number of attempts at producing tokens with some of the rounded vowels. Further analysis will provide an answer in this respect.

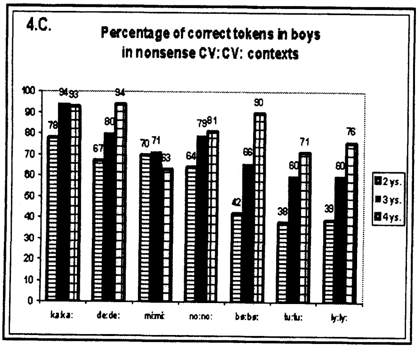
Another tendency in the results is the difference between boys and girls in willingness to produce the target forms. This phenomenon is also strongest at the age of 2;0 years. In general, girls' production data reflect a more developed skill level of vowel production at 2;0 years of age. It is more difficult to formulate a judgment about production level differences at 3;0 years of age on the basis of our data set since the group of 3;0 year-old girls did an outstanding job, as reflected by the high number of attempts through all conditions. Therefore, this question will be revisited at a later point during data analysis. At the age of 4;0 years, however, no major difference between the two genders is reflected by the results. Consequently, if boys appeared to have lower skills at vowel production at the age of 2;0 years as measured by the number of attempted targets, it appears that this gender difference disappears by the age of 4;0 years.

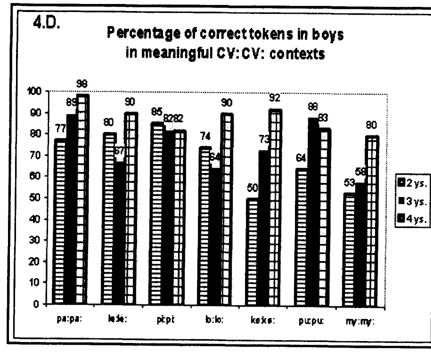
3.3.2 Percentage correct values

(4a-d) Percentage of correct tokens in boys displayed by vowel duration and nonsense vs. meaningful targets

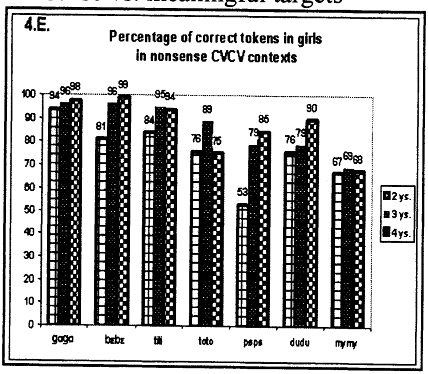


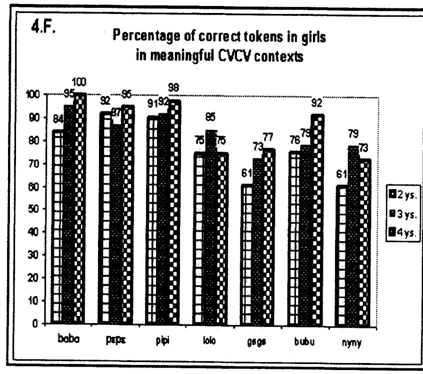


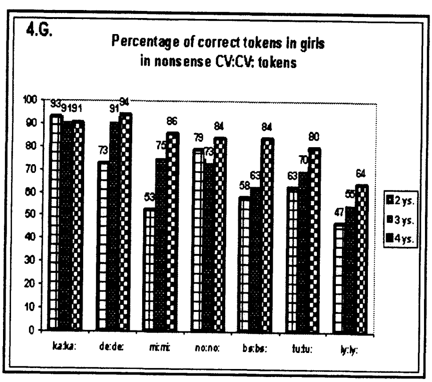


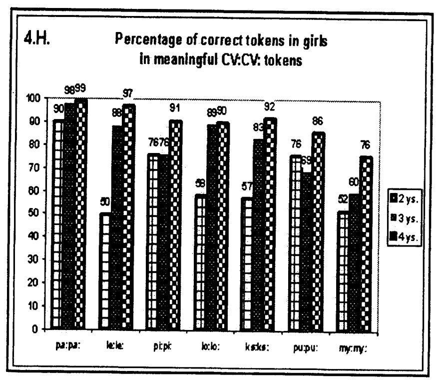


(4e-h) Percentage of correct tokens in girls displayed by vowel duration and nonsense vs. meaningful targets





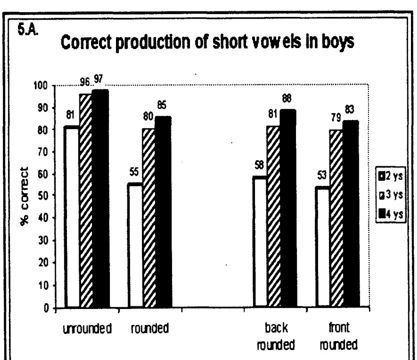


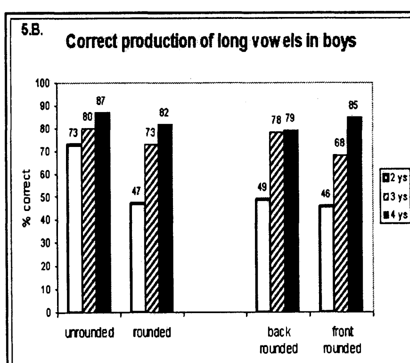


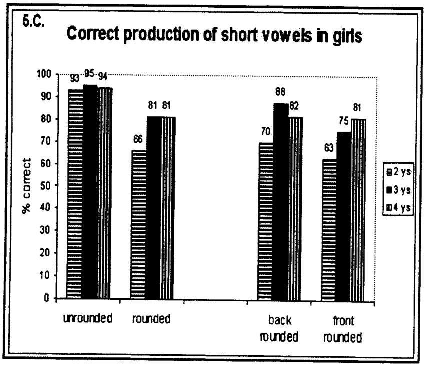
Data displayed in (4) suggest two main tendencies in children's speech development. The first one is that the effect of the sound environment (e.g. consonantal context) is a major determinant of young children's success at reproducing target tokens with the appropriate quality of vowels. For example, the vowel */e:/* (usually acquired later during speech development) has a 50 % success rate in 2;0 year-old girls if surrounded by the voiced lateral approximant */l/* (also a later acquired speech sound); however, with the voiced alveolar stop */d/* (an early acquired speech sound), the same group of girls manages to produce correct vowels in 73 % of productions. In some cases, it is more challenging to reason about the source of difficulty a child experiences due to the effect of consonantal environment on vowels. Even though similar tendencies of consonantal effect on the success rate of vowel production are mirrored in 2;0 years old boys' data, (e.g. the target */tu:tu:/* is successfully produced only 38% of the time; however, the target */pu:pu:/* has a higher success rate of 64%), it is difficult to explain this difference by acquisition order factors. One interpretation of the data is that a higher level of exposure and practice with the latter, meaningful word may result in a higher success rate in vowel production. Alternatively, ease of articulation between a labial stop and a rounded vowel may also be a factor in the higher success rate of */pu:pu:/.* Consonantal effect on vowel production also appears to operate, at a lesser extent, in 3;0 and 4;0 year olds. In short, children's production of vowels is heavily affected by the sound environment in which the vowels are embedded.

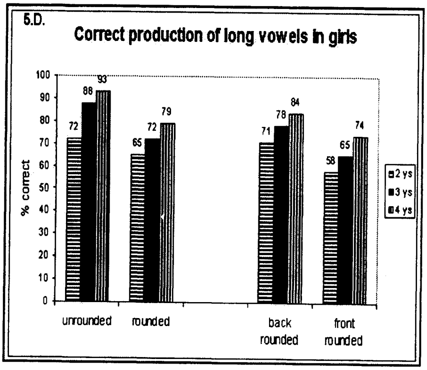
The second main tendency these data suggest is children's difficulty with the production of rounded vowels. Low success rates of rounded vowel production in both genders indicate that lip rounding activity is a major challenge to overcome during vowel acquisition. Re-grouping the data in (5) clarifies the tendency.

(5) Percentage correct values of vowel production in both genders (rounded vs. unrounded; division of rounded into back vs. front rounded sets)









The production of rounding appears to be hardest for boys at 2;0 years, especially in long (as opposed to short) vowels. One interpretation of this finding is that the motor challenges of lip rounding are enhanced by the difficulties of producing a long vowel. These processes, when faced with simultaneously, result in low levels of vowel accuracy, especially in the youngest children.

Difficulties with producing an appropriate level of rounding[1] appear to be substantial in the older age groups as well. In both genders, production accuracy of rounded vowels lags behind that of the unrounded ones. While girls at 2;0 years of age are somewhat more successful with formulating rounded vowels than their male peers, at the age of 3;0 years both genders experience similar difficulties with these sounds. At the age of 4;0 years, both genders' success levels in rounded vowel production are very similar and accuracy is comparable in short and long vowels (another indication of the similar levels of acquisition in short and long vowels).

A further division of the rounded vowels according to tongue position aids a deeper understanding of the matter. The data suggest that the production of front rounded vowels is more demanding for both genders in both vowel sets than that of the back ones. It is to be noted that, in adult speech, the front rounded /y/ and /y:/ sounds seem to require the most exaggerated lip protrusion that may correspond to the most sophisticated skill level of lip posture control. The front vowels /ø/ and /ø:/ and the back vowels /u/ and /u:/ also require a considerable level of lip rounding. Among the rounded vowels, /o/ and /o:/ are the least rounded[2]. Accordingly, in children it is the front vowels that should develop at a slower pace as compared to the back ones, due to a supposedly higher level of skill needed for their production. Overall, while the production of all rounded

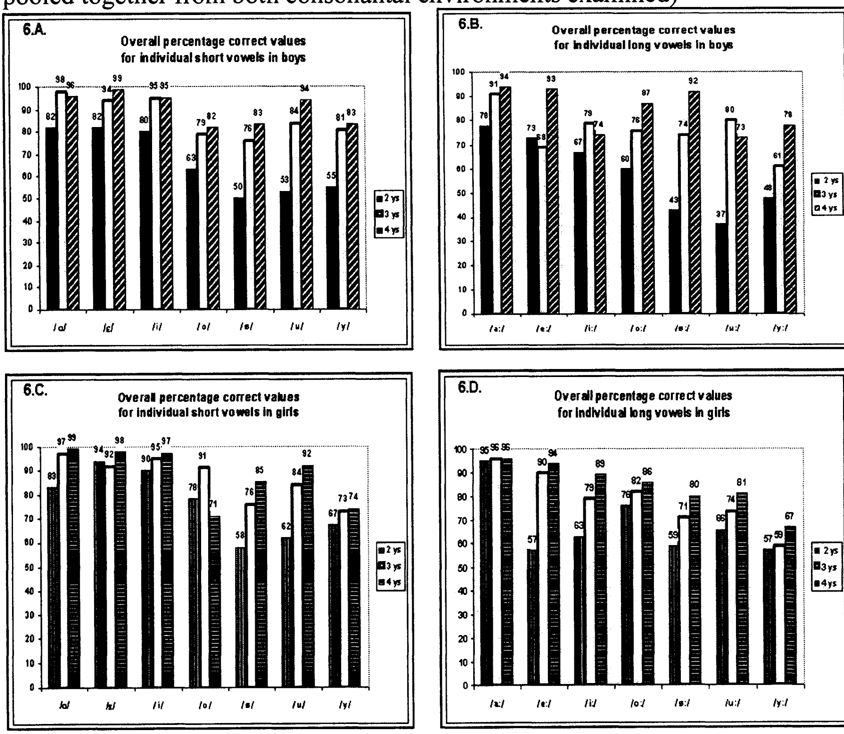
[1] To date, no study has documented the motor development of lip-rounding in children. In adult Hungarian, rounded vowels are "more rounded" than their English counterparts (Szabados 1990).

[2] Even though the vowel /ɑ/ is SLIGHTLY rounded, it frequently lacks rounding in running speech. Therefore, the calculations in this study included it as a non-rounded vowel.

vowels appears to test the motor abilities of children, it is the production of front rounded vowels that proves most challenging.

To develop a general understanding of vowel production development in Hungarian, it is interesting to examine the data set by pooling percentage correct values of individual vowels across different consonantal environments to generate an overall picture of vowel acquisition processes. Of course, the validity of these measures is limited, due to the effect of different consonantal environments on vowel production success. Keeping in mind the limitations of this interpretation, some general tendencies of vowel development may emerge through these calculations (see (6) below).

(6) Overall percentage correct values of vowel production (percent correct values pooled together from both consonantal environments examined)



To aid the interpretation of the data, three (somewhat arbitrary) categories of vowel acquisition level were selected: (1) a vowel that is perceived correct in 65% of the cases or less is in the *"emerging"* stage of acquisition; (2) a vowel that is identified correct in 66 to 89 % of the time is a *"established"* one; (3) a vowel that is identified correct 90 % or more of the time is *"mastered"*. Values in (7) on the next page are rounded to the closest whole number.

(7) Categories of vowel production displayed by vowel duration

7.A. BOYS Short vowels	% correct	2 ys.	% correct	3 ys.	% correct	4 ys.
ɑ	82	ESTABLISHED	98	MASTERED	96	MASTERED
ɛ	82	ESTABLISHED	94	MASTERED	99	MASTERED
i	80	ESTABLISHED	95	MASTERED	95	MASTERED
o	63	EMERGING	78	ESTABLISHED	82	ESTABLISHED
ø	50	EMERGING	76	ESTABLISHED	82	ESTABLISHED
u	53	EMERGING	84	ESTABLISHED	94	MASTERED
y	55	EMERGING	81	ESTABLISHED	83	ESTABLISHED

7.B. BOYS Long vowels	% correct	2 ys.	% correct	3 ys.	% correct	4 ys.
aː	78	ESTABLISHED	91	MASTERED	95	MASTERED
eː	73	ESTABLISHED	69	ESTABLISHED	92	MASTERED
iː	67	ESTABLISHED	79	ESTABLISHED	74	ESTABLISHED
oː	60	EMERGING	76	ESTABLISHED	85	ESTABLISHED
øː	43	EMERGING	74	ESTABLISHED	92	MASTERED
uː	37	EMERGING	80	ESTABLISHED	73	ESTABLISHED
yː	48	EMERGING	61	EMERGING	78	ESTABLISHED

7.C. GIRLS Short vowels	% correct	2 ys.	% correct	3 ys.	% correct	4 ys.
ɑ	84	ESTABLISHED	97	MASTERED	99	MASTERED
ɛ	94	MASTERED	92	MASTERED	98	MASTERED
i	90	MASTERED	95	MASTERED	97	MASTERED
o	78	ESTABLISHED	91	MASTERED	71	ESTABLISHED
ø	58	EMERGING	76	ESTABLISHED	85	ESTABLISHED
u	62	EMERGING	84	ESTABLISHED	92	MASTERED
y	67	ESTABLISHED	73	ESTABLISHED	74	ESTABLISHED

7.D. GIRLS Long vowels	% correct	2 ys.	% correct	3 ys.	% correct	4 ys.
aː	95	MASTERED	96	MASTERED	96	MASTERED
eː	57	EMERGING	90	MASTERED	94	MASTERED
iː	63	EMERGING	79	ESTABLISHED	89	ESTABLISHED
oː	76	ESTABLISHED	82	ESTABLISHED	86	ESTABLISHED
øː	59	EMERGING	71	ESTABLISHED	80	ESTABLISHED
uː	66	EMERGING	74	ESTABLISHED	81	ESTABLISHED
yː	57	EMERGING	59	EMERGING	67	EMERGING

One tendency that is easy to detect is that, in general, boys at the age of 2;0 and 3;0 years are less skilled at vowel production than girls. Girls have already mastered two short and one long vowels at the age of 2;0 whereas boys at this age have not mastered any vowels yet. Rounded vowels are still at the lowest level of acquisition in boys at the age of 2;0; girls already have one vowel pair, /o/ and /o:/, that is well established at this age. At the age of 3;0, level of mastery is still higher in girls, with four short and two long vowels mastered. Boys at the age of 3;0 have the three front unrounded short vowels and a long one acquired. However, at the age of 4;0 vowel production success in the short vowels are identical for the two genders; in terms of vowel acquisition in the long vowels, boys are more skilled than girls. So, by the age of 4;0, the boys who lagged behind during the previous years appear to catch up.

4. Summary

Results suggest that, in general, as it is reflected by data from 2;0, 3;0 and 4;0-year-old children,

- Long vowels are produced with less accuracy than short ones;
- Rounded vowels are produced with less accuracy than unrounded ones;
- At the age of 2;0 years, consonantal environment has a strong effect on vowel accuracy but this effect decreases with development;
- Girls' vowel accuracy is higher than boys' at 2;0 and 3;0 years of age;
- Boys' and girls' vowel accuracy is similar at the age of 4;0 years;
- Mastery of vowels is not complete in either gender at the age of 4;0 years.

Acknowledgements

This research was supported by the Graduate School Fund for Excellence and Innovation and the Center for Mind, Brain and Learning at the University of Washington. I thank my advisor, Dr. Carol Stoel-Gammon for her continuous support and enthusiasm for this project. I am indebted to Dr. Pat Kuhl for providing support and equipment for this study. Special thanks are due for generous advice to Drs. Klára Vicsi and András Illényi at the Laboratory of Speech Acoustics at the Department of Telecommunication and Telematics of the Budapest University of Technology and Economics.

References

Ball, Martin J. & Gibbon, Fiona E. 2002. Preface. In M.J. Ball & F.E. Gibbon (eds.) *Vowel disorders.* Boston, MA: Butterworth-Heinemann.

Ferguson, Charles A. & Farewell, Carol B. 1975. Words and sounds in early language acquisition. *Language* **51**: 419-439.

Kassai, Ilona. 1979. Időtartam és kvantitás a magyar nyelvben [Duration and quantity in the Hungarian language]. *Nyelvtudományi Értekezések* 102. Budapest: Akadémiai Kiadó.

Stoel-Gammon, Carol & Cooper, Judith A. 1984. Patterns of early lexical and phonological development. *Journal of Child Language* **11**: 247-271.

Szabados, Márta. 1990. Articulatory and acoustic comparison of the Hungarian and English vowel systems. In Gy. Décsy & A.J.E. Bodrogligeti (eds.) *Ural-Altaische Jahrbücher/Ural-Altaic Yearbook* **62**: 7-30.

Szende, Tamás. 1999. Hungarian. In *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet.* Cambridge, UK: International Phonetic Association.

Tarnóczy, Tamás. 1974. A magánhangzók akusztikai vizsgálatának problémái [Problems encountered during the acoustic examination of vowels]. *Általános Nyelvészeti Tanulmányok* **10**: 181-196.

Vihman, Marilyn M. 1976. From pre-speech to speech: On early phonology. *Stanford Papers and Reports on Child Language Development* **12**: 230-244.

University of Washington
Department of Speech & Hearing Sciences
1417 NE 42nd Street
Seattle, WA 98105

zajdo@hotmail.com

PARASESSION:
FIELD LINGUISTICS

Fieldwork as a Participatory Research Activity: The Mayangna Linguistic Teams

Elena Benedicto, Modesta Dolores, Melba McLean
Purdue University, URACCAN, CIDCA

0. Introduction

In dealing with endangered languages, focus has been shifting from the languages themselves and their scientific documentation to the communities that speak such languages. This paper is an attempt to examine how linguistic, scientific research can be conducted in a way that actively involves and empowers the target community—an approach which falls under the umbrella of what may be known as *participatory research.*

In this paper, we will address the particular goals and theoretical bases of a participatory-research-based approach to linguistic research (§1 and §2), we will present some of the projects that have been or are presently being conducted in Nicaragua (§3), and we will evaluate the preliminary results of those projects (§4 and §5).

1. The Goals of the Approach

When considering the variable of the community, doing linguistic research on an endangered language presents a series of issues having to do with power, control, and agentivity. The last of these issues—that of agentivity—is perhaps the most obvious one. The agent of the research is the linguist, usually an outsider to the community. The members of the community have little or no agentive behavior, but rather are simply "informants." Issues of control have to do, among other things, with questions like what is being done with the elicited materials, who has access to them, and how will they be used in the short and the long term. The first question is perhaps the most complex, and one which we will not even attempt to address as it deserves. For the purposes of this paper, let us suggest just one venue of thought: as one informant once put it, "If they don't explain anything to me, if they just use me as an informant [i.e., if there is no transfer of knowledge], I don't gain anything, I remain the same, what's the gain?" Knowledge allows people to make their own informed decisions about issues that affect them, such as language policies. In other words, knowledge is power; without it, the power to decide remains with others.

The general goals of the approach to linguistic research that we are suggesting here are, first, to empower the community by jointly (re)creating knowledge (thus, addressing issues of imbalance of power) and, second, to emphasize the documentation and preservation process as one for the benefit of the language community as well as of the scientific community (along the lines of suggestions such as those of Deloria (1969), Hale *et al.* (1992), Bach (1995)). More concretely, the specific goals are: (i) to involve the community more successfully in the linguistic activities having to do with their language; (ii) to establish a more equal-to-equal relationship between the community members and the linguist; and (iii) to transfer the linguistic "technology" back to the community, i.e., to foster an environment for native linguists to emerge.

2. The Ideological/Theoretical Bases

The goals established in the previous section are in line with what has been called Participatory Action Research (PAR), an approach developed in the (social) sciences in the context of environments of social imbalance (see Christie *et al.* (2000)). The approach taken by PAR intends, specifically, to reduce the power imbalance between "scientist" and "object of study"; to transform the relation "subject-object" into a "subject-subject," that is, to involve local members as agents in the process of knowledge generation; and to create knowledge that is useful, that can effect change in the lives of the original "objects." Research, then, is viewed as the articulation of a co-learning process that takes place at all levels (crucially involving what are now two "subjects"), and the subsequent change that it effects.

Christie *et al.* (2000) propose a broad methodological process, which includes three basic steps. First, research priorities need to be set and key problems and issues identified. Second, the causes underlying those problems and issues need to be analyzed. Finally, action needs to be taken to find both short-term and long-term solutions for the identified problems. We will use this approach with respect to the linguistic situation the Mayangna face.

In the context of the indigenous communities of Nicaragua and, more specifically, in its Northern Atlantic Region (RAAN), the Mayangna find themselves in a situation that, while not unique in its broad features, did present some idiosyncracies with respect to other indigenous or local communities in the area. The Mayangna are a relatively small indigenous community inhabiting the North Central area of Nicaragua, part of what is known as the RAAN. The language is spoken by a little over 10,000 speakers at varying degrees of endangerment (though it is strong and healthy for a good core of speakers). It is a member of the Misumalpan family, an isolate group (see Craig & Hale (1992)), and three dialectal variants are presently known: Tawahka in Honduras, and Tuahka and Panamahka in Nicaragua (see Benedicto & Hale (2001)). A bilingual program for primary school was begun in the early 1980's in Nicaragua and thus began a structural need for decision making in the linguistic realm. The projects described in §3 began in the late 1990's, after more than a decade of bilingual programs, thus inheriting a certain structural set up. All of these projects share a basic set of conditions, as established by Christie *et al.* (2000) (see above). First, the research priority was to enlarge the body of knowledge about the language as needed for primary school materials such as dictionaries and textbooks, but also for the particular linguistic contents to be included in the curriculum. Additionally, the teachers' rising level of training required new materials, such as grammars. Finally, the revision of school curriculum required culturally relevant materials such as collections of oral histories, stories, and traditions. The key problem identified for this project was the lack of local language experts. The most immediate cause for this situation was the absence of appropriate, locally available training mechanisms. The projects described below were born, as short- and long-term solutions to this situation, within a more general framework of creating a model of education relevant for the Atlantic Coast of Nicaragua.

3. The Projects

These projects are carried out in Nicaragua in conjunction with a series of institutions that work regularly in the area of the Atlantic Coast of Nicaragua (RAAN and RAAS), all of them committed to the development of bilingual intercultural education. These institutions are URACCAN (the University of the Atlantic Coast of Nicaragua), CIDCA (the Center for Research of the Atlantic Coast of Nicaragua), several NGOs involved in bilingual and intercultural education (the Italian Terra Nuova, Danish Ibis, Finnish Kepa, and Austrian OED), and last but not least the Program for Bilingual Intercultural Education, PEBI, a division of the Ministry of Education. In initiating and developing a program of linguistic research, the role of Ken Hale and the group Linguists for Nicaragua was absolutely instrumental beginning in the mid to late 1980's.

3.1. The Tuahka Project: TUYUWAYABA (1995-present)

Tuahka is the minority dialectal variant of Mayangna spoken by some 1,000 people around the area of Wasakin, near Rosita, RAAN (see Benedicto & Hale (2001) for an overview of its morphological properties). It is more endangered than its counterpart, Panamahka. In 1995, following a request by some of the members of the community, a group was formed to carry out an overview of the dialectal variant. This initial group grew into the team later called TUYUWAYABA, *Tuahka Yuln Walwi Yakwi Balna* 'Group for the Research of the Tuahka Language'.

3.1.1. Initial Problem(s)

The situation that prompted the members of the Tuahka community to organize into this group can be summarized as a lack of representation of this linguistic minority in the structure of the educational system. All of the materials and textbooks in PEBI were written in Panamahka, the majority variant, and there was no Tuahka member among the technical staff of PEBI (the bilingual section of the Ministry of Education).

3.1.2. Goals

The goals of the team were (i) to document the Tuahka linguistic variant, (ii) to train a group of technical staff in the linguistic "trade" (e.g., collection and handling of data), (iii) to pilot-produce some school materials in Tuahka, and (iv) to collect a set of culturally relevant materials (oral history, folktales, oral traditions, etc.).

3.1.3. Participants

The group is formed by eight members: three women and five men from the community of Wasakin. A member of the team has been, on and off, included in the PEBI technical staff. Unfortunately, however, the Ministry of Education has not yet officially recognized that person as the Tuahka representative.

3.1.4. Accomplishments

The efforts of the team have coalesced into three types of products. First, they produced a preschool Tuahka textbook, which was published and is currently in use; other textbooks (1st and 2nd grade) were also prepared but were not published (though they will be considered for the new curricular plan). Second, supporting cultural material has been collected: twenty-one folk stories (of which ten are currently being published), as well as a collection of oral history and traditions. Finally, all of the lexical material contained in the stories and folktales has been prepared as lexical entries for a dictionary. Originally, a small vocabulary was produced in 1996 (*Breve Vocabulario Twahka*); an expanded version is presently being included into the *Mayangna Dictionary* (see *§*3.3).

3.1.5. Funding

Funding for TUYUWAYABA has been provided by The Irish Student Group, Linguists for Nicaragua, the Foundation for Endangered Languages, and URACCAN University in Nicaragua.

3.1.6. Evaluation

As a result of these actions, two types of outcomes can be identified. One concerns what could be dubbed as "positive results" and another one as "new issues/problems identified."

In the first category we can list the fact that the dialectal variant has been studied and the community now has a "linguistic evaluation" of itself (which was the original concern of the community). More importantly, the community's "linguistic right to existence," so to speak, has been acknowledged. The PEBI section now has a permanent Tuahka staff member (though not recognized as such by MED). Furthermore, a considerable amount of culturally relevant material has been collected (though not completely published). Finally, the group is beginning to set up new goals and organize activities on their own, without the participation of the outsider linguist.

In the second category of outcomes we can include the realization of the role of old, well-established dynamics. On the one side, internal divisions within the small community, reproduced in the group, can paralyze its smooth functioning. On the other side, we could find an external factor: the role of the community at-large and the local politics at play. Old rivalries between the majority group Panamahka and the minority Tuahka have found their way into the linguistic debate, though it seems that, at least up to now, they are being treated in the climate of an open and civilized conversation.

3.2. The Women's Project (1997-2000)

3.2.1. Initial Problem

There were no women working as technical staff in PEBI. Generally speaking, very few women were trained to carry out a position of responsibility, though a number of them were teachers or teachers-in-training, and at least one was a university student.

3.2.2. Goals

The initial goal of the project was to train women in linguistic techniques so that they could eventually opt at positions as technical staff members at PEBI in matters related to language. More specifically, they were to be trained in how to collect linguistic data (taping, transcribing, etc.); how to prepare a text for publication (using punctuation, standard orthography, etc.); and how to prepare a lexical item as a dictionary entry.

The broader goals of the project were to develop leadership and organizational skills in the members of the team and to prepare women to successfully compete for well-paying and well-regarded jobs.

3.2.3. Participants

A total of five women participated in the program. One of them worked as the local coordinator of the team. She was a speaker of Panamahka and one of the few Mayangnas—and the only woman—with a University degree. Two other women in the program were also speakers of Panamahka (one in Awastingni/Bilwi, one in Sakalwas), and the remaining two participants were speakers of Tuahka (in Wasakin).

3.2.4. Accomplishments

The team has produced two collections of women's life stories, one of them already published (Benedicto *et al.* 1999). Two illustrated dictionaries for preschoolers and 1st and 2nd graders have also been produced and are now awaiting publication. These are bi-dialectal dictionaries (in Panamahka and Tuahka), the first of their kind in the Nicaraguan context. Finally, lexical data in the women's stories were prepared as lexical entries for the *Mayangna Dictionary* (see §3.3).

3.2.5. Funding

Funding for this project came, originally, through the *Seminari d'Estudis de la Dona* ('Women's Studies Seminary') (SIED) via a cooperation program of the *Paeria de Lleida* (the City Hall of Lleida) in Catalunya, Spain. Further funding was provided by the Endangered Languages Foundation at Yale University and by the Linguistics Program at Purdue University.

3.2.6. Evaluation

As with the previous project, two sorts of outcomes can be identified—one that acknowledges the positive effects of the project, and one that discovers new problems or issues.

On the positive side, we can cite the fact that the women in the group now know and use their linguistic skills fully. One of the women is now working as a technical staff member in the PEBI and another was named director of the Bilwi office of CIDCA. A third participant has organized a new independent research group and has initiated several projects.

The new problems and issues that were identified concern, once again, questions of the internal politics of the community, more concretely, issues of power and the role of women within the community. In fact, this is not an unexpected problem but a common outcome of any kind of activity that questions the roles traditionally assigned to women. As an internal process within the community, it needs to follow its course of discussion. With respect to the issue of building self-confidence, the project was quite successful in its internal dynamics. However, once the women were integrated into mixed, male dominated groups, their participation dropped dramatically. Finally, the questions arose of what to do with the collected texts, who has "ownership" of them, who has control over them, who has access to them and under which circumstances, etc. The women who told their stories agreed that they be published for use in the bilingual system. Now, some groups have requested to have translations made in other languages, but some of the women may not be alive anymore for consent. These questions only add to the growing need for establishing guidelines that recognize both individual and community rights to the linguistic materials collected during research.

3.3. The Linguist Team Project (2000-present)

3.3.1. Initial Situation

As already pointed out at the end of §2, one of the key problems identified was the lack of local, indigenous linguists. As a result, the community was not self-sufficient and had to rely on outsiders for linguistic matters ranging from setting consistent orthographic standards, to preparing and revising textbooks for the bilingual program, to editing and publishing texts.

3.3.2. Goals

In preparing this project, three goals were established. The first, general, long-term goal was to get the language community in a position to be self-sufficient in linguistic issues. In order to achieve this, the second, concrete, goal was to create and train an indigenous team of future linguists. The duties of such a team would be to assist the bilingual program (PEBI) in linguistic matters, to assist other institutions in the publication of materials in Mayangna (e.g., with respect to orthographic standards), and to develop deeper knowledge of their own language (that is, to become involved in linguistic research). Finally, the third, specific goal was to obtain reference materials on Mayangna relevant for the community: bidialectal monolingual dictionaries (with definitions in Mayangna and lexical entries in the two dialects), teaching grammars, etc.

3.3.3. Participants

A total of nine people are participating in this project—six women and three men. Six of them are teachers in the bilingual program, PEBI. The regional coordinator is a speaker of Panamahka; she is also the director of the CIDCA office in Bilwi. Five of them are speakers of Panamahka and three are speakers of Tuahka. Except for the regional coordinator, who already has a University degree, all of them are students in the Bachelor's Degree program for Bilingual Intercultural Education conducted by URACCAN University (see below, §3.4).

This project has fed from the experience and participation in the two previous projects. Five of the six women in this project participated in the Women's Project and, thus, had quite a bit of experience in these matters. Three of them are Tuahka speakers and had also participated in the Tuahka Project, so their experience was double.

3.3.4. Activities

Work has concentrated on creating and revising lexical entries for a bi-dialectal (Tuahka and Panamahka), monolingual (with definitions in Mayangna) dictionary. The dictionary is currently in the final revision stage. The next project is a grammar for teachers in the bilingual program, written in Mayangna.

The members of the team have worked both individually and as a group. They meet monthly for self-evaluation and feedback and they have two training workshops a year, where the previous phase is evaluated and goals for the next one are set.

3.3.5. Funding

The main sponsor of the project is CIDCA, with funds from AID/MED and substantial support from SAHWANG (an organization formed by the NGO's Terra Nuova, Ibis, and Kepa). The Linguistics Program at Purdue University has also provided support sporadically.

3.3.6. Evaluation

On the positive side, the level of training has indeed risen and, thus, the level of linguistic sophistication, a substantial step towards the initial goal of creating an indigenous local team of linguists. In the same way, the level of awareness and ability to identify problems has grown, as well as the level of personal commitment to the project among the participants (i.e., the ability to see the project as one's own, and not as one more outsider's project). Finally, the first volume of the dictionary is about to come out and the second volume will soon follow.

The problems and issues that were identified are not new. In fact they had already been talked about in Christie *et al.* (2000). Though the level of organizational and leadership skills augmented over the course of the project, it became painfully obvious that organizational training was far more important than originally thought; it can never be stressed enough how this point can affect the outcome of a project. Furthermore, faction and gender power struggles internal to the community at large were reproduced inside the group, though one must acknowledge that as the group's awareness and sense of identity grew, those factors decreased. At another level, one could mention the difficulties posed by old, deeply rooted pre-conceptions, in particular, the idea that "the foreigner knows best and is the one to decide on issues," and also the difficulty recognizing that the community is itself a source of knowledge and that they themselves can create knowledge. Finally, certain economic expectations and dependency on outsiders were created that led some to consider the project more as source of easy money than as a source of empowerment and intellectual growth. This was by no means a product of this particular project, but rather a reflection of the general climate and attitudes in the region towards external cooperation.

All of these issues do nothing else than underscore the fact that this is an ongoing process of co-learning and that, through the identification of new "problems," better solutions can be found.

3.4. Other Projects (1996, 1998-present)

The picture sketched here would not be complete without mentioning the role that the Bachelor's Degree in Bilingual Intercultural Education has had and has in providing crucial logistic, ideological, and intellectual support for the task of building a local scientific community. The degree program offered by URACCAN (the University of the Atlantic Region and Caribbean Coast of Nicaragua) has provided the pool of students to participate in these projects and has fostered the intellectual and participatory climate without which they would not have flourished.

Additionally, the Mayangna Girls and Women's Project (http://icdweb.cc.purdue.edu/~benedict/sumu), an independent project administered by CIDCA-UCA, has worked towards augmenting the pool of women feeding the university, by increasing their participation in secondary education.

In sum, as already mentioned at the beginning of §3, the projects referred to in this paper have benefited greatly from a general climate of cooperation and mutual knowledge creation among different institutions and organizations.

4. General Evaluation

In evaluating these projects as steps or modules within a single multi-focal process, itself part of a larger task of creating a multilingual intercultural model of education relevant for the Atlantic Coast of Nicaragua, some recurrent patterns emerge. Let us consider them in turn.

Though the specific and concrete goals were achieved (the dictionary is in the process of being published, culturally relevant material has been or is being published, some members of the community are better trained in linguistics and are continuing the process), it became obvious that attention needed to be shifted to, or at least needed to include, the process itself. The generation of a product (such as a dictionary) is merely an indicator, but not the only one. In evaluating the process, three variables emerge, which, of course, should not be considered an exhaustive list.

First, more detailed organization of the work is needed, especially under the current circumstances, where members of the team are not in permanent contact. Both in the case of the Women's Project and of the Linguists Team, the members were scattered throughout a number of distant and often difficult to reach communities. Local members met once a month and the outsider linguist met with them twice a year, though telephone and e-mail contact was maintained when and where possible.[1] Though a plan of work was designed at the beginning of a six-month period and evaluated at every monthly meeting, more detail in the planning, in the actions to be taken, and in the treatment of contingencies would have gone a long way towards improving the co-learning process. A second lesson learned was that it is equally important to assign individual responsibilities as it is to assign group responsibilities. The former increase the participants' sense of self-worth and of accomplishment, the latter emphasize the aspect of being part

[1] Local (URACCAN, CIDCA and Terra Nuova) offices in Managua acted as points of contact for telephone and e-mail. Lately, too, Bilwi has obtained e-mail access by satellite. The towns of Rosita and Bonanza (closer to some of the communities) had telephone lines, however unreliable.

of a larger more far-reaching and long-term project. Though this is in principle a good assessment of the situation, it must be said that it interacts with variables as of yet poorly understood. More concretely, increasing individual responsibilities greatly improved the dynamics of the group in the case of the TUYUWAYABA project, while it did nothing for the Linguists Team. In the case of the Women's Project, individual and group responsibilities worked in near-perfect combination. The third observation, which interacts with this last one, has to do with introducing more specific and process-related self-evaluation mechanisms, especially some that would push the members to look for causes and possible solutions and improvements.

These observations are by no means intended to be exhaustive. But they do give us a good point of the departure to re-evaluate the process and make the necessary improvements. It is, no doubt, an on-going process.

5. Final Remarks

In this paper we have presented the experience of three different linguistic projects, where the role of the native speaker went beyond that of mere informant to become an agent in the process of uncovering knowledge. Though an outsider linguist was present, the whole process took the shape of a co-learning process where each member learned from the others and from the process itself. The goal of such an approach was to create a more equal-to-equal relationship between the different participants of the linguistic research in the context of an endangered language. Within a background framework of participatory action research, a concomitant goal was that of effecting change in a situation of disadvantage, that is, to empower the community. Have these goals been attained? Have these actions effected any substantial change? Is the community more empowered as a result of such a process?

These are difficult questions to answer right now. Obviously, these projects did attain the specific goals that were set up (see §1): (i) the community has become more involved in the linguistic activities, by having some of its members actively participate in them and by having town meetings ('*validaciones*') where the materials were presented (i.e., returned to the community) and evaluated by the community itself; (ii) presumably, a more equal-to-equal relationship was established, since the local members were able to make their own decisions in the process, though the role of the outsider linguist was still distinct; and (iii) there was transfer of "linguistic technology" in the sense that a group of native linguists is emerging.

Obviously, there has been change, but is the community at large more empowered? The challenge now is to integrate those changes more fully into the community itself, without provoking a loss of cultural identity. Though most of the work in these projects came out of the communities, it is still difficult to be self-sufficient, since most do not have electricity or are difficult to access.

Moreover, some of the members of the teams have been offered jobs at institutions working out of town, so in a way there has been a sort of brain drain.

It cannot be stressed enough that this is an ongoing process, where all of us are responsible agents and where continuing self-evaluation is a tool of utmost importance to keep at the task of working towards our goals.

References

Bach, E. 1995. Endangered languages and the linguist. In Ik-Hwan Lee (ed.) *Linguistics in the Morning Calm 3: Selected Papers from SICOL 1992*, 31-43. Seoul: Seould Hanshin Publishing Company.

Benedicto, E. & K. Hale. 2001. Mayangna, a Sumu language: Its variants and its status within Misumalpan. In Benedicto, E. (ed.) *The UMass Indigenous Languages Volume*. UMOP [University of Massachusetts Occasional Papers] 20. University of Massachusetts, Amherst.

Benedicto, E., M. McLean, M.Dolores, A.Jacobo, C.Poveda & E.Salomón. 1999. *Mayangna Yayal Sangninaka Yulni / Mayangna Wawana Yalalahwa Yuln.* ('Mayangna Women's Life Stories'). Managua, Nicaragua: OED and Terra Nuova.

Breve Vocabulario Twahka. 1996. CIDCA, MIT, UMass.

Christie, P., D. Bradford, R. Garth, B.González, M.Hosteler, O.Morales, R.Rigby, B.Simmons, E. Tinkam, G.Vega, R.Vermooy, & N.White. 2000. *Taking Care Of What We Have. Participatory Natural Resource Management On The Caribbean Coast Of Nicaragua.* Managua and Ottawa: Centro de Investigación y Desarrollo de la Costa Atalántica (CIDCA) and International Development Research Centre (IDRC).

Craig, Colette, and Ken Hale 1992. A possible Macro-Chibchan etymon. *Anthropological Linguistics* **34**:173-201.

Deloria, V.Jr. 1969. *Custer Died For Your Sins: An Indian Mannifesto*. New York: McMillan.

Hale, K. 1996. UG and the roots of linguistic diversity. In J.D. Bobaljik, R. Pensalfini & L.Storto (eds.) *Papers on Language Endangerment and the Maintenance of Linguistic Diversity*, MITWPL **28**:137-161.

Hale, K., M. Krauss, L. Watahomigie, A. Yamamoto, C. Craig, L. Masayesva Jeanne & N. England. 1992. Endangered languages. *Language* **68**:1-42.

Mayangna Dictionary. in prep. CIDCA.

McLean, Melba. 1996. *Diccionario Panamahka. Sumo-Español-Sumo*. Managua, Nicaragua: CIDCA.

Norwood, Susan. 1988, 1997. *Gramática de la Lengua Sumu*. Managua, Nicaragua: CIDCA.

Elena Benedicto
Linguistics Program
Heavilon Hall 307B
Purdue University
West Lafayette, IN 47907

benedict@omni.cc.purdue.edu

Modesta Dolores
c/o IPILC – URACCAN
Radio Rosita
Rosita, RAAN (Nicaragua)

ipilc.uraccan@ifxnw.com.ni

Melba McLean
CIDCA
del campo de beisbol, 50v. al mar
Bilwi, RAAN (Nicaragua)

estermelmclean@yahoo.com

Tracing Dialect Death: The Texas German Dialect Project

HANS C. BOAS
University of Texas at Austin

0. Introduction

This paper reports on current efforts of the Texas German Dialect Project to record, archive, and analyze the remnants of the rapidly eroding Texas German dialect spoken in central Texas.[1] The paper is structured as follows. Section one gives a brief history of German settlements in central Texas. Section two presents in detail the on-going activities of the Texas German Dialect Project. They include conducting fieldwork in three representative speech communities and the creation of an on-line digital archive of Texas German. Section three discusses the change of a selected number of lexical, morphological, syntactic, and phonological features of Texas German that took place over the last four decades. More specifically, it compares data from thirty to forty years ago with data recorded during current fieldwork in fall of 2001 and spring of 2002. Finally, section four gives an overview of research questions that the Texas German Dialect Project will be tackling as it will be collecting and analyzing more data over the coming years.

1. Brief History of German Settlements in Central Texas

The Texas German speakers who live across central Texas are in large part the descendants of settlers that started emigrating from Germany beginning around 1830 (Biesele 1928, Salmons 1983).[2] The first large group of about 5000 settlers came to Texas in 1845/46 as part of an organized effort by the "Society for the Protection of German Immigrants to Texas" that was organized by a group of German noblemen with the goal of systematically settling Germans in Texas.

The different dialects of the settlers' native homes in geographically diverse places such as Hesse, Hesse-Nassau, Rhenish Prussia, Westphalia, and Hannover

---

[1] The Texas German Dialect Project (TGDP) gratefully acknowledges the financial support of the Dean of Liberal Arts, the Vice President for Research, and the Center for Instructional Technologies, all of the University of Texas at Austin.

[2] According to Gilbert (1965), the majority of Texas German speakers "are located in the eastern part of the state between Austin and Houston and on the eastern part of the Edwards Plateau west of Austin and San Antonio" (1965: 102).

formed the basis from which the dialects of New Braunfels, Fredericksburg, and other Texas German communities evolved (cf. Eikel (1949: 278), Gilbert (1965: 102)). Because of their isolation at the western frontier, these towns became important regional cultural centers which enabled German immigrants and their descendants to maintain their culture and language for an exceptionally long time (Salmons 1983, Guion 1996, Boas, in press).

However, due to the wave of anti-German sentiment caused by the two World Wars, the prestige of German in the United States suffered severe blows and the generally stable linguistic situation of Texas German began to collapse.[3] The social, demographic, and economic changes that took place in central Texas following World War II led to schooling exclusively in English in a culture that had an extensive system of German schools.[4] Moreover, English gained in prestige among younger speakers because of the practical and economic advantages associated with being primarily English-speaking (Salmons 1983, Guion 1996, Boas, in press).

The decline in prestige associated with speaking Texas German led to the dialect not being transmitted to children to any significant extent. Whereas for generations Texas German was acquired in early childhood at home, this form of transmission disappeared in the years following World War II. The continuation of these trends over the last five decades has resulted in a sharp decrease in the number of fluent Texas German speakers, thereby causing a rapid language shift to English in the Texas German community. At present, English has become the primary language for most Texas Germans in both private and public domains, whereas the reverse would still have been true as late as the 1940s (Salmons 1983, Boas, in press). As Guion (1996) points out:

> The last two environments outside the home that afforded a possibility of speaking German, namely business transactions and church, have been lost. Presently, the only surviving register is an informal, familial one. It is important to note here that German is strictly used for oral purposes. The only written German found in German-speaking homes is an occasional family Bible. English is the sole language written by the people interviewed (p.447).

The remaining number of fluent Texas German speakers in central Texas, most of whom are age 60 or older, is estimated to be between 6-7000 (compared to ca. 70,000 in the early 1960s, according to Gilbert (1965: 102)) (see Boas (in

[3] See Eikel (1949) and Salmons (1983) for a more detailed discussion.

[4] One of the most influential measures affecting the prestige of Texas German has probably been the passing of English-only laws by the Texas Legislature at the beginning of World War I. This legislation required English to be the language of instruction in public schools as well as for all official transactions (Guion 1996: 444). Although in 1938 the English-only policies were loosened to allow for instruction in German above the second grade, the prestige of Texas German had already suffered severe blows such that "the young people [who] had been schooled in English (...) could much more easily identify with the more prestigious English and disappear into the melting pot" (Salmons 1983: 188).

press)). The rapidly shrinking number of fluent Texas German speakers as well as the limited domains and registers of use puts this dialect on the list of about 1200 moribund dialects world-wide that are expected to go extinct within the next 20-25 years (see Crystal (2000)).

2. The Texas German Dialect Project (TGDP)

As a response to this situation, the Texas German Dialect Project (henceforth TGDP) was founded in September 2001 with the goal of recording, archiving, and analyzing the remnants of the rapidly eroding Texas German dialect. The following sections give an overview of the work flow of the TGDP. The first step in setting up the TGDP consisted of collecting and reviewing all publicly available previous work on Texas German. The purpose of this preliminary phase was to identify the relevant linguistic features of Texas German recorded by earlier studies[5] in order to set up procedures that would allow for eliciting as much relevant data from informants as possible.

2.1. Design of Questionnaire

The second step was to derive a strategy that would allow for a broad-scale collection of natural data representing the largest possible number of linguistic features of Texas German. One option considered for collecting data consisted of using structured word lists and sentences to be translated by informants from English to German. In addition, structured word lists and sentences in German were considered for serving as a basis for eliciting information on how informants pronounced words and sentences differently.

However, due to their narrow scope (each covered only a limited number of words and sentences), both methods would have only yielded a limited amount of linguistic information. Another disadvantage of this approach is that the data collected would not have been elicited in a natural setting. That is, the informants' daily use of Texas German would not have been replicated or understood by simply having them read or translate structured word and sentence lists.

In order to overcome these problems, an eight page long questionnaire was drafted to serve as a basis for sociolinguistic interviews that are more thorough. The first section of the questionnaire contains questions about informants' personal history (date and place of birth, place of origin of informants' ancestors, etc.). The second section of the questionnaire consists of about 140 questions in German about topics including childhood activities, the community, religion, education, living conditions, tourism, government, language, and current activities. The goal of these questions is to produce casual, relaxed conversation in which informants are given the chance to talk naturally. The final section of the

[5] For an overview, see the works cited in the bibliography at the end of this paper.

questionnaire contains twenty English sentences that informants are asked to translate into German.[6]

Thus, conducting sociolinguistic interviews with open-ended questions enables data to be gathered on informants' natural speech, as opposed to the artificial/simulated speech imposed by the structured word and sentence lists of previous studies (e.g., Eikel (1967), Gilbert (1972)). Furthermore, by allowing informants to speak freely it becomes possible to discover new linguistic features of Texas German that may have previously gone unnoticed because elicitation methods for them were not included in the research methodology of previous studies.

2.2. The interview process

Since the majority of previous studies are concerned with the Texas German varieties spoken in and around Fredericksburg (Gilbert 1963, 1965, 1972, Salmons 1983, Guion 1996) and New Braunfels (Eikel 1949, 1954, 1965, 1967), both towns were natural choices in which to begin fieldwork. Informants were found through a social network tracing process beginning with students at the University of Texas at Austin. Through students enrolled in my classes during the academic year 2001/2002 I was able to make contact with their family members or friends who were at least third generation Texas German speakers residing in Fredericksburg and New Braunfels. Furthermore, colleagues in my department helped me to establish contact with informants in two other locations, namely Round Rock and Freyburg.

Each interview totaled between forty-five and sixty minutes in length. Interviews were conducted in different locations and situations including in informants' homes and on their farms, in hospitals and nursing homes, or while participating in transactions with local business merchants.

Currently, the main task of the TGDP lies in recording as many hours of interviews as possible in order to create as large a data pool of Texas German as possible. This documentation is becoming increasingly pressing because more and more speakers of Texas German are passing away, thereby taking their dialect with them. So far, the youngest informant was 68 years old, and the oldest was 92 years old.

2.3. The Texas German Dialect Archive (TGDA)

Interviews are recorded on Mini Disc and/or digital Mini Video, subsequently transferred to the TGDP's main computer, and finally converted into a variety of digital formats for further dissemination. The informants' identity is kept anonymous by assigning each interview a specific number and by deleting the

---

[6] These English sentences are the same sentences used by Susan Guion and John Kaufmann during their fieldwork in Fredericksburg in 1992 and which subsequently served as the basis for Guion (1996). I am grateful to Susan Guion for making a copy of their questionnaire available to me.

first section of the interview in which informants volunteer their personal information.

In order to preserve the recordings for further generations, each interview is burnt onto a number of CDs to be stored in different locations. Furthermore, each interview (or sections thereof) is transcribed and translated. Finally, interviews are stored in the Texas German Dialect Archive (TGDA), a password-protected on-line archive aimed to make information on Texas German available to a broader audience (see Boas (in press)).[7]

The graphical user interface designed to access the web-based TGDA consists of a digitized map of central Texas listing the different fieldwork sites where members of the TGDP are conducting fieldwork. By clicking on a specific location, e.g., Fredericksburg, users see a pop-up window containing a list of file names giving the length of each available file for that location. Each file name is linked to edited audio/video and text files that contain portions of recordings of linguistic interviews. By clicking on a file name, a Quicktime window opens and plays the file with the combined audio, video, and text data.[8] Users are able to play the entire file or only parts of it in order to conduct a linguistic analysis of the dialectal features of different speakers of Texas German. While the file is playing, the pop-up text window contains a transcript of the interview and a translation. This feature enables users to understand the recordings more easily, thereby facilitating linguistic analysis of dialectal features. By combining audio, video, and text data and delivering them over the web, users of the TGDA feel as if they are sitting directly across from the Texas German informants as they talk.[9]

As members of the TGDP continue conducting fieldwork, more interviews will be added to the TGDA. Once the TGDA contains a large enough number of recordings of various speakers from different locations across central Texas, it will be possible to start analyzing individual linguistic features of Texas German in more detail. This investigation will not only shed light on the current state of Texas German. It will also yield valuable insights regarding how the dialect has changed since the last in-depth descriptions and analyses of Texas German were conducted some four decades ago.

The following section gives an overview of a number of selected linguistic features extracted from the interviews conducted in spring 2002. Note that the choice of examples is in no way the result of a systematic sampling method—but

[7] The web-based TGDA is part of the TGDP web site which also contains information about Texas German heritage and culture, references about Texas German and other German dialects spoken in the United States, and links to other sites concerned with the preservation of endangered languages and their dialects.

[8] The majority of interviews are audio only since informants typically are uncomfortable with them being video taped.

[9] The TGDA is expected to go on-line at the end of August 2002. For a more detailed description, see Boas (in press). Eventually, the content of the recordings as well as their transcriptions and translations are to be aligned. This will allow users to listen to (and view) each interview with subtitles without having to follow the transcription and translation texts in separate windows.

only a collection of representative examples randomly taken from the first twenty hours of recordings.[10]

3. Texas German in 2002

3.1. The lexicon

As has been pointed out before, Texas German has systematically borrowed English words because German "lacked many names to describe the new environment" (Gilbert 1965: 110). Among the borrowings are words "for new plants, animals, and concepts encountered (e.g., 'armadillo', 'live oak', and 'ranch')" (Guion 1996: 449) or "names of entirely new cultural innovations or technical improvements" such as *hay-bailer*, *sonic boom*, and *jello* (Gilbert 1965: 111).

The fieldwork conducted in spring 2002 confirms these observations but also reveals a range of borrowings that is far greater than that described by Gilbert (1965) or Guion (1996). For example, a number of Texas German informants used the words *tree*, *food*, *church*, and *week* instead of their German counterparts *der Baum* 'tree', *das Essen* 'food', *die Kirche* 'church' and *die Woche* 'week'. The use of these nouns in Texas German cannot be attributed simply to borrowing due to a lack of a similar word in German. Rather, for each of the borrowed nouns, Texas German already has its own German noun to describe the objects or concepts denoted by the loanwords. This observation suggests that Gilbert's (1965: 110) claim that "the actual proportion of English words in German utterances remains small, probably less than 5 percent" needs to be seriously re-evaluated once enough data are recorded that reflect the current state of Texas German.

Other areas of the vocabulary that have been influenced by English include prepositional calques (*mitaus* 'with+out', *mitohne* 'with+out') (see Guion (1996: 449)) and discourse markers such as *well*, *sure*, and *right*, among others. One of the most interesting current developments in Texas German includes the borrowing of verbs. Whereas previous studies such as Gilbert (1965), Eikel (1967), Salmons (1983), and Guion (1996) report on the borrowing of nouns, discourse markers, and conjunctions (among others), there exists to my knowledge no previous description of verbs being borrowed into Texas German. However, the data recorded in spring 2002 include a number of sentences by different speakers in which English verbs are borrowed into Texas German as the following set of examples illustrates:

(1) a. Ich habe mich behaved.	b. Ich habe mich benommen.
I have myself behaved	*I have myself behaved*
'I behaved myself.'	'I behaved myself.'

[10] At this preliminary stage it is too early to conduct a detailed analysis of the current state of Texas German (see, e.g., Guion (1996)) because the data have not yet been systematically transcribed, categorized, and analyzed.

In (1a) the English verb *behave* is borrowed into Texas German, an effective substitution for the Standard German counterpart *sich benehmen* 'to behave' as seen in (1b). Note that the borrowed verb *behave* does itself not have a restructuring effect on the syntax of the German sentence in which it occurs, i.e., the structure of the main declarative clause in (1a) remains unaffected by the borrowing of *behave*. A preliminary analysis of the spring 2002 data reveals a number of other verbs such as *beat* and *come* that have apparently made their way into the speech of various Texas German informants. Since the borrowing of verbs into Texas German seems to be a relatively underdescribed phenomenon, it is first necessary to gather more data before making any claims about the range of verb borrowings.[11]

3.2. Plural Morphology

The distribution of plural morphemes in Texas German is another feature that has undergone significant changes over the last four decades. For example, Eikel (1967) reports that

> [t]he plurals of nouns in NBG are often formed like the plurals of native nouns in Standard German (SG): a. With a plural identical with the singular form, or with vowel modification. b. with the addition of *–e* to the singular, with or without vowel modification. c. With the addition of *–er* to the singular, with or without vowel modification. d. With the addition of *–n* or *–en* to the singular (p.83).[12]

Eikel's (1967) findings are very close to Gilbert's (1963) which also identify the frequency with which individual plural morphemes are used. In particular, Gilbert points out that [-ən] and [-s] are the most frequently occurring plural morphemes, the latter usually being used to pluralize English loan words. While the observations made by Eikel and Gilbert in the 1960s describe a rather rich plural morphology that is relatively close to Standard German, Salmons (1983: 194) points out with respect to paradigm reduction that "we may see the beginning of some breakdown in the language system." Citing the example of *Koffer* 'suitcase', Salmons observes that the plural form of this noun "is formed with –n, -s, or –ø" (1983: 194).

A preliminary analysis of current fieldwork data confirms Gilbert's 1963 description of [-ən] and [-s] as being the plural morphemes employed most frequently. Our preliminary analysis also confirms Salmons' 1983 observation regarding the instability of different plural morphemes occurring with nouns. For example, there are a large number of nouns of German origin whose plurals are

[11] The current Texas German data strongly suggest that the amount of English words borrowed into the dialect is steadily increasing, an observation already made by Salmons (1983: 193): "[Y]ounger informants used more English loans than older informants." If this trend continues – and there is no evidence to the contrary – Gilbert's (1963: 110) claim that "the actual proportion of English words in German utterances remains small, probably less than 5 percent" is outdated.

[12] NBG means New Braunfels German.

formed either by suffixation of an [-s] as in (2b) or by employing the zero morpheme as in (2c).

(2)	a. das Kind	b. die Kinder-s	c. die Kinder-ø
	the child	*the child-pl*	*the child-pl*
	'the child'	'the children'	'the children'

Although both plural morphemes are used with nouns that occurred exclusively with the zero morpheme some forty years ago, the [-s] variant occurs with a much higher frequency than the zero variant in the 2002 data. This observation suggests that the [-s] plural morpheme is becoming more productive than previously thought.

There also exist a number of nouns that have undergone a different change in how their plural variants are formed. These nouns have previously taken the [-ə] morpheme only to mark their plurality. In contrast, the 2002 data reveal an unsystematic pattern of plural formation as the following examples illustrate.

(3)	a. die Freund-en	b. die Freund-s	c. die Freund-ø	d. die Freund-e
	the friend-pl	*the friend-pl*	*the friend-pl*	*the friend-pl*
	'the friends'	'the friends'	'the friends'	'the friends'

The data in (3) show that the plural morpheme attached to the nominative *Freund* 'friend' currently has four different allomorphs in Texas German as opposed to only one (i.e., [-ə]) some four decades ago. The seemingly unsystematic variation in plural morphology exemplified by (3a–d) is not restricted to a small number of nouns but seems to be rather widespread according to a preliminary analysis of the first twenty hours of current recordings.

Unfortunately, there is at this point no large enough corpus of current Texas German data available to conduct in-depth analyses that would reveal the full range of systematic and unsystematic variation in plural morphology. However, if it turns out that similar inconsistent patterns for plural formation prevail across the lexicon then this variation might very well be what Salmons called "the beginning of some breakdown of the language system" (1983: 194).

3.3. The Case System

Another area of interest in the study of Texas German is its case system which "has been greatly reduced over its recorded history" (Guion 1996: 454). Whereas the genitive had basically dropped out of Texas German by the 1960s (except for a few fixed phrases)[13], the dative has been in recession for some time as well[14]. For example, Eikel (1967: 91) observes that "the use of the dative decreases from

[13] See Eikel (1967: 88-89) for a list of examples.
[14] See Eikel (1967: 89-91) for a list of examples.

generation to generation" and that "the accusative is used in many instances for the dative."

A preliminary analysis of the 2002 recordings confirms previous observations regarding the recession of the genitive and dative cases. Of particular interest is the development of the dative case which was used much less frequently than reported by Eikel (1967) ("about half the time expected"). This development has led to the accusative markers taking over most of the functions previously held by genitive and dative markers, thereby leading to the creation to what Gilbert (1965: 109) labels a "new case" or "simply oblique or non-nominative – as opposed to the nominative case." The following examples from the 2002 data illustrate how the accusative in (4a) and (5a) has taken over the functions of the dative in (4b) and (5b), respectively.

(4) a. Helf mich! *help me: ACC* 'Help me!'	b. Hilf mir! *help me:DAT* 'Help me!'

(5) a. wegen den Tisch *because the table:ACC* 'because of the table'	b. wegen des Tisches *because the table:GEN* 'because of the table'

Besides the accelerated recession of the dative case, the current recordings reveal another interesting tendency in the development of the Texas German case system. Whereas previous data show that Texas German has reached a stage in its development where the only case opposition is between nominative and accusative (the "new" oblique case), it appears as if even this distinction is beginning to erode. That is, in some instances informants used the nominative instead of the expected accusative as the following examples illustrate.

(6) a. nach der erste Krieg *after the first war:NOM* 'after the first war'	b. nach den ersten Krieg *after the first war:ACC* 'after the first war'

Example (6b) contains the preposition *nach* 'after' which is recorded by Eikel (1967: 93) as governing the "[a]ccusative instead of dative" case. This example by itself is not new since it is an instance of the more general Texas German tendency of dative governing prepositions to govern the accusative case. However, what makes this example interesting is the fact that some of the 2002 informants do not use the accusative case following *nach* on a regular basis any more. Instead, the nominative case is used to mark noun phrases governed by *nach* as in (6a). So far, none of the informants using the nominative following *nach* use this case exclusively, but rather in a number of instances while retaining the accusative in other cases. Our preliminary analysis of the data also shows that nominative case marking in oblique case marking positions is not an isolated phenomenon. Similar patterns of case erosion are observed with prepositions that have traditionally governed the accusative such as *durch* 'through'. Example (7b) shows how four decades ago, *durch* used to be a preposition "that regularly require[d] an object in the accusative" (Eikel 1967: 92). In contrast, (7a) exemplifies the use of *durch* followed by a nominative in the 2002 recordings.

(7)	a. durch der erste Krieg	b. durch den ersten Krieg
	through the first war: NOM	*through the first war:ACC*
	'throughout the first war'	'throughout the first war'

Examples such as (6) and (7) suggest that although the functions of the genitive and dative cases have been taken over by accusative markers, even this "new case" (Gilbert 1965: 109) may not be stable as of 2002. At this point of the analysis it is too early to come to definite conclusions about the full range of accusative case erosion. However, the data analyzed so far do suggest that even the accusative case may not be immune to erosion under the influence of other cases, in this case the only remaining case, i.e., the non-oblique nominative.

3.4. Phonetics/Phonology

The change in pronunciation of Texas German is one of the most obvious changes that has taken place over the last forty years. In particular, some of the Texas German vowels seem to have undergone a quite dramatic change since their features and distribution were last described in detail by Eikel (1966) and Gilbert (1972), among others.

For example, Eikel (1966) characterizes the vowel in the third person singular of *gehen* 'to go' as a long mid-high front vowel [e:] as in (8a). In contrast, a preliminary analysis of the 2002 recordings reveals that the majority of informants produce a diphtongized version of this vowel as in (8b) (see Boas, in press). Another change that has taken place in Texas German is in environments in which vowels precede nasals. Eikel reports for *siebzehn* 'seventeen' a regular long [e:] as in (9a). The 2002 data in (9b) show that this vowel has become both diphtongized and nasalized (see Boas (in press)).

(8)	a. [ge:t]	b. [geyt]	'goes'
(9)	a. [si:ptse:n]	b. [si:ptsẽyn]	'seventeen'

While both diphtongization and nasalization of long vowels in environments such as (8) and (9) occur quite frequently, it is not entirely clear what the conditioning environments for the nasalization in *Vieh* 'cattle' and *Tag* 'day' in the following examples are (see Boas (in press)).

(10)	a. [fi:]	b. [fĩ:]	'cattle'
(11)	a. [ta:x]	b. [tã:x]	'day'

Besides changes in the vowel system, Texas German has undergone significant changes in its consonant system, in particular in word-initial position. For example, Gilbert (1972) reports that the word-initial consonant in *Zimmer* 'room' has two variants, i.e., the voiceless affricate [ts] as in (12a) as well as the voiceless alveolar fricative [s] as in (12b). Our preliminary analysis of the 2002 data shows that both variants occur rather infrequently and are instead substituted by the voiced alveolar fricative [z] as in (12c) (see Boas (in press)). Other substitutions observed include cases in which a voiceless palato-alveolar fricative [ʃ] in word-initial position as in (13a) is substituted by a voiceless alveolar fricative [s] in positions in which it is followed by a voiced labio-dental fricative as in (13b) (see Boas, in press).

(12) a. [tsimər] c. [ziməɹ] 'room'
b. [simər]

(13) a. [ʃvimən] b. [svimən] 'to swim'

As with the other linguistic changes briefly outlined in this paper, it is at this point of the investigation too early to give a full-fledged description of the current state of Texas German, let alone a full analysis of the underlying factors that have triggered these changes. Our overview of how a number of representative linguistic features of Texas German have drastically changed over the last forty years shows that much work remains to be done. As members of the TGDP are recording more data that are subsequently stored in the TGDA, we hope to build a large enough corpus (at least 100 hours of interviews) that will eventually allow us to produce a detailed inventory of linguistic features of Texas German as it is spoken at the beginning of the new millennium.[15]

4. Research Questions

In addition to describing the current state of Texas German, the TGDP will also seek answers to the following research questions in the years to come: (a) Which linguistic features of Texas German have changed over the last four decades and why? (b) Which of these changes can be attributed to contact with English? (c) Are these changes similar to those of other rapidly eroding dialects that are in contact with other languages? (d) Is change in dialect death different from other types of language change? (e) How close is Texas German to becoming extinct? (f) What are the sociolinguistic causes leading to language shift from German to English? (g) What types of German words are retained by members of the Texas

[15] To this end, members of the TGDP will not only have to conduct more interviews with informants who live in the areas that comprise the three current fieldwork sites (Fredericksburg, New Braunfels, Freyburg, and Round Rock). It is also planned to add further fieldwork sites across Washington, Austin, Fayette, DeWitt, and Medina counties to the research area of the TGDP.

German speech community who have shifted to English? (h) In which situations do bilinguals code-switch from Texas German to English and vice versa?

References

Barbour, S. and P. Stevenson. 1990. *Variation in German.* Cambridge: CUP.

Boas, H.C. in press. The Texas German Dialect Archive as a tool for analyzing sound change. *Proceedings of the Workshop on Methods and Tools in Field Linguistics. Held in conjunction with the Third International Conference on Language Resources and Evaluation.* Las Palmas, Spain.

Biesele, R.L. 1928. A History of the German Settlements in Texas. Ph.D. dissertation, UT Austin.

Crystal, D. 2000. *Language Death.* Cambridge: CUP.

Dorian, N.C. 1982. Linguistic models and language death evidence. In L.K. Obler and L. Menn (eds.) *Exceptional Languages and Linguistics*, 31-48. New York: Academic Press.

Eikel, F. 1949. The use of cases in New Braunfels German. *American Speech*, **24**: 278-281.

Eikel, F. 1954. The New Braunfels German Dialect. Baltimore:

Eikel, F. 1966. New Braunfels German: Part II. *American Speech* **41**: 254-160.

Eikel, F. 1967. New Braunfels German: Part III. *American Speech* **42**: 83-104.

Gilbert, G. 1963. The German Dialect spoken in Kendall and Gillespie Counties, Texas. Ph.D. dissertation, Harvard University.

Gilbert, G. 1965. English loanwords in the German of Fredericksburg, Texas. *American Speech* **40**: 102-112.

Gilbert, G. 1972. *Linguistic Atlas of Texas German.* Austin: UT Press.

Guion, S. 1996. The death of Texas German in Gillespie County. In P.S. Ureland and I. Clarkson (eds.) *Language contact across the North Atlantic*, 443-463. Tübingen: Niemeyer.

Salmons, J.C. 1983. Issues in Texas German language maintenance and shift. *Monatshefte* **75**:187-196.

Hans C. Boas
Department of Germanic Studies
E.P.S. 3.102, C3300
University of Texas at Austin
Austin, TX 7812-1190, U.S.A.

hcb@mail.utexas.edu

Leggbo Verb Inflection: A Semantic and Phonological Particle Analysis

LARRY M. HYMAN[1], HEIKO NARROG[2], MARY PASTER[1], & IMELDA UDOH[3]
University of California, Berkeley[1], Hokkaido University[2] & University of Uyo[3]

0. Introduction

One of the exciting aspects of working on a "field language" is that virtually everything is open to investigation. Not only can and must researchers address issues that arise simultaneously in the phonology, morphology, syntax, and semantics etc., but also the complex interrelations that sometimes exist between the different "modules" of a language. What is particularly satisfying is when phonological, grammatical, and semantic properties converge to allow one coherent, overarching statement. The present paper presents one such result concerning verb inflection in Leggbo, an Upper Cross minority language spoken by an estimated 60,000 people in Eastern Nigeria (Grimes 2000).[1] The goals of this paper are to present, first, a featural analysis of the inflectional system of oppositions which are explicitly marked on Leggbo verbs, specifically, aspect, mood, polarity and clause type; and second, an account of how this system of oppositions is realized in morphological terms. It will be seen that the inflectionally marked categories are organized in terms of a fixed hierarchy of privative features or "particles" which compete for expression within the Leggbo verb paradigm.

The paper is organized as follows. First, we present the verb inflection system of Leggbo. Then we discuss how the proposed morphological features occur in combination, focusing especially on cases where their individual spell-outs conflict. We conclude with a brief discussion of the implications of our findings.

1. Verb inflection system of Leggbo: the components

A most striking first fact about Leggbo is that it does not distinguish tense. Thus, the sentence in (1a) is underspecified as to present vs. past time reference:

[1]Leggbo, which is spoken by the Agbo people in two local government areas, recently named Abi and Yakurr, in Cross-River State, is the language being investigated in our field methods course (Linguistics 240ab) this year—which we hope will result in a published grammar of the language. We are grateful for the contributions of the other full- and part-time members of the class (Jeff Good, Ahmadu Kawu, Julie Larson, Ian Maddieson, Keith Sanders, and Tess Wood). We would like especially to thank the Wenner-Gren Foundation for Anthropological Research (for travel support to Imelda Udoh in March 2001), the Committee on Research, George Breslauer, Dean of Social Sciences at UC Berkeley, and the African Studies Center for providing support for Dr. Udoh to be at Berkeley for the year. Finally, Drs. Narrog and Udoh wish to thank Hokkaido University and the University of Uyo, respectively, for granting their sabbatical leaves to be in Berkeley.

(1) a. ba zee icéji — 'they see/saw Icheji'
3pl see Icheji

b. ba zee icéji lEgbàl amma — 'they see Icheji now'
3pl see Icheji time this

c. ba zee icéji lEgbàl ámmE — 'they saw Icheji then'
3pl see Icheji time that

As seen in (1b) and (1c), the time reference can be made explicit by the addition of the appropriate present or past time adverbial.

While tense is not marked on the verb, the inflectional features which do receive an overt expression belong to the categories Aspect, Mood, Polarity and Clause Type. For the purpose of this talk, we assume the simplified verb structure in (2).

(2)
verb
prefixes — stem
stem: root — (suffix)

The verbal unit consists of a stem preceded by one or more prefixes, including a subject agreement marker[2]. The stem in turn consists of a root and possible suffix. Recognizing the constituencies in (2), we now can indicate, as a first approximation, how aspect, mood, polarity and clause type are inflected on the Leggbo verb (3):

(3) Inflectional features ("particles") marked on verbs [first approximation]

				Segmental Marking		Tonal Marking	
				Prefix	Suffix	Prefix	Stem
	Aspect	Progressive	**P**		-i	m	H-m
		Habitual	**H**	nà-		m	L-l
	Mood	Irrealis	**I**			h	L-l
	Polarity	Negation	**N**	aà-		l	H-m
	Clause	Consecutive	**C**			l	L-l

The letters P-H-I-N-C stand for privative features (particles) which are either present or absent in the underlying representation of inflected verbs. Particles are spelled out segmentally or tonally. A tonal exponent may be assigned to a prefix (e.g. the subject marker) or to the stem. In the last two columns of (3), the alternating H/L root tone is in upper case, while prefix and suffix tones are indicated in lower case.[3]

The elements P, H, I, N, and C may occur alone on a verb or may occur in combination with each other. In the latter case, when the morphological spell-outs of two or more of these elements conflict, the surface form of the verb will be determined by the feature whose spell-out is ranked highest. In addition, as we have already seen in (1), a verb may completely lack any of these features, in which case a default perfective is obtained, which may have either present or past meaning.

[2]As in many Niger-Congo languages, the one verb form not requiring a prefix is the singular affirmative imperative, which also serves as the base entry of verbs in our lexicon, e.g. *zee* 'see!'.
[3]In cited forms, an acute accent (´) marks H(igh) tone, a grave accent (`) marks L(ow) tone, and a vowel lacking an accent carries M(id) tone. For more on Leggbo tone, see Paster (2002).

Additional examples are given in (4).

(4) Ø-marked verb (perfective): non-habitual, non-progressive, realis, affirmative

a. activity verbs = past

ba	mɔ́ɔ́ŋɔ	'they returned'		ba	dzi	lídzil	'they ate food'
3pl	return			3pl	eat	food	

b. stative verbs = present/past

ba	nná	'they shine/shone'		ba	dzɛlɛ	icéji	'they know/ knew Icheji'
3pl	shine			3pl	know	Icheji	

Activity verbs (4a) *mɔ́ɔ́ŋɔ* 'return' and *dzi* 'eat' have past meaning, while stative verbs (4b) *nnà* 'shine' and *dzɛlɛ* 'know' can have present or past interpretations. Thus, the Ø-form has a perfective meaning, presenting events as a whole. In our account it is the unmarked form characterized by absence of aspect-mood-polarity features: it is non-progressive, non-habitual, realis, affirmative, and main clause.

Activities which are not completed, that is, which are ongoing at the time of the speech event, have to be marked as "progressive." As seen in (5a), which involves the activity verb *dzi* 'eat', the progressive form is completely neutral between present and past interpretations, which can be disambiguated by adding the appropriate present or past time adverbial, as in (5b,c).

(5) Progressive (P) marked by *-i* suffix (on the activity verb *dzi* 'eat')[4]

a.	ba	dzi-i	lídzil			'they are/were eating food'
	3pl	eat-P	food			
b.	ba	dzi-i	lídzil	lɛgbàl	amma	'they are eating food now'
	3pl	eat-P	food	time	this	
c.	ba	dzi-i	lídzil	lɛgbàl	ámmɛ	'they were eating food then'
	3pl	eat-P	food	time	that	

The examples in (6) show that the progressive of stative verbs such as *kkù* 'stay' is also neutral between present and past meaning:

(6) Progressive (P) marked by *-i* suffix (on the stative verb *kkù* 'stay')

a.	ba	kkú-i	ḿmɛ			'they are/were staying here'
	3pl	stay-P	here			
b.	ba	kkú-i	ḿmɛ	lɛgbàl	amma	'they are staying here now'
	3pl	stay-P	food	time	this	
c.	ba	kkú-i	ḿmɛ	lɛgbàl	ámmɛ	'they were staying here then'
	3pl	stay-P	food	time	that	

Another aspectual feature, habitual, marked by a *na* prefix, denotes an event

[4]While all progressive verb forms involve an -i suffix, there are two potential complications. First, most verbs also undergo fortition of one or both consonants, e.g. mɔ́ɔ́ŋɔ 'return' –> mɔ́ŋŋ-i 'be returning'. Other verbs use /-azi/, which does not condition fortition, e.g. kum 'pierce' –> kum-azi 'be piercing'. In other cases, -i (+fortition) or -azi may indicate pluractionality.

occurring regularly or an activity being performed habitually. As indicated in (7a), the habitual shares with the progressive the property of not carrying any implications in regard to temporal reference. As before, (7b,c) show that temporal reference can be disambiguated by means of time adverbials:

(7) Habitual (H) marked by *nà* prefix

a.	ba	nà-dzi	lídzil	ŋ́-kɛ etekpan			'they eat/used to eat food outside'
	3pl	H-eat	food	outside			
b.	ba	nà-dzi	lídzil	ŋ́-kɛ etekpan	lɛgbàl	amma	'they now eat food outside'
	3pl	H-eat	food	outside	time	this	
c.	ba	nà-dzi	lídzil	ŋ́-kɛ etekpan	lɛgbàl	ámmɛ	'they then used to eat food outside'
	3pl	H-eat	food	outside	time	that	

The same aspectual distinctions are found in the negative. As seen in (8), negation not only involves the prefix *aà*, but also the use of a 3pl subject pronoun *bɛ̀*, different from its affirmative counterpart, *ba*:[5]

(8) Negative (N) marked by *aà* prefix

a.	bɛ̀	aà-nná		'they do/did not shine'
	3pl	N-shine		
b.	bɛ̀	aà-kkú-i	ḿmɛ	'they aren't/weren't staying here'
	3pl	N-stay-P	here	

In addition, the examples in (9) show that the object precedes a negative verb:

(9) SOV word order in the negative

a.	bɛ̀	icéji	aà-dzælɛ			'they don't/didn't know Icheji'
	3pl	Icheji	N-know			
b.	bɛ̀	lídzil	aà-dzi-i			'they aren't/weren't eating food'
	3pl	food	N-eat-P			
c.	bɛ̀	lídzil	dzɛ̀	aà-dzi	ŋ́-kɛ etekpan	'they don't eat/didn't use to eat outside'
	3pl	food	H	N-eat	outside	

The one complication in (9c) concerns the additional verbal auxiliary *dzɛ̀*, which derives from the homophonous verb dzɛ̀, meaning 'to finish'.

Finally, continuing our demonstration that Leggbo lacks tense, note that what we call the irrealis mood in (10), marked by the high tone on the subject agreement marker *bá* (vs. mid tone *ba* in the realis), can have either a future or conditional meaning, as in (10a), or a subjunctive function, as in (10b).

(10) Irrealis (I) marked by H tone on subject agreement marker *bá*

a.	bá	dzi	lídzil	'they will/would eat food'	(cf. ba dzi 'they ate')
	3pl-I	eat	food		

[5]The affirmative subject markers are *m* '1sg', *a* '2sg', ɛ '3sg', *mɛ* '1 pl.' and *ba* '2pl/3pl'. The four subject prefixes are identical in the negative, where, however, bɔ̀ '2pl' and bɛ̀ '3pl' are distinguished, both distinct from *ba*.

b. m-vɔ́ŋi bɛ ta <u>bá</u> dzi lídzil 'I want them to eat food'
1sg-want 3pl comp 3pl-I eat food (I want them that they eat food)

On the other hand, a L tone subject prefix is used on consecutive or serialized verbs:

(11) Consecutive (C) marked by L tone on subject agreement marker *bà*

ba númi ɛ <u>bà</u> nìì bɛ̀ɛ́ 'we gave it to children' (lit. we
they take it they-C give children took it & gave to children')

As we shall discuss below, tone is implicated in the realization of all of the marked inflectional features except the progressive aspect. Before moving on to tone, however, let us summarize and elaborate on the segmental properties seen thus far.

2. Segmental marking

As shown in (12a), most lexical verb stem entries consist of one or two syllables:

(12) Verb-stem shapes in Leggbo

a. most verb stem entries

CV : dzi 'eat', nnà 'shine', zu 'exist', kkù 'stay'
CVV : zee 'see', zai 'wash', vyài 'wash and squeeze', dua 'hide'
CVCV : fina 'touch', zumi 'extinguish', dzɛlɛ 'know'
CVVCV : mɔ̀ɔ̀ŋɔ 'return', vèèli 'lend', taali 'draw a line'

b. mostly derived verb stems (pluractional or reduplicated CV)

CVCVCV : fin-azi 'touch-pl', zum-azi 'extinguish-pl', fi-fina 'really touch', zu-zumi 'really extinguish'
CVVCVCV : mɔ̀ɔ̀ŋɔzi 'return-pl'
CVCVVCV : mɔ̀-mɔ̀ɔ̀ŋɔ 'really return'

c. with reduplicated CV + -azi

CVCVCVCV : fi-fin-azi 'really touch-pl', zu-zum-azi 'really extinguish', mɔ̀-mɔ̀ɔ̀ŋ-ɔzi 'really return-pl'

As exemplified in (12b), most longer verbs either contain the progresive/pluractional suffix *-azi* or have reduplication of their first syllable (with "intensive" meaning). The maximum number of syllables possible in a verb stem is, thus, four, as in (12c).

Considering just CVCV verbs, V-V combinations are distributed as in (13):

(13) V1-V2 distribution in 147 CVCV verbs (out of 356 verbs in current lexicon)

V1 / V2	**i**	**e**	**ɛ**	**u**	**o**	**ɔ**	**a**	**Totals:**
i	3	0	0	0	0	0	12	15
e	7	9	0	0	0	0	0	16
ɛ	2	0	13	0	0	0	0	15
u	7	0	0	0	0	0	13	20
o	1	0	0	0	6	0	0	7
ɔ	5	0	0	0	0	27	0	32
a	15	0	0	0	0	0	27	42
Totals:	40	9	13	0	6	27	52	147

As proposed in (13), these verbs consist of a CVC (or CVVC) root plus a suffix, either /-i/ in (14a), which can occur after all V1 vowels, or /-a/ which assimilates to a preceding mid vowel, as in (14b).[6]

(14) Analysis of CVCV verbs as /CVC-i/ and /CVC-a/ (including CVVCV)

a.	/CVC-i/	:	V1 can be any vowel					
b.	/CVC-a/	:	/CiC-a/, /CuC-a/, /CaC-a/ (no change)					
			/CeC-a/	∅	CeC-e	/CoC-a/	∅	CoC-o
			/CɛC-a/	∅	CɛC-ɛ	/CɔC-a/	∅	CɔC-ɔ

The underlying V2 /-i/ and /-a/ are thus frozen lexical suffixes found only on some verbs. As schematized in (15a), the progressive /-i/ overrides the lexical suffix /-a/:

(15) Progressive /-i/ overrides lexical /-a/

a.	fin-a + P	∅	finn-i	'be touching'
	mɔ̀ɔ̀ŋ-ɔ + P	∅	mɔ̀ŋŋ-i	'be returning'
b.	bin-i + P	∅	bin-azi	'be carrying'
	vìl-i + P	∅	vìl-azi	'be cutting'

As also seen, consonant fortition, here written as double, frequently accompanies the progressive suffix *-i*. (In addition, a preceding long vowel will be shortened before a fortis consonant.) Approximately 1/4 of Leggbo verbs, including those indicated in (15b), instead use the underlying suffix /-azi/ without fortition.[7] In many cases the same verb may undergo (15a) in the progressive, but (15b) to express pluractionality, or (15a) or (15b) may be ambiguous between progressive and pluractional meaning.

By contrast with the progressive, which modifies the stem in the two ways indicated in (16), the segmental marking of the habitual is as a prefix, which fuses with the subject marker as indicated in the table in (17).

(16) Habitual marking of bila 'climb'

Person	Habitual form	cf. Ø-form
1sg	nim̀-bila	m bila
2sg	naà-bila	a bila
3sg	neè-bila	e bila
1pl	ma nè-bila	me bila
2pl	ba nà-bila	ba bila
3pl	ba nà-bila	ba bila

The negative, in its simplest form, is also marked by adding a prefix, *aà*, which fuses with a slightly different set of subject markers, as in the second column in (16).

[6]Given difficulties of analysis, we have omitted verbs from the table whose intervocalic consonant is a weakly articulated velar approximant "gh" or "ghost-h".

[7]The vast majority of these either end in *-i* or already have a fortis consonant, suggesting that *-azi* is used to make the progressive more different from the corresponding non-progressive forms.

(17) Negative Ø (perfective) of bila 'climb'

Person	Negative form	Negative Habitual
1sg	mm̀-bila	... (dzɛ̀)... mm̀-bila
2sg	aà-bila	... (dzɛ̀)...aà-bila
3sg	eè-bila	... (dzɛ̀)...eè-bila
1pl	mɛ̀ eè-bila	...(dzɛ̀)... mɛ̀ eè-bila
2pl	bɔ̀ aà-bila	...(dzɛ̀)... bɔ̀ aà-bila
3pl	bɛ̀ aà-bila	...(dzɛ̀)...bɛ̀ aà-bila

The third column shows the extra marker *dzɛ̀* in the negative habitual.

3. Tonal marking

With the segmental marking now established, we now turn to the question of tone. As seen in (18a), Leggbo has three surface tones: H(igh), M(id), and L(ow), which contrast on noun roots:

(18) a. Noun roots, which usually take a prefix, exhibit a three-way contrast

L-L	:	lè-kɔ̀l	'neck'	M-L	:	le-dùl	'bundle'
L-M	:	lè-tol	'head'	M-M	:	li-bul	'bow, arrow'
L-H	:	lì-tɔ́l	'ear'	M-H	:	li-kól	'mat'

b. Verb roots show only a two-way opposition

"M-tone verbs"		**"L-tone verbs"**	
dzi	'eat'	sì	'make, do'
tʌm	'send'	nùm	'take'
mana	'catch, hold'	fìna	'touch'
beeli	'escort'	mɔ̀ɔ̀ŋɔ	'return'

As seen in (18b), there is only a two-way distinction among verbs with respect to tone. We term the two classes "M-toned verbs" and "L-toned verbs", with the tone label referring to the underlying tone of the root. Important for our study, the inflectional particles under examination here may contribute a tone that associates to the prefix, root, suffix, or a combination of these. Throughout the verbal paradigm, M-toned verbs surface with M on the root almost without exception, as illustrated in (19a). L-tone verb roots, however, alternate between L and H, as shown in (19b).

(19) a. M root tone is stable

m̲ana	'catch!'
ba m̲ana	'they caught'

b. L root tone alternates with H

f̲ìna	'touch!'
ba f̲ína	'they touched'

Therefore, since their root tone alternates, L-toned verbs better exemplify the full effect of the grammatical tone assignments operating in different parts of the inflectional paradigm.

The table in (20) shows the prefix and stem tones of each inflectional category.

(20) Tonal marking of inflectional features on subject prefix and verb stem

	Prefix tone	Stem tone
Ø/P	m	H-m
H	m-l	L-l
I	h	L-l
N	(l) - ml	H-m
C	l	L-l

We begin with the prefix tone, illustrated (21).

(21) Examples of tonal realizations

	prefix		**M root**	**/**	**L root**
Ø	ba		mana	/	fĩna
P	ba		mann-i	/	fĩnn-i
H	ba nà-		manà	/	fĩnà
I	bá		manà	/	fĩnà
N	(bɛ̃) aà-		mana	/	fĩna
C	bà		manà	/	fĩnà

We take the M tone of the subject prefix *ba* to be unmarked in the Ø, progressive, and habitual forms. We also consider the first part of the negative marker *aà* to carry this M tone. As seen in (22a), we analyze the irrealis as assigning a H morphological tone—or particle—to the subject prefix:

(22) H and L subject prefixes

a. Irrealis H subject prefix : $[\text{ H }]_{I}$ + ba ∅bá

b. Consecutive L subject prefix: $[\text{ L }]_{C}$ + ba ∅bà

Similarly, in (22b), the consecutive assigns L morphological tone to the subject prefix. When a clause is both irrealis and consecutive, the effect is cumulative (23):

(23) Consecutive (C) marked by L(ow) tone on subject agreement marker *bà*

a. consecutive realis (L assigned to subject)

Ø = [M]			L			
ba	númi	ɛ	<u>bà</u>	niì	bɛ̀ɛ́	'we gave it to children'
they	take	it	they-C	give	children	

b. consecutive irrealis (L and H assigned to subject)

H			H+L = [M]			
bá	nùm	ɛ	<u>ba</u>	niì	bɛ̀ɛ́	'we will give it to children'
they-I	take	it	they-I-C	give	children	

In the realis in (23a), repeated from (11), no tone is assigned to the first *ba*, which is therefore realized as default M, but a L particle is assigned to the second *bà*. In (24b), the first *bá* receives the irrealis H tone, while the second *ba* receives both the irrealis H and the consecutive L features. The result is a fusion, whereby H+L is realized M, as indicated. On the basis of the subject prefix tone, we arrive at the partial hierarchy in (24):[8]

(24) Hierarchy for subject tone assignment

C, I >> H, P, Ø

What this means is that the consecutive L and the irrealis H will override Habitual, Progressive, and Ø, all three of which have unmarked subject tone.

We now consider stem tones, which are summarized in (25).

(25) Stem tones by aspect-mood-polarity-clause

	MCA	SRA	CCA	NEG
Ø/P	H-m	L-m	L-l	H-m
H	L-l	L-l	L-l	H-m
I	L-l	L-l	L-l	L-l
IH	L-l	L-l	L-l	L-l
Imperative		L-m	L-l	

MCA = main clause affirmative (non-subject relative clause = same as MCA)
SRA = subject relative clause affirmative
CCA = consecutive clause affirmative
NEG = all negatives (MCA, SRA, CCA etc.)

Recall that M verbs do not change their root tone, while the tone of L verb roots alternates with H. This is what is shown in upper case in (25)—followed by one of two suffix tones, low or mid (in lower case). We note the following generalizations:

(26) a. L-l in both the Irrealis and Consecutive, across the board, as shaded in
b. H-M in the Negative in the absence of I
c. L-l in the Habitual, in the absence of Negative
d. H-m in MCA in the absence of H or I
e. L-m in the SRA and Imperative in the absence of H or I

Restated in terms of a hierarchy of tonal assignments, we get the following in (27).

(27) C, I>> N >> H >> MCA, SRA/Imper (= Ø, P)

The above hierarchy is crucially established by the double outlined boxes in (25). As shown in (28), the Irrealis L-l pattern overrides the Negative H-m pattern, but the Negative H-m pattern overrides the Habitual L-l pattern.

(28) Irrealis >> Negative>> Habitual
L-l H-m L-l

As indicated in (29), the Habitual L-l pattern, in turn, overrides the H-m of the main

[8] The negative subject prefix has peculiarities of its own and will not concern us at this point.

clause affirmative and the L-m of the subject relative affirmative and imperative:

(29) Habitual >> MCA, SRA/Imper
L-l H-m L-m

When we combine (28) and (29) and bring in the consecutive, we get the hierarchy in (27). Since both the irrealis and the consecutive have L-l on the stem, it is impossible to tell from the surface forms which wins out between these two, so we assume that, as with the prefix tones, irrealis and consecutive are unranked with respect to each other in the stem tone override system. The same unranked relation holds of the bottom two inflectional features: main clause affirmative and the combined subject relative affirmative/imperative. Because of the unnatural class that this latter constitutes, our intuition is that the imperative is the default. While this would take us too far afield, our suspicion is that the H variant of L verbs is from a prefixal H tone that has been assigned to it, as indicated in (30):

(30) Possible origin of H root variant of L from proto *L and *H

a. default realizations: *L > L , *H > M
b. * H - [L > HL > H
c. * H - [H > H > M

Our hypothesis in (30a) is that proto *L tone is realized L, while proto *H is realized M. However, as shown in (30b), when a floating H prefix, circled, preceded a L root tone, a HL falling tone was first produced, which was simplified to a H tone. On the other hand, when the root was a H tone, as in (30c), the floating H and the root H simply fused as one H, realized M. If correct, we can interpret the H/L alternation of L tone verb roots as due to the presence vs. absence of a preceding floating H tone.

4. Further complications

To summarize thus far, we have seen how the output realization of the inflectional P, H, I, N, and C particles are determined by a ranking that reflects the scope relations that hold between these elements. Before drawing our conclusion, we need to point out that we have presented the major, but not all of the possible aspectual and clause-type distinctions that can be made in Leggbo.

Let us just consider one further form, which has particular interest. We have said all along that Leggbo does not mark tense. There is one exception to this, which is the presence of an anterior past form, illustrated in (31), which is distinguished from the Ø perfective only in a main clause affirmative:

(31) Main clause affirmative Anterior "tense" distinguished by tone

a. ba fìna 'they have/had touched' cf. ba fína 'they touched'
ba fìn-azi 'they have/had touched-pl' (L-m)
b. ba mànásu 'they have/had caught' cf. ba mana 'they caught'
ba màn-ázi 'they have/had caught-pl' (L-hm)

In both negatives and non-main clauses, the perfect has the same realization as the perfective. As seen in (32a), if the proximate time reference is the time of speaking, the meaning will be present perfect. However, if the time reference is already in the past, the result will be an anterior past, as in (32b).[9]

(32) Anterior past (to present or past time reference)

a.	ba <u>bbǒ</u> \`da they die-A already	'they have already died'
b.	ba <u>bbǒ</u> \`da bèle m̀-bɔ̀lɔ̀ m̀-wèl they die-A already before I-just I-arrive	'they had already died before I arrived'

While this may represent a small corner where tense is expressed in Leggbo, it is significant that it is relative anteriority being marked, not exact time reference. Just as in the case P, H, I, N, C, the Anterior form does not explicitly indicate when the action took place with respect to the time of speaking. In any case, this restricted form is low in the hierarchy, since it only occurs in the main clause affirmative.

A second, and last, complication we will consider concerns the perfective itself. Thus far we have implied that it is unmarked. In fact, as shown in (33a), many verbs, including all CVC roots, take an *-i* suffix in the perfective:

(33) a.	CVC	:	nùm 'take'	→	ba núm-i	'they took'	
			tɔ̀l 'pull'	→	ba tɔ́l-i	'they pulled'	
b.	CVCV	:	bila	→	ba bila	'they climbed'	(cf. P ba bidd-i)
			mɔ̀ɔ̀ŋɔ	→	ba mɔ́ɔ́ŋɔ	'they returned'	(cf. P ba mɔ́ŋŋ -i)
c.	CV	:	dza	→	ba dza-i	'they are/were good'	
			nnà	→	ba nná	'they shine/shone'	

When the verb already has a second vowel, as in (33b), the perfective does not have an *-i* suffix. Finally, when the verb has the shape CV, as in (33c), some take an *-i* suffix, while others don't. The data in (33b) suggest that in such verbs, the lexical second vowel, an *-a* suffix, overrides the spell-out of the Ø feature we have called perfective. This contrasts with the progressive *-i* which, as shown in parentheses, replaces the second vowel of CVCV verbs. We note that the gerund suffix -ɛ̀ has the same distributional property as perfective *-i*, but we leave this to further study.

4. Conclusion

In the preceding sections we have presented an analysis of the underlying features of the aspect-mood-polarity system of Leggbo as well as their morphological and phonological realizations. We have seen that the posited inflectional particles are ranked in one of two hiearchies in (34):

(34) a. Tonal properties

C, I >> N >> H >> P, Ø

[9]The tone pattern is also unique: *bbó* 'die' is a M verb. When we add the /-azi/ pluractional suffix, we obtain *bbò-ózi* 'die-pl', i.e. L-h-m. Compare this with the L verb, *fin-azi* 'touch-pl', i.e. L-m-m. We know the initial L is from a floating L- prefix, but we cannot at present explain the H.

b. Segmental properties
P >> Ø

The hierarchy in (34a) determines morphological tone assignment, while the hierarchy in (34b) is responsible for segmental overwriting in the progressive. The significance of these findings is as follows: Much of the work concerning complex inflectional morphology addresses issues of concatenation. Whether citing Bybee's (1985) semantic notion of relevance or Baker's (1985) syntactic mirror principle, it is easy to cite examples where linear ordering of concatenated affixes reproduces inherent scope relations. Basically, outer affixes have scope over inner affixes.

On the other hand, but still dealing with concatenative morphology, Anderson (1986) has been concerned with cases where more than one affix vies for the same "slot." In this case it has to be determined which affix wins out. Anderson's proposal is to establish "disjunctive rule blocks", but the same kind of hierarchy that we have proposed in (34a) will essentially do the same trick.

What Leggbo and many other African tone languages show is that in addition to segmental affix ordering and segmental affix disjunction, non-concatenative spell-outs, especially via prosodic features such as tone, also show the same hierarchical, scope effects. While many authors show verb tone patterns in tabular displays, only some have attempted to order or rank morphological spell-outs in a systematic manner (see, for example, Hyman & Byarushengo 1984 for Haya; Hyman & Olawsky, in press, for Dagbani). This is a rich area for future comparative work on the use of tone in morphology and for typological research on the semantics and morphology of inflectional morphology in general.

References

Anderson, Stephen R. 1986. Disjunctive ordering in inflectional morphology. *Natural Language and Linguistic Theory* 4.1-31.
Baker, Mark. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16.273-416.
Bybee, Joan. 1985. *Morphology: a study of the relation between form and meaning*. Amsterdam: Benjamins.
Grimes, Barbara F., ed. 2000. *Ethnologue: Volume 1, Languages of the World,* 14th Edition. Summer Institute of Linguistics.
Hyman, Larry M. & Ernest Rugwa Byarushengo. 1984. A model of Haya tonology. In J. Goldsmith & G.N. Clements (eds), *Autosegmental studies in Bantu tone*, 53-103. Dordrecht: Foris.
Hyman, Larry M. & Knut Olawsky. In press. Dagbani verb tonology. In Victor Manfredi (ed.), *Proceedings of 31st Annual Conference on African Linguistics.*
Paster, Mary. To appear. Tone Specification in Leggbo. In John M. Mugane, ed. *Linguistic Description: Typology and Representation of African Languages. Trends in African Linguistics 8.* Trenton, New Jersey: Africa World Press.

Department of Linguistics
1203 Dwinelle Hall
Berkeley CA 94720-2650

hyman@socrates.berkeley.edu

Phonetics in the Field

IAN MADDIESON
University of California, Berkeley

1. Phonetics in grammars

For many linguists the noticeable growth of interest in descriptive fieldwork over the past two decades or so has been a welcome development within our profession. In considerable part this phenomenon is linked to concern about language loss, reflected in, for example, the creation of the Endangered Languages Committee of the Linguistic Society of America, and research programs funded by the Volkswagen Foundation in Germany and the Hans Rausing Fund in Britain. Descriptive work has been undertaken both in association with efforts to reinvigorate the use of a language or out of a wish to document a disappearing resource for its cultural and scientific value. An increased interest in typological linguistics has further stimulated new descriptive work on lesser-known languages since typological studies require the broadest knowledge possible of how human languages differ one from another. The publishing industry has responded with new or relaunched grammar series, notably from Mouton de Gruyter, Routledge, and LINCOM EUROPA and there seems to be a renewed acceptance of the practice of writing a descriptive grammar as a doctoral dissertation in linguistics. Practical issues in field linguistics (e.g. Newman & Ratliff 2001, Vaux & Cooper 1999) and language maintenance (Hinton and Hale 2001) have also received more attention in the recent professional literature.

Although all this attention to descriptive linguistics is welcome it seems generally the case that little detail on specifically phonetic matters is provided in a typical grammar, nor is there much use of phonetic techniques to provide insights on other matters, such as adding precision to observations of phonological alternations or testing whether supposed syntactic ambiguities are actually disambiguated at the phonetic level. While syntactic patterns are documented with example sentences, often from natural discourse or texts, the phonetic facts are rarely if ever documented by the presentation of hard evidence.

In order to see if this impression was justified a survey of twenty grammars published or submitted as doctoral dissertations in the period of a dozen years from 1989 to 2000 was conducted. The grammars were selected to represent a wide range of language families broadly distributed around the world, as well as to sample some of the range of publishing sources and centers of research effort. The languages represented are all 'minor' languages spoken by from a handful to some tens of thousands of persons. The grammars examined are listed in (1), which is arranged alphabetically by the names of the languages and includes also the language family, author, publication or submission date, and publisher or university. Full references are given at the end of the article.

(1) The sample of grammars examined.

Language	Family	Author and date	Publisher
Aleut	Eskimo-Aleut	Bergsland 1997	Alaska Native Language Center
Chalcatongo Mixtec	Otomanguean	Macaulay 1996	U. of California Press
Cubeo	Tucanoan	Morse & Maxwell 1999	SIL
Evenki	Tungusic	Bulatova & Grenoble 1999	LINCOM EUROPA
Kisi	Niger-Congo	Childs 1995	Mouton de Gruyter
Koiari	Trans-New Guinea	Dutton 1996	LINCOM EUROPA
Kolyma Yukaghir	Ural-Altaic	Maslova 1998	U. Bielefeld (Ph. D.)
Lavukaleve	E. Papuan	Terrill 1999	Australian National University (Ph. D.)
Lekeitio Basque	Basque	Hualde, Elordieta & Elordieta 1994	Universidad del Pais Vasco
Lezgian	N. E. Caucasian	Haspelmath 1993	Mouton de Gruyter
Lillooet	Salishan	van Eijk 1997	UBC Press
Meithei	Sino-Tibetan	Chelliah 1997	Mouton de Gruyter
Miya	Afro-Asiatic	Schuh 1998	U. of California Press
Nandi	Nilo-Saharan	Creider & Creider 1989	Buske
Nivkh	"Paleo-Siberian"	Gruzdeva 1998	LINCOM EUROPA
Oneida	Iroquoian	Abbott 2000	LINCOM EUROPA
Rapanui	Austronesian	Du Feu 1996	Routledge
Semelai	Austro-Asiatic	Kruspe 1999	U. of Melbourne (Ph. D.)
Sonora Yaqui	Uto-Aztecan	Dedrick & Casad 1999	U. of Arizona Press
Wardaman	Australian	Merlan 1994	Mouton de Gruyter

(2) presents an analysis of the proportion of each of these grammars directly devoted to discussion of the sound system of the language being described. The metric chosen is crude — the proportion of numbered pages devoted to phonetic and phonological topics as a percentage of the overall numbered page count — but this gives a reasonable idea of how much attention is given to these aspects of the language. This table is arranged in increasing order of these calculated percentages.

(2). Proportion of grammar devoted to phonological/phonetic topics.

Language	"Ph" content	Page count	"Ph" percent
Cubeo	5 pages on phonology	197	2.5%
Yaqui	15 pages on phonology	411	3.6%
Evenki	3 pages on phonology	64	4.7%
Lavukaleve	24 pages on phonology and morphophonology	486	4.9%
Rapanui	13 pages on phonology and morphophonology	217	6.0%
Oneida	4 pages on phonology	65	6.2%
Semelai	47 pages on phonology	729	6.4%
Yukaghir	41 pages on phonology	375	6.6%
Wardaman	46 pages on phonology	617	7.5%
Lezgian	43 pages on phonology	567	7.6%
Aleut	31 pages on phonology	360	8.6%
Nandi	15 pages on phonology	172	8.7%
Mixtec	27 pages on phonology	298	9.1%
Lillooet	35 pages on phonology	279	9.3%
Meithei	54 pages on phonology and tonal phonetics	539	10.0%
Nivkh	9 pages on phonology	66	13.6%
Koiari	12 pages on phonology	77	15.6%
Miya	65 pages on phonology	414	15.7%
Kisi	77 pages on phonology, incl 2 pages tonal phonetics	370	20.8%
Basque	73 pages on phonology, incl 12 pages of F0 data	314	23.2%

The grammars vary in page count from 60-70 pages for some of the LINCOM EUROPA *Languages of the World/Materials* series to ten times this for some of those which were submitted as doctoral dissertations. This variation reflects both variation in richness of description as well as technical factors such as choice of type size and line spacing and decisions on pagination. The mean page count is 331. The mean number of pages explicitly devoted to phonetic/phonological topics is around 32 pages, hence the mean percentage of pages devoted to these topics is just under 10%. However only five of the twenty grammars devote more than one tenth of their pages to phonetic, phonological, or morphophonological issues and the harmonic mean is just 7.3%. Not surprisingly there is a tendency for those authors whose professional interests are more oriented towards phonology (e.g. Hualde, Schuh, Childs) to devote a greater proportion to this area. Schuh's and Macauley's grammars provide substantial additional information on certain phonological aspects in sections describing morphology (these pages are not counted in (2)). In the other grammars the morphological description adds little or nothing to the phonological insights provided, being concerned principally with matters such as affix shape and order, paradigm classes and so on.

All of these grammars provide an inventory of vowels and consonants, but only

9 of them furnish any evidence of the contrastive status of the segments listed. In the majority of cases the phonetic realizations of each of these units is described only by a categorical label and/or by the choice of a symbol for its representation. In 9 of the 20 grammars there is no commentary on phonetic realization of the segments going beyond the placement of a symbol on a labeled chart. The result is frequent ambiguity and imprecision. In at least 14 of the grammars there are one or more quite significant uncertainties about what kind of segment is intended by a symbol or label (some of these uncertainties will be described in a later section.)

Only three of the grammars include anything at all in the way of phonetic documentation. Two, those on Meithei and Kisi, include some exemplification of F0 patterns in tones, and one, on Lekeitio Basque has some documentation of intonation patterns. Childs's grammar of Kisi also presents some data on vowel duration but is not specific on the number of measurements made. Apart from this there is no hard data on any phonetic characteristics of the language being described in these grammars. For most of these languages, as for the great majority of the world's languages, there is little or no specialist literature on their phonetics to fill the gaps.

The relatively short shrift given to phonological topics in these grammars is regrettable but the almost total neglect of phonetics, often to the point of even failing to provide clear descriptions of typical segmental realizations, is even more distressing. Factors in the current professional socialization processes of the discipline of linguistics partially account for this situation. Many modern linguists receive mininal training in phonetics, especially if they elect to specialize in syntax. Many doctoral programs in linguistics implicitly or otherwise encourage students to select an identity primarily as either a "Ph" (phonetics and phonology) person or an "S & S" (syntax and semantics) person. Since a good grasp of the syntax of a language is fundamental to writing a grammar, not surprisingly many of those who write grammars have their deepest professional education in syntax. For some, enough phonology to develop a workable transcription is the limit of their ambition in dealing with the sound system of the language.

But this is probably far from the only factor involved. It seems that the 'market' for grammars has changed in a significant way during the past century. In the earlier decades of the twentieth century a good proportion of the linguistic description being published was targeted at those who wished to actually learn to speak the language being described, including in the case of the more 'exotic' languages such people as anthropologists preparing for field work, missionaries, and colonial administrators. Also at this period phonetics and phonology were not divorced from each other, as they later became with the development of structural linguistic theories that treated the sound structure of languages as primarily a system of contrasts between abstract phonemes or features, and moreover one in which minimizing the number of entities involved was a highly prized goal. Consequently earlier descriptions often aim to provide enough guidance to enable a native speaker's pronunciation to be imitated and they avoid the reductionism which later structuralism encouraged.

In the second half of the century, grammars are targeted primarily to an audience of professional linguists, phonology is regarded as a separate sub-discipline from phonetics, and language manuals take over the role that grammars used to play for those trying to learn to speak a language. Some grammars of the 1960's and 1970's, e.g Carrell 1970, contain no phonetic information whatsoever and are largely impenetrable even to linguists trained a generation later. Although in the closing decades of the century, some of the major movements in phonology began to re-emphasize the phonetic foundation on which sound patterns rest, this concern is not reflected in any overall greater attention to phonetics in grammars of this period, as we have seen.

In the next section of this paper it will be argued that explicit phonetic documentation should be a basic part of any grammar which aims to give a general description of a language. The third section of the paper discusses briefly what phonetic properties should be the prime targets for investigation and exemplifies some of the techniques for obtaining appropriate data under field conditions.

2. The need for phonetic documentation

There are several reasons why a grammar should be regarded as incomplete when it lacks careful phonetic description and documentation. The first is that when segmental phonetic properties are only characterized by symbols and descriptive labels the results are often imprecise or ambiguous, or even uninterpretable. But in any case more than precise segmental descriptions are required. The phonetic properties of a language are the foundation for non-arbitrary patterns in its phonology. Morphemic or syntactic constituency, or the organization of phrases into larger phonological units may be indicated by phonetic boundary markers or other phonetic variations. Segments vary in context and interact with each other and with non-segmental properties in distinct ways in different languages, and this too is part of the grammar of the language. In short, languages have a "phonetic grammar" and the highlights of this should be covered in an overall grammar, just as the highlights of its syntactic and other more conventionally "grammatical" properties are covered.

The frequent use of ambiguous symbols and the occurrence of descriptions which are vague or worse is partly the fault of the standard frameworks of phonetic and phonological description, but is often aggravated by lack of precision by users. For example, the grid of categories and symbols provided by the IPA in its chart of consonants is unhelpful in a number of ways. A case in point is the column headed 'retroflex'. This term has been used for consonants with three fairly distinct articulatory configurations (Ladefoged & Maddieson 1996); one in which the underside of the tongue tip makes contact with the front part of the hard palate (e.g. as in Telugu or Arrernte), one in which the tongue tip makes contact a little behind the alveolar ridge (e.g. as in Hindi or the Norwegian 'retroflex' consonants written 'rd,' 'rn,' 'rs', etc); and one in which the articulation is a laminal post-alveolar (as in the 'retroflex affricates' of Polish and Mandarin). In fact, the way the IPA consonant chart is laid out tends to discourage careful description of consonants in the coronal area as a whole. Distinctions of place (dental, alveolar, post-alveolar), tongue position (apical, laminal), and sibilance, which may be crucial, are among those conflated on the chart. A symbol such as /t/ or /s/ may casually be used without specifying the actual pronunciation that normally occurs.

A symbol may also be ambiguous if it has been employed with different values in different traditions. For example, the letter /c/ is used in the IPA for a voiceless palatal plosive, but in Americanist tradition it usually represents a voiceless dental or alveolar affricate. If an author fails to specify which tradition they are following the reader must infer or guess the answer, and will not always get it right.

Further problems can be caused by the use of labels that have never had a very specific meaning, or have had their meaning 'bleached' by being constantly (mis)used in imprecise ways. These include terms such as 'fortis' and 'lenis'. Unless it is explicitly spelled out, a reader does not know what mix of voicing, duration, amplitude, and other properties are intended to be conveyed by these terms. Similarly, when vowels are described as differing with respect to the phonological feature [ATR] it is rarely clear if the language truly has independent use of tongue root position among its vowel parameters or if this feature is being used essentially as a diacritic prop to disguise an inadequate set of vowel height features.

Individual grammars may add to problems of this general nature by using terms in vague or idiosyncratic ways, or in other ways failing to provide sufficient information. For example, in Merlan's Wardaman grammar the symbols /k, g/ are placed on a chart of consonants in a column headed "velar" but the text states that "velars are dorso-palatal." Does this mean that these are actually true palatal plosives similar to Hungarian /c, ɟ/, or fronted velars similar to Kwakw'ala /k̟, g̟/, or perhaps velars with a palatal offglide, /kʲ, gʲ/, similar to the palatalized velars of Hausa? In both the short Nivkh and Evenki grammars mentioned above, the symbol /h/ is placed in a column headed "pharyngeal." In Russian linguistic tradition pharyngeal and laryngeal segments are often grouped into a single category, but without further information, knowing this fact doesn't resolve the uncertainty of what /h/ is intended to represent in these cases. Laryngeal and pharyngeal segments are also very obscurely described in van Eijk's Lillooet grammar. Liquids are particularly often inadequately characterized. Dutton's Koiari description uses the symbol /r/ and provides no description of it beyond "vibrant" (it quite likely is a post-alveolar lateral flap). Bulatova & Grenoble's Evenki grammar uses the symbols /l/ and /r/ but does not describe their phonetic nature in any other way, and this imprecision is compounded by a proof-reading error which sees both these symbols placed on the chart of consonants in a column headed "bilabial."

Even when attempting to provide phonetic information, ambiguities often occur. Du Feu's Rapanui grammar describes glottal stops in intervocalic position as being pronounced as creaky voicing "which shows up as white noise on spectrograms." The problem here is that creaky voice does not have an acoustic pattern that is anything like white noise, which would be closer to a description of [h]. The reader doesn't know whether Rapanui has creaky voice or an [h]-like pronunciation in the position being discussed.

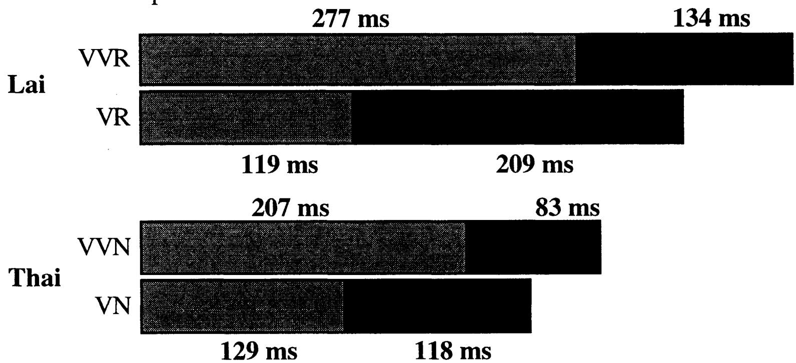
Outside the sample of grammars examined above, a favorite example of a difficult-to-understand description of segments comes from Stell's (1972) discussion of the phonology of Nivakle (also known as Axluxlay or Chulupi), a Mataguayan language spoken along the Argentine/Paraguay border. Stell gives puzzling descriptions of two segments transcribed as [x͡ɬ] and [k̠͡l] and described respectively as "A voiceless oral dento-alveolar velarized lateral fricative (with simultaneous articulation of a voiceless velar fricative)" and "A voiceless oral dento-alveolar velarized lateral affricate (with simultaneous articulation of a voiceless velar stop)" (my translations). There are multiple unclarities in both of these descriptions. The simultaneities claimed are extremely improbable, but it is unclear if what is being described are sequentially-complex elements, or perhaps laterals with a primary constriction at the velar place of articulation — the affricate maybe being similar in articulation to the velar lateral ejective affricate of Zulu.

The frequent ambiguities and unclarities in the literature provide a strong reason for wishing that field workers would do a better job of describing the phonetic properties of the languages they are working on. However, besides this rather negative argument there are very strong positive reasons to encourage closer attention to phonetics. Aside from the strictly contrastive distinctions that form the core of a phonological characterization of a language there are many other aspects of the sound pattern that must be correctly controlled by a native speaker and hence are part of what the speaker must learn in order to become fluent. In short, languages have a "phonetic grammar" as much as they have a syntactic one.

Consider the case of stop releases: languages differ in when a stop will ordinarily have an audible release. Tlingit requires an audible release of pre-pausal stops, whereas the audible release is often suppressed in this position in English. In Tlingit the first of two abutting stops within or between word boundaries also has an audible release in the great majority of cases if it is not homorganic with the second (Maddieson & Smith 1999). Thus in a phrase like /tɬeet kaa/ "white man" the final /t/ of /tɬeet/ is released before the closure for the velar plosive is made. On the other hand in an English phrase such as "white car" the velar closure for the initial plosive in "car" is almost invariably formed before the alveolar stop at the end of "white" is released (and that /t/ may also have its audible release further suppressed by an overlapping glottal closure).

There are also striking differences between languages with respect to the implementation of quantity contrasts, whether between long and short vowels or single and geminate consonants. Not only do languages differ in the ratios of long to short C and V durations but also in properties such as how durations of adjoining segments are adjusted in relation to quantity contrasts. Hakha Lai (Maddieson 2002) and Standard Thai (Mixdorff et al 2002) both show a compensatory duration pattern in which coda consonant durations are longer after a short vowel and shorter after a long vowel. But the pattern is more marked in Lai — final consonants after short vowels in Lai are actually longer in duration than the preceding vowel — and hence following consonant duration can be expected to play a larger role in cuing the vowel quantity distinction in Lai than in Thai. Mean duration of long and short vowels (light stippling) and following sonorants (heavy stippling) from the two studies cited are shown in (3) (The Lai data include words with final nasals and laterals, the Thai data only final nasals).

(3) Duration compensation in coda sonorants in Hakha Lai and Thai.



A few of the phonetic patterns by which different languages differ can be briefly mentioned for exemplification. One is vowel-to-vowel coarticulation: languages differ in the extent to which one vowel affects the near edge of another across an intervening consonant. For example, in Eggon (a Platoid language in the Niger-Congo family) there is strong anticipation of V2 in V1 especially if V1 is unrounded and V2 is rounded (personal fieldwork). In Sotho, with a broadly similar 7-vowel system, there is little effect of this kind of one vowel on its neighbor (Manuel 1990). Languages also differ in the alignment of tonal targets with segmental landmarks. For example, Mandarin (Xu 1997) appears to have later target alignment than Thai (Gandour et al 1994), which in turn has later alignment than Lai (Maddieson 2002). To speak a language with native-like pronunciation, all such patterns in its "phonetic grammar" must be learned.

The phonetic patterns of a language also provide the foundation for non-arbitrary patterns in its phonology. For example, the tendency for the retroflex stops consonants to be absent from word-initial position in Australian languages has been noted often. This distributional restriction becomes understandable when the precise phonetic properties of these 'retroflexes' are understood: they are characterized by markedly shorter closure duration than other coronal stops and by a forward movement of the tongue tip in the releasing gesture, resulting in differences in onset and offset formant transitions. These cues to segment identity are not detectable when there is no preceding context (Anderson 2000).

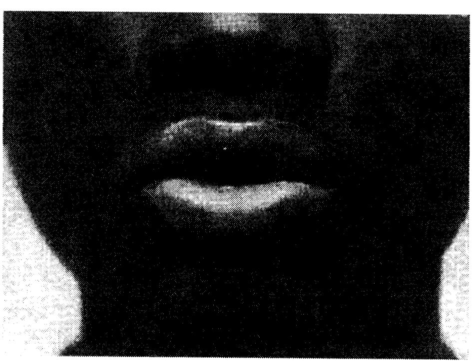
There is often also phonetic marking of morphemic or syntactic constituency, of sentence type, and other aspects of varying levels of linguistic organization. For example, different degrees of final lengthening depending on the rank of a following prosodic boundary have been found in English and other languages (e.g Beckman & Edwards 1990). Sentence types are frequently marked by intonation, and in some languages (e.g. French, Leggbo and Ghotuo) questions may be distinguished from statements by no other property.

All these considerations argue for greater care and more attention to be devoted to phonetic aspects of a language in a basic description.

3. Phonetics in the field

Phonetic documentation as a part of field work has a considerable history. Classic field studies from the early decades of the twentieth century not infrequently include explicit evidence of articulatory or acoustic properties of the language studied. For example Goddard's (1907) work on Hupa includes a number of palatograms obtained using a moulded artificial palate in order to show details of the contact position of coronal consonants. Boas's grammatical outline of Tlingit (1917) included kymograms to show timing of aspiration and the relative fundamental frequency values of high and low tones. Doke's pioneering work in the 1920's on Shona includes photographs of the position of the lips in several sounds, including one to illustrate the so-called "whistling fricatives" reproduced here as (4). This clearly shows that the lips are not in a typical rounded position in these sounds (Ladefoged and Maddieson 1996).

(4) Lip position in Shona "whistling fricative" from Doke (1931)



The value of such documentation is twofold. First it enriches the description by providing more detail, but crucially it also provides evidence that can be independently evaluated. Thus, by including the photograph in (4) Doke does more than just show the lip position, he shows it it in a way that can be scrutinized by others who can therefore decide for themselves if a description provided is appropriate.

The remainder of this section will discuss suggestions about what phonetic data should be included in a field report, as well as reviewing some of the issues that phonetic investigation can address, exemplifying the phonetic investigations that can be readily carried out in field situations, and discussing some of the practicalities involved. An ideal to have in mind is to provide enough phonetic information that a reader would be able to sound like a native speaker. A grammar writer might imagine he or she is describing the speech of the last generation of fluent speakers and aims to do it well enough that their grandchildren will be able to learn to sound like them and react like them to hearing the language spoken. Like other ideals, this is unattainable, but serves as a goal to reach towards.

The target should be to show representative speech behaviors of the linguistic community. But simple exemplification of raw phonetic patterns is not helpful. Phonological perspectives, especially contrast, are required to provide the framework for organizing phonetic data into appropriate categorical form. The phonetic investigation should therefore address such matters as the basis of syntagmatic contrasts, the nature of contextual segmental (or tonal, etc) variability, and the phonetic marking of prosodic or other constituency relationships. Phonetic investigations begin with direct use of the visual and aural senses, followed by the collection of audio exemplars. They can also encompass acoustic, articulatory and perceptual investigations.

Sound recordings provide the most basic form of phonetic documentation. Recordings, especially of texts, also serve to record lexical, syntactic and other data. The recording environment and the type and positioning of the microphone used make more difference to resulting recording quality than the choice of recording medium and machine (Maddieson 2001). A recording of a single speaker intended for careful listening or acoustic analysis is often best made using a head-mounted microphone. The close-range pick-up minimizes background noise, and the speaker-to-microphone distance is kept fairly constant so that relative amplitudes are more faithfully preserved than when the speaker is free to move or turn away from a free-standing microphone.

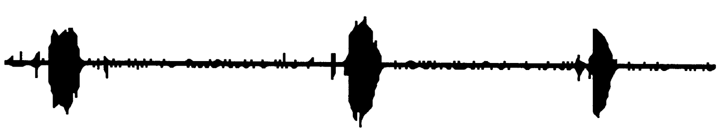
To make a good recording one should seek the least reverberant environment free of continuous noise sources. In some situations this is likely to be an outdoor setting. Sound reverberates in closed spaces, especially when the surfaces (walls, floors) are smooth and hard. This can cause considerable problems, particularly when one wishes to make measurements in the time domain. Indoor spaces are often noisier than may initially be apparent. Human hearing can focus selectively on signals of interest and "tune out" a lot of background noise; a microphone picks up every sound in its range. An important skill to develop is to listen for any persistent interfering sources of sound, such as wind noise, a refrigerator running, building air-conditioning, noisy flurorescent lighting, traffic noise, etc, and eliminate or avoid as many as possible. Short-term noises — the crowing rooster, or single passing truck — are less of a problem than these persistent noise sources. You can always re-record an item if a short-term noise interrupts, but persistent noise overlays the entire recording.

Outdoor recording is illustrated in (5), taken during fieldwork on Wichí in collaboration with Marisa Censabella (photo courtesy of Jose Braunstein). Relatively constant levels of wind and insect noise, and intermittent bird calls and voices of other speakers were considerably muted by the use of the head-mounted short-field microphone worn by the speaker being recorded in the foreground, and reverberation was eliminated. A sample of the resultant waveform is shown in (6).

(5). Outdoor recording of Wichí at Lote 42, Las Lomitas, Argentina.

This waveform covers about three and half seconds and contains three repetitions of the word /t'u/ "grave marker." The relative salience of the voice with respect to the background noise is apparent. The signal-to-noise ratio is good enough that quite precise measurements can be made of properties such as the timing of the glottal release and the onset of voicing after the initial ejective stops. Other available settings or different equipment would have yielded considerably less usable recordings.

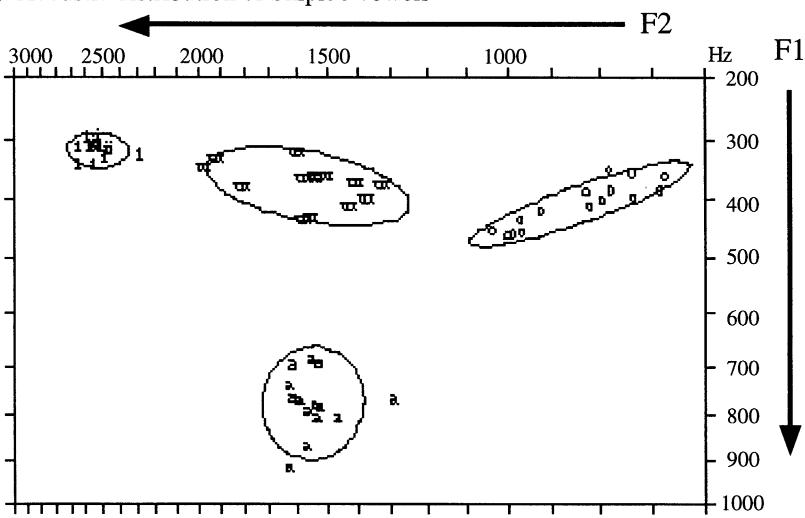
(6). Sample of waveform (see text)



Analysis of timing and other acoustic properties is now easily performed by anyone with access to a personal computer using readily available commercial or free software. A certain level of experience is required to avoid the risks of misinterpretation, but linguists who may feel they lack this experience will find their more phonetically-trained colleagues eager to assist.

Acoustic analysis is an especially good way to provide objective data on vowel quality as well as being an easy way to obtain accurate timing information. Basic information on vowel formant frequencies not only supports decisions on how to categorize vowels but also provides data on more gradient properties such as their relative positioning. A plot of the four vowels of Shipibo (Valenzuela et al 2001) in a perceptually-scaled two-formant space is shown in (7). Each symbol represents a single token while the ellipses enclose an area encompassing roughly 90% of the variance among tokens of each vowel type for this speaker.

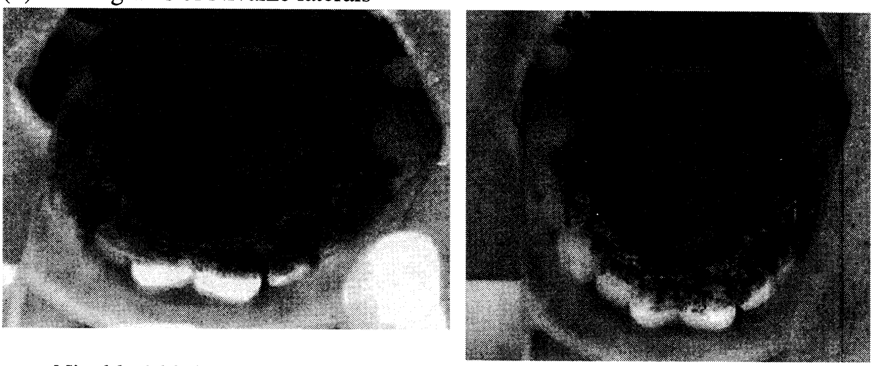
(7). Acoustic distribution of Shipibo vowels



As (7) shows, the three non-low vowels are all quite distant from the single low vowel, but become progressively lower as one moves further back in the acoustic space. This provides much more subtle information on these sounds than is provided by, as here, choosing the symbols /i ɯ a o/ to represent them. The reason for hesitation between representing the back rounded vowel with the symbol /o/ or with the symbol /u/ in the literature on this language becomes apparent.

As remarked earlier the articulation of coronal segments is often poorly described. A good way to document their articulatory position is to use the old-established technique of palatography, as well as linguography, its counterpart for seeing the part of the tongue involved in making the arrticulation. This is an easy field operation which subjects often find quite entertaining. Ladefoged (1997) describes how to do it in some detail, but a simple photograph of a sample palatogram can be sufficient to answer important questions. Recall the speculation, based on Stell's discussion of Nivakle, that this language might have velar laterals. Recent fieldwork employing a combination of acoustic analysis and palatograms has been able to show that the segments given such ambiguous descriptions by Stell consist of less exotic alveolar laterals, as demonstrated in (8). The dark markings behind the teeth visible in the reflection of the palate in the mirror in the two pictures in (8) show where the tongue, painted with an olive oil and charcoal mixture, contacted the alveolar ridge in pronouncing the two words illustrated.

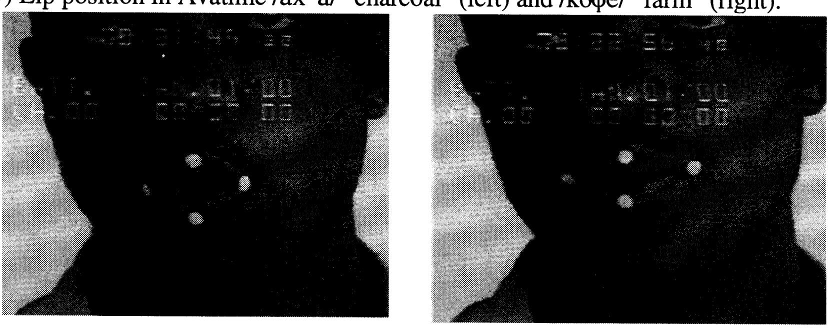
(8). Palatograms of Nivakle laterals



Nivakle /ekle/ "parrot" — simple sequence of /k/ + /l/

Nivakle /ɬa/ "fruit" — voiceless alveolar lateral fricative

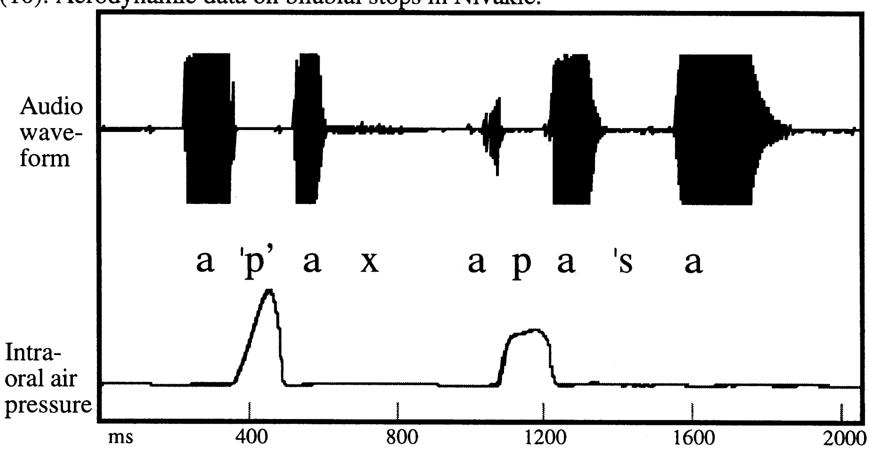
9) Lip position in Avatime /axʷa/ "charcoal" (left) and /koɸe/ "farm" (right).



The pictures in (8) were filmed with a video camera, but direct video recording of a speaker in action is also is a good non-invasive way to document certain articulatory movements, especially of the lips. A stabilized rather than hand-held camera and a constant camera-speaker distance greatly facilitate comparison across utterances, so a tripod or other device should be used and the speaker seated on a firm chair facing the camera, or at right angles to the camera for a side view. Position yourself in front of the speaker to keep his or her attention directed and cue what you want said without having the speaker look down to read. Putting small markers on the face makes it easier to track movements. If a ruler placed in the plane of the speaker's face is also filmed it then becomes possible to measure distances between points and calculate the extent of movements. Two frames from a video study of Avatime (Maddieson 1998) are shown as (9). These show the maximum contriction of the lips in the consonants /xʷ/ and /ɸ/ and show how unlike each other these consonants are, although previously they had been described as both having a bilabial fricative component. The aperture between the lips is just 1.2 mm high in this token of /ɸ/ but is 5.3 mm high in the example of /xʷ/, during which the lips are also obviously protruded in a rounding gesture.

Most of the techniques used by phoneticians to track articulatory movements in detail, such as electromagnetic articulography or MRI imaging, are not adapted to use in field environments. However, comparatively simple computerized field units, such as the MacQuirer and PCQuirer from SCICON, are available to collect data on the aerodynamics of speech production. This data can provide valuable insights on articulatory timing, on laryngeal control, on airstream mechanisms in use, on nasalization, and so on.

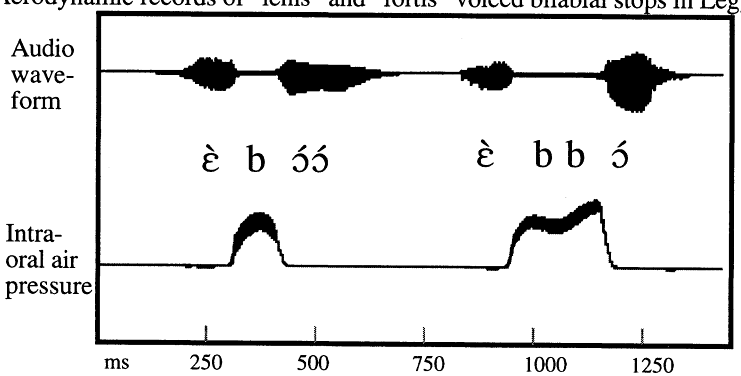
(10). Aerodynamic data on bilabial stops in Nivakle.



Among the many rather imprecise phonetic labels that are encountered is the term "glottalized." When applied to stops this may imply use of the ejective mechanism, with the raising of the larynx creating the pressure behind the closure for an explosive release, or something quite different, such as the "glottalization" of voiceless stops in English, which blocks an audible release of the oral closure by preventing pressure building up behind it. Nivakle is one of many languages where stops have been described as "glottalized." Air pressure behind the lip closure in bilabial stops was observed by having speakers say words while holding a thin plastic tube between their lips and recording the result to a computer using MacQuirer. Sample records of the words /a'p'ax/ "snake" and /apa'sa/ "your lips" are shown in (10). In the first of these, on the left of the figure, the pressure within the oral cavity builds up rapidly to a high level and declines as soon as the peak is reached. By contrast, in the second word the pressure builds up to a lower peak but maintains an elevated level for a relatively extended time. The latter is the pattern commonly found in plosives, where the pressure increase comes from the nearly-constant driving force of the lungs. The more 'ballistic' pressure curve seen in the first word is the signature of an ejective stop, in which the pressure increase is generated by a short-term movement of the closed larynx upwards. Acoustic records also show the release of the glottal closure before the onset of the vowel. Together these data document that this word contains a true ejective stop.

Aerodynamic records sometimes reveal facts that are difficult to observe in other ways. In Leggbo, a Cross River language studied in the field methods class at UC Berkeley during the 2001-2 academic year, consonants have been classified as "fortis" and "lenis" (Bendor-Samuel & Spreda 1969). Work by myself and Julie Larson has shown that the consonants labelled "fortis" have about twice the acoustic duration of their "lenis" counterparts (where pairs exist), suggesting that a transcription as geminate vs singleton consonants is justifiable, as in the words /ɛ̀bbɔ́/ "branch" vs /ɛ̀bɔ́ɔ́/ "hand." Aerodynamic records reveal something that cannot be seen in the acoustic data: the "fortis" closure in /ɛ̀bbɔ́/ has two separate pressure peaks, suggesting that it is organized more like a sequence of two separate segments than like a single lengthened one. This finding reinforces the case for considering that the "fortis" /bb/ is serving as both a coda and an onset.

(11). Aerodynamic records of "lenis" and "fortis" voiced bilabial stops in Leggbo.

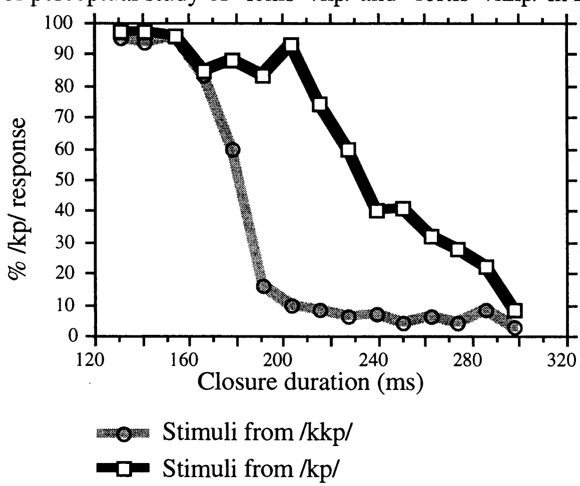


The final type of phonetic data that will be discussed relates to hearing and perceiving speech. Auditory processing of sound is broadly similar for all people (without significant hearing loss), but perceptual processes are greatly influenced by linguistic experience and are less universal. When a listener can see a speaker's face there is also the influence of the visual information on perception to consider.

It can be quite easy to devise and carry out perceptual experiments in the field by presenting unaltered or edited recordings for listeners' judgments. Such experiments can, among other things, search for native speakers' impressions of perceptual similarity, explore which acoustic characteristics play a greater role than others in cueing recognition, and test for the location of category boundaries. Common models for such experiments would include simply asking subjects to identify what they hear, having them assign a 'goodness rating' to stimuli, or asking if they can discriminate between two samples. Natural tokens or ones edited with software designed for the manipulation of sound can be used, depending on the target of investigation.

An experiment prepared by Julie Larson and administered in the field by Imelda Udoh was designed to examine if other properties of the "fortis/lenis" consonant pairs in Leggbo besides the duration contribute to their identification. As part of this experiment, natural tokens of two minimally distinct items, /ɛkpa/ (a man's name) and /ɛkkpa/ "he paid" were edited by adding or removing some of the silent duration in the middle of the consonants closure (/kkp/ represents a fortis, or geminate, voiceless labial-velar plosive). Two continua were thus produced with consonant durations ranging from about 130 ms to about 300 ms, these end-points being the natural durations of the two original tokens. Properties of the consonant onsets and releases were not modified. If duration is the only parameter involved in the contrast, then tokens lengthened from /kp/ and tokens shortened from /kkp/ should both reveal the same critical duration value at which the preponderance of judgements switches from "lenis" to "fortis."

(12). Results of perceptual study of "lenis" /kp/ and "fortis" /kkp/ in Leggbo.



Nine subjects gave their judgments as to which word they heard in response to each token. The results in (12) show that duration is only one factor that governs the perception of the distinction between these words. Lengthening /kp/ by about 100 ms switches its perception to /kkp/, showing that duration can be a sufficient cue to "fortisness," but shortening /kkp/ does not shift its perception to /kp/ until the duration is down to about 180 ms. Evidently other properties than duration contribute to the distinction between these sounds. This finding in turn motivates a more extensive study of the acoustic and articulatory differences between such pairs to discover what these properties might be.

4. Summary

The intent of these remarks is obvious — to encourage more attention to the description of the basic phonetic facts of the world's languages as part of the reviving interest in language diversity. The paper has attempted to do this both by showing where much linguistic data collection fails in this regard and by showing how easy it can be to do better. This author and many of his phonetician colleagues stand ready to help other linguists to take the plunge into documenting the phonetic patterns of the languages they study if that help is needed.

Acknowledgements

Thanks to the National Science Foundation for supporting the work reported here through grant BCS-9817345 for research on the phonetic structures of endangered languages.

References

Abbott, C. 2000. *Oneida*. LINCOM EUROPA, München and Newcastle.

Anderson, V. B. 2000. *Giving Weight to Phonetic Principles: The Case of Place of Articulation in Western Arrernte*. Ph D. dissertation, University of California, Los Angeles.

Beckman, M. E. & J. Edwards. 1990. Lengthenings and shortenings and the nature of prosodic constituency. In J. Kingston & M. B. Beckman (eds.) *Papers in Laboratory Phonology 1: Between the Grammar and Physics of Speech*, 152-177. Cambridge University Press, Cambridge.

Bendor-Samuel, J. T. & K. W. Spreda. 1969. Fortis articulation: A feature of the present continuous verb in Agbo. *Linguistics* **52**, 20-6.

Bergsland, K. 1997. *Aleut Grammar: Unangam Tunuganaan Achixaasix̂*. Alaska Native Language Center, University of Alaska, Fairbanks.

Boas, F. 1917. *Grammatical Notes on the Language of the Tlingit Indians*. University of Pennsylvania, Philadelphia.

Carrell, P. L. 1970. *A Transformational Grammar of Igbo*. Cambridge University Press, Cambridge.

Chelliah, S. L. 1997. *A Grammar of Meithei*. Mouton de Gruyter, Berlin and New York.

Childs, G. T. 1995. *A Grammar of Kisi, A Southern Atlantic Language*. Mouton de Gruyter, Berlin and New York.

Creider, C. A. & J. T. 1989. *A Grammar of Nandi*. Helmut Buske, Hamburg.

Dedrick, J. M. & E. H. Casad. 1999. *Sonora Yaqui Language Structures*. University of Arizona Press. Tucson.

Doke, C. M. 1931. *A Comparative Study of Shona Phonetics*. University of the

Witwatersrand, Johannesburg.
Du Feu, V. 1996. *Rapanui*. Routledge, London and New York
Dutton, T. E. 1996. *Koiari*. LINCOM EUROPA, München and Newcastle.
van Eijk, J. 1997. *The Lillooet Language: Phonology, Morphology, Syntax.* University of British Columbia Press, Vancouver.
Gandour, J., S. Potisuk, and S. Dechongkit. 1994. Tonal coarticulation in Thai. *Journal of Phonetics* **22**: 477-92.
Goddard, P. E. 1907. The Phonology of the Hupa Language: Part 1, The Individual Sounds. *University of California Publications in American Archaeology and Ethnology* 5: 1-20.
Gruzdeva, E. 1998. *Nivkh*. LINCOM EUROPA, München and Newcastle.
Haspelmath, M. 1993. *A Grammar of Lezgian*. Mouton de Gruyter, Berlin.
Hinton, L. & K. Hale (eds.). 2001. *The Green Book of Language Maintenance and Revival.* Academic Press, San Diego.
Hualde, J. I., G. Elordieta & A. Elordieta. 1994. *The Basque Dialect of Lekeito.* Universidad del Pais Vasco, Bilbao.
Kruspe, N. D. 1999. *Semelai*. Doctoral dissertation. University of Melbourne, Melbourne.
Ladefoged, P. & I. Maddieson. 1996. *Sounds of the World's Languages.* Blackwells, Oxford.
Ladefoged, P. 1997. Instrumental techniques for linguistic phonetic fieldwork. In W.J. Hardcastle & J. Laver (eds.) *Handbook of Phonetic Sciences*, 137-66. Blackwells, Oxford.
Macaulay, M. 1996. *A Grammar of Chalcatongo Mixtec.* University of California Press, Berkeley and Los Angeles.
Maddieson, I. & C. L. Smith. 1999. The stops of Tlingit. To appear in *Living Language, Living Culture: Language Maintenance and Revitalization in Athabaskan Communities*. University of New Mexico Press, Albuquerque.
Maddieson, I. 1998. Collapsing vowel harmony and doubly-articulated fricatives: Two myths about Avatime phonology. In I. Maddieson and T. J. Hinnebusch (eds) *Language History and Linguistic Description in Africa (Trends in African Linguistics 2)*, 155-66. Africa World Press, Trenton, NJ.
Maddieson, I. 2001. Phonetic fieldwork. In P. Newman and M. Ratliff, eds, *Linguistic Fieldwork,* 211-229. Cambridge University Press, Cambridge
Maddieson, I. 2002. Timing and alignment: A case study of Lai. Paper presented at the 8th International Symposium on Chinese Languages and Linguistics, Academia Sinica, Taipei, November 8-10 2002.
Manuel, S. 1990. The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *Journal of the Acoustical Society of America* 88: 1286-98.
Maslova, E. 1998. *Kolyma Yukaghir*. Doctoral dissertation. Universität Bielefeld, Bielefeld.
Merlan, F. C. 1994. *A Grammar of Wardaman, a Language of the Northern Territory of Australia.* Mouton de Gruyter, Berlin and New York.
Mixdorff, H., S. Luksaneeyanawin, H. Fujisaki & P. Charnvivit. 2002. Perception of tone and vowel quantity in Thai. *Proceedings of the 7th International Conference on Spoken Language Processing*, Denver: 753-6.
Morse, N. L. & Maxwell, M.B. 1999. *Cubeo Grammar (Studies in the*

Languages of Colombia 5). Summer Institute of Linguistics and University of Texas, Arlington.

Newman, P. & M. Ratliff, eds. 2001. *Linguistic Fieldwork.* Cambridge University Press, Cambridge.

Schuh, R. G. 1998. *A Grammar of Miya.* University of California Press, Berkeley and Los Angeles.

Stell, N. N. 1972. Fonología de la lengua axluxlaj. Centro de Estudios Lingüísticos, Universidad de Buenos Aires, Buenos Aires.

Terrill, E. 1999. *A Grammar of Lavukaleve: A Papuan Language of the Solomon Islands*. Australian National University, Canberra.

Valenzuela, P. M., L. M. Pinedo & I. Maddieson. 2001. Shipibo. *Journal of the International Phonetic Association* **31**: 281-285

Vaux, B. & J. Cooper. 1999. *Introduction to Linguistic Field Methods.* LINCOM EUROPA, München and Newcastle.

Xu, Yi. 1997. Contextual tonal variations in Mandarin. *Journal of Phonetics* **25**: 61-83.

"Peculiar to Themselves": Idioms in the Dictionary[1]

PAMELA MUNRO
UCLA

0. Introduction

That preeminent Berkeley fieldworker Mary R. Haas declared that people describing a language should produce a grammar, a dictionary, and a body of texts. Of these, my particular love is the dictionary, in part because the dictionary making process generally winds up teaching me a lot about most aspects of grammar. The reason for this is, of course, that to prepare an insightful dictionary we need to know how to write, define, and classify (that is, analyze) the words we're collecting—and to do any of those, we must understand a lot about grammar.

Some items are especially difficult to put into a dictionary, either because it is hard to decide what form of the word or phrase to enter or because it is hard to decide how to translate the chosen entry and to explain or illustrate its grammatical use. In this paper (which I originally planned to call "How Should I Put That in the Dictionary?"), I will survey some of the problems various such expressions pose for the fieldworker constructing a bilingual dictionary of an unfamiliar target language.

Most of my examples in this paper, taken from dictionaries of four unrelated languages illustrating an extensive typological range—Chickasaw (Munro & Willmond 1994), Mohave (Munro, Brown, & Crawford 1992), San Lucas Quiaviní Zapotec (SLQZ; Munro & Lopez et al. 1999), and Wolof (Munro & Gaye 1997)[2]—will involve verbs, and in particular complex verbal idioms each of which has a meaning or syntax "peculiar to itself" (in the words of my American Heritage dictionary). I will consider idioms of the familiar semantically opaque type as well as other expressions whose non-standard syntax must be elucidated for the user, and will discuss "literal translations" and idiomatic structures influenced by majority contact languages. Finally, I will describe how an extensive dictionary corpus can illustrate families of related idioms and the restructuring of syntactic idioms to fit a standard model.

My purpose in this paper is not to dictate a certain way to put idiomatic expressions in the dictionary, though I will mention many relevant considerations.

---

[1] Thanks to everyone who has helped shape my ideas on the dictionaries and issues discussed here, in particular my principal collaborators, the late Nellie Brown, Dieynaba Gaye, Felipe Lopez, and Catherine Willmond, but also William Frawley, Michael Galant, Rodrigo Garcia, Larry Gorbet, Lynn Gordon, the late Ken Hale, Kenneth Hill, Judith Hummel, Margaret Langdon, Olivia Martínez, and Carson Schütze; Heriberto Avelino, Karen Dakin, Lance Foster, Russell Schuh, Miriame Sy, Harold Torrance, and members of the BLS audience provided additional useful input.

[2] I have shortened or otherwise adapted many of the dictionary entries presented as examples. Orthography and other features of each language are described in the dictionaries referenced.

Such decisions are clearly determined in part by the esthetic preferences of dictionary compilers and in part by the structure of the target language. Instead, this paper will be a reminder—to you and to myself—that dictionaries are much more than lists of words, and that a dictionary can be an important resource for linguists as well as the heritage community that may be its first and most basic audience.

My illustration of difficult dictionary entries will begin with something other than verbs, dictionary entries for "expletives" in SLQZ (a Zapotecan (Otomanguean) language of Oaxaca with VSO word order). As (1) shows, this dictionary is trilingual (but with a considerable English bias, as we'll see later), and the expletives exemplified here are not words like *it* and *there*, but exclamations or oaths:[3]

(1) **nih bèi'nyande'eh** (mild expl.) {Bribàa'iny Rumoro's zhi', nih bèi'nyande'eh "There was a landslide in Rumoroso, oh shoot | *Hubo un derrumbe en Rumoroso, que suerte*" }
§§ Zap. expl. are used in many different syntactic constructions (as described in the Introduction, section 4.37)..
nih wrihiny (mild expl.) {Xi nih wrihiny bèi'nyùu'? "What the heck did you do? | *¿Qué diablos hiciste?*" }
nih wzha'ahn, nih wrihiny nih wzha'ahn, nih wrih wzha'ahn (strong expl.) {Chih byehtënn, a nih wzha'ahn myee'gr ri'cygah, bdi'cah myee'gr zèenèeg myee'gr dannoohnn a'xta' Tijwa'nn steeby zhi' "When we got down, the fucking migras were right there, the immigration agents appeared and took us to Tijuana again | *Cuando nos bajamos, la chingada migra ya estaba allí, los agentes de inmigración aparecieron y nos llevaron a Tijuana otra vez*"; Que'ity rgui'ihzhdi' nih wzha'ahn la'anng lùa' "He (the son of a bitch) doesn't pay me | *El hijo de la chingada no me paga*"; Que'ity rgui'ihzhdi' nih wzha'ahn lìu' lùa' "You don't pay me, you son of a bitch | *No me pagas, tú, hijo de la chingada*"; Nih wzha'ahn lìu'! "Screw you! | *¡Maldito seas!, ¡Vete a la chingada!*" }

The entries in (1) show that SLQZ expletives are grammatically slippery. They seem to constitute an independent part of speech, with a morphological structure that looks like either a relative clause, a nominalization, or a certain type of quantifier (Munro & Lopez et al. 1999: 27-28), but unlike these items they can be used in an unexpected variety of syntactic constructions (underlines illustrating these have been added to the translations in (1)), but cannot occur alone as independent exclamations. Many of these uses are like those of comparable words in English or Spanish, but no English or Spanish expletives can be used in all these ways.

[3] Abbreviations in the dictionary entries include expl. = expletive, ger. comp. = verb with gerundive complement, impers. id. = impersonal idiom, impers. pssr. id. = impersonal possessor idiom, intr. = intransitive, irr. = irrealis, mod. sscmp. = verb with modal same-subject complement, obj. id. = object idiom, poss. id. = possessor idiom, poss. obj. = possessed object idiom, prep. v. = preposiitonal verb, pssd. = possessed, pssr. obj. id. = possessor object idiom, sg. = singular, subj. = subject, subj. poss. = subject possesses object idiom, tr. = transitive, Zap. = (SLQ) Zapotec, 3-at = nominative marked noun. Some of these are discussed in the text; the remainder are explained in the dictionaries referenced. Examples in the dictionary entries are translated but not analyzed. Abbreviations in other examples include acc = accusative, asp = aspect, dat = dative, hab = habitual, neut = neutral, nom = nominative, obj = object, perf = perfective, pron = pronoun, prox = proximaate, pt = past/perfective, s = singular, ss = same subject, Vfoc = verb focus, with 1, 2, 3 used for person and I, II, III used for the Chickasaw inflectional classes.

In this case, I've chosen a cop-out: I just give illustrative examples. In the rest of this paper, I will concentrate on issues involving verbs, where this cop-out should not be available: if a verb or verb phrase exists in the language, speakers are able to use it, and the ideal dictionary must tell a learner how they do that.

2. Verbs in the Dictionary

Dictionary makers often have to deal with many difficult analytical issues in making decisions about dictionary entries for verbs. The first thing is to decide how to spell the words in the dictionary, how to select the "base" (dictionary entry) form of a verb, how to define this once it is selected (see Munro to appear),[4] and how to convey relatively regular inflectional information.

In Chickasaw (a Muskogean language with SOV word order spoken in Oklahoma) and Wolof (a West Atlantic language with SVO word order spoken in Senegal), verbs can be entered in the dictionary in the unaffixed ("bare") form, used in Chickasaw with third person subjects, in imperatives, and in infinitival-like complements, and in Wolof in non-negative main clauses:

(2) *Chickasaw*
imilhlha to be scared, wild, afraid (III)
malili to run (mainly sg. subj.), to go (of a machine); to make a run (in baseball); to run for office (I)
tikahbi to be tired (II)

(3) *Wolof*
fecc to dance (intr.)
laabir to be compassionate; to be open, friendly (of a person) (intr.)

For many languages, the decision as to what verb form to list in the dictionary is more complex. When a bare stem form of verbs is not an independent word, I believe it should not be used as a dictionary entry, a feeling generally echoed by native speakers. In Mohave (a Yuman language with SOV word order spoken on the California-Arizona border), for example, "realis" (citation) forms of verbs appear with either *-k* or *-m* (4);[5] these two suffixes are used both on independent verbs and in a variety of grammatical constructions, though they drop before other affixes:

(4) *Mohave*
imkwilyk=k get up (out of bed)
imkwiilyk=m lie in bed turning around; be upset (of the stomach) § 'Iitonych imkwiilyktm. My stomach is upset. | 'Imkwiilykm. I lie in bed tossing and turning.

In SLQZ, verbs have six or seven "aspectual" forms differentiated by prefixes. In our dictionary, verbs are listed in the habitual (5), though this means that a third of the dictionary entries start with *r-* (see Munro 1996):

[4] I cannot consider here reasons why in some languages it is best to define verbs with 'to' but in others it is not (but see Munro to appear). (Regarding selection of a base, see also Albright 2002.)
[5] The analysis described here was influenced by the work of Gordon (1986) on Maricopa and by the form in which verbs are cited in Gordon, Heath, and Munro (in preparation).

(5) *SLQZ*
rbahb itches | *le pica (le da comezón)* {Ua's rbahb ni'a' "My foot really itches | *El pie de veras me pica*" }
rdùu'b sweeps | *barre*

Language-specific heuristics should tell the user how to use the listed dictionary entry verbs in sentences. In some dictionaries these are just assumed, but it is better if they are stated clearly somewhere, probably in the introduction to the dictionary.

If there are different inflectional classes of verbs (comparable to the different conjugation classes in many European languages), the dictionary entry should specify this, as with the Mohave *-k* or *-m* verbs. Chickasaw presents a more complex case. The language has a simple nominative-accusative case marking opposition for noun phrases (6) (accusative marking can be omitted in such sentences, with no apparent affect on meaning; cf. Munro 1999). But Chickasaw intransitive verbs exhibit three different morphological agreement patterns (7),[6] and transitive and ditransitive verbs can display an even greater range of inflectional possibilities, just a few of which are exemplified in (8). (The intransitive verb classes are largely semantically governed (Payne 1982), though there are many exceptions (Munro & Gordon 1982).)

(6) Ihoo-at chipot-a̱ sho̱'ka-tok.
woman-nom child-acc kiss-pt 'The woman kissed the child'

(7) Malili-li. 'I run' (inflection class I)
run-1sI
Sa-tikahbi. 'I am tired' (inflection class II)
1sII-be.tired
Am-ilhlha. 'I am afraid' (inflection class III)
1sIII.dat-be.scared

(8) Chi-sho̱'ka-li. 'I kiss you' (inflection classes I,II)
2sII-kiss-1sI
Chin-taloowa-li. 'I sing to you, I sing for you' (inflection classes I,III)
2sIII.dat-sing-1sI
Ofi'-a̱ sa-banna. 'I want the dog' (inflection class II, noun object)
dog-acc 1sII-want
Ofi'-a̱ am-alhkaniya. 'I forget the dog' (inflection class III, noun obj.)
dog-acc 1sIII.dat-forget
Ofi'-a̱ chim-a-li. 'I give the dog to you'
dog-acc 2sIII-give-1sI (inflection classes I,III, noun object)

The dictionary user also needs to know which word in multi-word verb entries is the (main) verb (for inflectional and other syntactic purposes). In Chickasaw (9) and Mohave (10), for example, the verb is the last word of such entries, while in Wolof (11) and SLQZ (12) it is the first word of the entry.

[6] We mark the three classes as (I), (II), and (III), as shown in (2). A similar approach is adopted by Martin & Mauldin (2000). Other dictionaries, such as Sylestine et al. (1993), follow the approach of listing the first person singular form (an especially good approach for Alabama, which not only has different classes but which allows different positioning of the agreement affixes). For more about these how we handle the data in (8) for Chickasaw, see Munro & Willmond (1994).

(9) *Chickasaw*

naalhpisa' oppani to break the law (I) § Naalhpisa' ishoppani ishhonkopakmat. If you steal you'll break the law.
˜ *naalhpisa'* 'law', *oppani* 'to break' :: 'to break the law'
The example's first verb has the second person singular I prefix *ish-*.

(10) *Mohave*

mavar suuvii=k make gravy § Mavar 'asuuviik. I made gravy.
˜ *mavar* 'flour', *suuvii=k* 'make mush' :: 'to make flour into mush'
The verb in the example has the first person prefix *'-*.

(11) *Wolof*

jot kepp to fit just right (tr.) {Jot na ma kepp. It fits me just right.}
˜ *jot* 'to fit', *kepp* (ideophone; no "literal" meaning) :: 'to fit'
The verb in the example is followed by the third person singular neutral clitic *na* and the first person singular object clitic *ma*.

(12) *SLQZ*

rdàa' guehehll harvests corn | *pizca maíz* {Caldàa' Gye'eihlly guehehll Mike is harvesting corn | *Miguel está pizcando maíz*}
˜ *rdàa'ah* 'breaks', *guehehll* 'cornfield' :: 'to break the cornfield'
The progressive aspect prefix *ca-* replaces habitual *r-* on the example's verb.

On the line beginning with ˜ after multiword dictionary entries I list the meanings of component words; at the end of this line, following ::, I give the "literal" meaning of the expression. The examples in (9-12) illustrate a variety of types of "idiomatic" expressions: none of them could be predicted on the basis of knowledge of the component words and their syntax. (9) is a transparent collocation (perhaps calqued from English), included in the dictionary since not all languages would express the concept with this metaphor; (10) is easily understood but unpredictable; (11) is a verb plus an ideophonic complement; and (12) is a completely unpredictable idiom. In each case, the example in the entry confirms that the first (9, 10) or last (11, 12) word of the entry is the verb, as explained following the ˜ line after each entry.

I'm now going to go on to discuss problems for this general approach to verb entries: cases where it is hard to write definitions and, in particular, idioms, both typical semantic idioms and "syntactic idioms" using nonstandard constructions.

3. Difficult Definitions

Difficulties arise when there is a mismatch between the syntax of the target language and the semantics or syntax of the defining language. The translation of SLQZ *ryu'làaa'z* (13), for example, is completely unproblematic in English—this verb is remarkably like English *like* or *love.* But only one of the four definitions of this verb (the second) was easy to translate into Spanish. Our dictionary followed a rule that the syntax of definitions should match that of entries—but the most natural way to say 'like' in Spanish is with the verb *gustar,* which expresses the liker as an indirect object and the like-ee as the grammatical subject. Therefore, in definitions 1, 2, and 4 the natural Spanish translation with *gustar* appears in parentheses following a syntactically more parallel definition. (I'll return to another *gustar* problem later.)

(13) **ryu'làaa'z** 1. likes (something) | *disfruta, quiere, ama (le gusta) (algo)* (tr.) {Ryu'là'aza' coloory nsua'll "I like (the color) blue | *Amo el (color) azul (Me gusta el (color) azul)*"}; 2. loves, desires (someone) | *quiere, desea, ama (a alguien)* {Ryu'làa'za' lìu' "I love you | *Te amo*"}; 3. likes to | *disfruta (le gusta)* {mod. sscmp.; used with irr.; Ryu'làaa'z Gye'eihlly ygyàa'ah "Mike likes to dance | *Miguel disfruta bailar, A Miguel le gusta bailar*"; Ryu'làa'zëng ygyàa'ng "He likes to dance | *Disfruta bailar, A él le gusta bailar*"}; 4. likes | *disfruta (le gusta)* {ger. comp.; Ryu'làaa'z Gye'eihlly gahll rgyàa'ah "Mike likes dancing | *Miguel disfruta bailar*"; Ryu'làa'zëng gahll rgyàa'ah "He likes dancing | *Él disfruta bailar*" }

The examples in (14) illustrate a different problem. The seemingly "logical" subject of these verbs is a non-surfacing agent (the subject's parents) for the first entry and a non-subject (the child) for the second. *Have* is a useful verb that often allows construction of English definitions in cases like these, but this approach is seldom possible in Spanish and may be awkward in English (in some cases, speakers interpret such 'have' constructions as causatives). In these Spanish definitions we used an 'is the one that' construction whose implications are different from those of the Zapotec and English, but which works to show the verbs' grammatical relations.

(14) **rahcgaan** has her betrothal arrangements negotiated (of a young woman) | *es por quien se negocian arreglos esponsales (dícese de una joven)* {A guhcgaan zhyàa'p "The girl has had her betrothal arrangements negotiated (her parents have agreed to the arrangements) | *La muchacha es por quien se han negociado arreglos esponsales (sus padres han aprovado los arregloss)*" }
rbèe' te'ihby laad has (his child) move out of the family home (for instance, when the child gets married) | *es a quien se le muda (un hijo o una hija) de la casa familiar (por ejemplo, cuando el hijo se casa)* {Blèe'ng zhìi'inyëng te'ihby laad "He had his child move out | *Él fue a quien se le mudó un hijo de la casa familiar (su hijo ya salió de la casa familiar)*" }
~ *rbèe'eh* 'puts', *te'ihby* 'one', *laad* 'side' :: 'to put to one side'

4. Verbal Idioms

Idioms raise similar, but often even more tricky problems of definition and presentation. An idiom, as we normally use the term, is a phrase whose meaning is not transparently computable from its component words.[7] I will be concerned here only with idioms containing verbs, like some of the examples already presented. In English, the majority of verbal idioms are verb phrases—they can be predicated of a subject. Familiar examples like *to bark up the wrong tree, to spill the beans,* and *to let the cat out of the bag* require only a single subject argument; they thus have an open (or, in the terminology of Marantz (1984: 27), free) subject position.[8]

Some similar idioms from Wolof and Chickasaw are given in (15) and (16):

[7] English idioms (cf. Marantz 1984; Nunberg, Sag, & Wasow 1994; Manaster-Ramer 1993) are cited with *to* before the verb. An included subject is preceded by *for;* with the following free position shown by an ellipsis. Other non-subject free positions are represented by "*(someone)*".
[8] Marantz calls these "object idioms".

(15) *Wolof*

bëgg dee to take lots of risks, live dangerously; to be exhausted, really tired (intr.)

˜ *bëgg* 'to want', *dee* 'death' :: 'to want death'

tëgg tulli sabaru to keep changing one's story, keep saying different things (intr.) {Tëgg na tulli sabaru. He kept saying different things.}

˜ *tëgg* 'to beat (a drum)', *tulli* 'to accompany a lead drummer', *sabaru* 'to dance the *saba*r with a drum accompaniment' :: 'to drum and dance the *sabar*' (a traditional dance)[9]

(16) *Chickasaw*

chipota hayoochi to get pregnant (often, accidentally) (I) § Chipota hayooshtokoot ittihaalalla'chi. She's going to get married because she got pregnant.

˜ *chipota* 'child', *hayoochi* 'to find' :: 'to find a child'

holisso kashof? to get divorced (I)

˜ *holisso* 'paper', *kashoffi* 'to clean' :: 'to clean the paper'

English and other languages have additional more complex types of idioms with open subject positions. In some, the object must be possessed by the subject, as in *to shoot one's wad, to hold one's horses,* or (with two possessed non-subjects) *to wear one's heart on one's sleeve.* Wolof and SLQZ examples are in (17) and (18):

(17) *Wolof*

ànd ag ay buumi nafaam to hold one's horses, not get carried away (intr.: poss. obj.) {Àndal ag say buumi nafa! Hold your horses!}

˜ *ànd* 'to go', *ag* 'with' *ay buum* 'ropes', *nafaam* 'his traditional purse or pouch':: 'to go with the strings of one's purse'

topp nafsoom to live one's life without direction, act without thinking (intr.: poss. obj.) {Bul topp sa nafsu. Don't act without thinking. (line from a Kiné Lam song)}

˜ *topp* 'to follow', *nafsu* 'nose' :: 'to follow one's nose'

(18) *SLQZ*

rchi'ih ru'ni' shuts up, gets quiet | *se calla* {subj. poss.; Bchi'ëng ru'ëng "He shut up | *Él se calló*" }

˜ *rchi'ih* 'fills in (a hole)', *ru'uh* 'mouth' :: 'to fill in one's mouth'

rgwèèe' dìi'zh nyèu' lohni' says one thing but means another | *dice una cosa pero quiere decir otra* {Rgwèèe' Gye'eihlly dìi'zh nyèu' lohni' "Mike says one thing but means another | *Miguel dice una cosa pero quiere decir otra*" }

˜ *rgwèèe'* 'says', dìi'zh 'word', *nyèu'* 'closed up', *lohoh* 'face, to' :: 'to say a closed up word to oneself'

Another type of verb phrase idiom is transitive, with a open object (or possessor of object) position as well as a open subject position, much like English *to give (someone) the shirt off one's back* or *to clean (someone's) clock,* or SLQZ (19):

[9] This expression may actually be some kind of verb-verb-verb compound.

(19) **ràa'nny gue'ehcy x:tàa'ah** keeps a secret for | *le guarda un secreto a* (prep. v.) {pssr. obj. id.; Ràa'nnya' gue'ehcy x:tàa'ng "I'm keeping the secret for her | *Le estoy guardando un secreto*" }
˜ *ràa'nny* 'sits down on', *gue'ehcy* 'head / on top of', *dàa'ah* 'petate (woven mat)' :: 'to sit down on top of (someone's) petate'
rtyu'uh x:bààa'n baptizes | *bautiza* {pssr. obj. id.; Rtyu' bxuuhahz x:bàa'nëmm "The priest baptizes him | *El sacerdote lo bautiza*" }
˜ *rtyu'uh* 'cuts off', x:*bààa'n* 'tail' :: 'to cut off (someone's) tail'

There are, however, other types of idiom which do not have open subject positions. One of these is the clausal idiom, which is fully specified lexically, but can occur in different constructions, varying for tense/aspect/modality and polarity. (Clausal idioms are thus different from proverbs, whose non-lexical features are fixed.) The best known English example (to linguists) is *for the shit to hit the fan* (where the citation with *for* and *to* specifies an idiom with a fully specified subject, verb, and other arguments, only whose non-lexical component is open); others are *for the shoe to be on the other foot, for the cat to be out of the bag,* and *for the fat lady to sing.* Other languages also have such idioms, for instance Chickasaw (20):

(20) **Siitanat imihoo fammi** for Satan to whip his wife: (idiom) for there to be rain and sun at the same time

Like the English examples, this expression can be used in different tenses and in various constructions.

And still other idioms have open non-subject positions, with an open object or other oblique position, comparable to English idioms like *for what to be eating ...?, for the vultures to be circling..., for fortune to smile on ...,* or *for a little bird to tell ....*[10] Mohave, for example, has metaphorical idioms that work like (21):

(21) **'anyach ka'aak=k** get sunstroke (idiomatic object expression: "the sun kicks" the affected person, the object of ka'aak) ['anya+ch] § 'Anyach nyaka'aakm. I got sunstroke ("the sun kicked me").

In the example sentence in (21), the person with sunstroke is a grammatical object (the speaker, 'me'), the syntactic subject is the word 'sun'.

Mohave also has many idioms in which the open position is the possessor of the subject. The most basic ways to say 'be happy' and 'be sad' in Mohave are

(22) **iiwanych 'ahot=k** be happy (idiomatic possessor expression) {pl. iiwanych 'ahuut} § 'Iiwanych 'ahuuttaahanm. We're very happy.
˜ *iiwa* 'heart', *'ahot=k* 'be good' :: > 'for ...'s heart to be good'
iiwanych 'alay=k be sad (idiomatic possessor expression) § Miiwanych 'alayk. You're sad.
˜ *iiwa* 'heart', *'alay=k* 'be bad' :: > 'for ...'s heart to be bad'

These examples also have parallels in English, such as *for ...'s tongue to be hanging out, for ...'s heart to bleed, for ....'s stock to be going up,* or *for ....'s*

[10] The existence of this type of English idiom has been the subject of some controversy (Marantz (1984: 27, 313)) but these structures are far from rare in other languages, as the examples illustrate.

chickens to come home to roost. Such idioms often involve body parts. As (22) shows, an idiom with a non-subject open position may often be translated with normal syntax (thus, the syntax of 'for ...'s heart to be good' does not match that of *to be happy*).

Finally, the open position may be the possessor of a non-subject, as in English *for the ball to be in ...'s court* or the following SLQZ idiom:

(23) **zh:ààa'n** 1. bottom, buttocks, rear end | *trasero, nalga* {pssd. only; zh:àa'nùu' "your bottom | *tu trasero*"} ...
zi'ihcydihzy càa gyi'biihahz zh:àa'nëng he's still inexperienced and ignorant, he's still wet behind the ears (he still has dry excrement in his bottom) | *él todavía es ignorante y sin experiencia, él todavía es un escuincle (todavía tiene excremento seco en el trasero)* {impers. pssr. id.}
~ *zi'ihcydihzy* 'just', *càa* 'is hanging', *gyi'biihahz* 'dry excrement', *zh:ààa'n* 'buttocks' :: 'for there to be dried excrement hanging in ...'s bottom'

As the examples above show, my dictionaries have varied considerably in terms of when and how I give "literal" translations for idioms (as in (20-21) and (23) above) and how I specify their grammatical structure. Colorful metaphors are intriguing to everyone, so including some of these paraphrases can enliven a dictionary. Sometimes these translations may have scholarly value, revealing the influence of other languages (I recently learned that the Chickasaw (20) may well be calqued from English, for example, and the second SLQZ idiom in (19) may come from Spanish (Lance Foster and Heriberto Avelino, p.c.'s)). However, I feel that including all such cutesy paraphrases has the effect of suggesting inappropriately that the target language of the dictionary conveys all such expressions through fully accessible metaphors, though surely the metaphors involved in the majority of such idioms are defunct. This is why I usually do not provide such paraphrases (as in (10), (12), (15-19) and (22)), including them only when speakers themselves point them. But such decisions clearly involve rather delicate judgment calls.

Making the grammatical structure of some idioms clear to the dictionary user can be a more challenging task. The only idiom above with an open non-subject whose structure is fully apparent from the dictionary entry presented above is (21). Although I discussed "idiomatic object expressions" like this in the introduction to the Mohave dictionary, there were so few of them that I could take the space to give a user-friendly presentation. Normally, however, because there is such a variety of different idiomatic construction types, I describe their structure carefully only in the introduction to the dictionary, where syntactic labels like those in the entries above are explained. There are pitfalls in this approach, of course, since many users (even linguists!) often skip an introduction. But when dealing with idiomatic syntactic structures of the sort I'll describe more fully in the next section, it does not seem feasible to have a long description of the structure repeated in each entry.

5. Idiomatic Syntax

The popular conception of idioms mentioned in the last section refers only to meanings that don't add up to the actual sense of the idiomatic phrase. However, a phrase may contain all the right semantic elements, but be syntactically idiomatic in that its grammar does not fit the standard (or most usual) patterns of the language. This is in fact one of the normal meanings of *idiom*: my American Heritage dictionary's first definition of this word is "a speech form or an expression in a given language that is peculiar to itself grammatically or cannot be understood from the individual meanings of its elements, as in *keep tabs on*" (1993: 674). In this section I will discuss systematic cases of deviation from expected syntax.

5.1

Let me start with some data from Chickasaw that has been discussed at length elsewhere (Munro & Gordon 1982,[11] Munro 1999). As we saw in (6), ordinary Chickasaw transitive sentences have nominative-accusative case marking on nouns. This is true even when the verb includes a dative applicative prefix (and would trigger III-dative agreement like that in (7-8) with a non-third person argument):

(24) Ihoo-at chipot-a̱ in-taloowa-tok.
woman-nom child-acc dat-sing-pt
'The woman sang for the child'

Some Chickasaw sentences containing two nouns, however, have more than one nominative, as in (25):[12]

(25) Chipot-aat ofi'-at im-illi-tok. 'The child's dog died,
child-nom dog-nom dat-die-pt The child had his dog die'

(25) is related by Possessor Raising to the more basic sentence pattern shown in (26). This construction can be used when the speaker wishes to highlight the discourse salience of the first nominative marked noun (the derived subject).

(26) Chipota im-ofi'-at illi-tok. 'The child's dog died'
child dat-dog-nom die-pt

(26) is an intransitive sentence whose subject is a possessed noun phrase (in Chickasaw, genitive nouns are unmarked, possessed nouns agree with non-third person possessors, and alienably possessed nouns have a dative prefix). Chickasaw has several types of multiple nominative sentences, all of which share two important properties. First, it is the "derived" subject ('child' in (25)) rather than the "old" or original subject ('dog') which passes a variety of syntactic subject tests. Second, the old subject is syntactically inert—unlike the derived subject, it cannot be freely moved or focussed, for example.

Chickasaw idioms like those in (27) are Possessor Raising structures with a specific noun—always marked nominative—filling the old subject role:

(27) **chipotaat imalla** to have a baby, give birth (III) § Chipotaat amalla. I had a baby.
˜ *chipota* 'child', *imalla* 'to arrive for' :: 'for ...'s child to arrive'[13]

---

[11] I do not have space to discuss a second multiple nominative construction, Oblique Subject, in which an underlying dative object is promoted to subject. Both constructions are lexically governed, as marked in the dictionary.

[12] As far as I know (25) has identical truth conditions to the simpler sentence in (26). The 'have' translation in (25) is used to reflect the subjecthood of the first noun in (25); note, however, that *have* does not indicate an indirect causative in this translation.

[13] Alternatively, this example might come from 'for a child to arrive to ...', indicating an Oblique Subject origin (cf. fn. 11). It is sometimes not easy to tell whether a given multiple nominative originated via Possessor Raising or Oblique Subject.

holissaat inkashofa to get a divorce (III) § Maryat holissaat inkashofatok. Mary got a divorce.
˜ *holisso* 'paper', *kashofa* 'to be clean' :: 'for ...'s paper to be clean'

Superficially, these expressions look similar to clausal idioms like (28) (comparable to (20), but with an intransitive verb), in that the entries in (27) and the one in (28) contain a nominative marked noun followed by a verb.

(28) **Chokfaat Achiili** for it to be Easter (when the rabbit lays eggs—a joke) § Chokfaat Achiilikma... on Easter (in the future)...
˜ *chokfi* 'rabbit', *achiili* 'to lay eggs' :: 'for the rabbit to lay eggs'

However, (28) is a complete clause, while the expressions in (27) have an open subject position, like many of the idioms considered in the last section. The syntax of the idioms in (27) is different from that of most of the verb entries in (15-19), since those in (27) do not include an object, but rather a nominative-marked old subject. The entries in (27) follow the rule of the Chickasaw dictionary that (unless otherwise indicated) subject inflection (here with the III/dative set, specified by "(III)") appears on the last word of the verb's entry.

5.2

Wolof and SLQZ have a much wider range of types of idiomatic verbs whose syntax does not follow standard patterns. Despite the enormous typological disparity between these two languages, their idiomatic structures are very similar. In this section I will survey several types of Wolof and SLQZ idioms with highly parallel syntax and discuss their treatment in dictionaries of these languages.

Wolof "object idioms" and SLQZ "impersonal idioms" are like Mohave (21), with an open object position. In the following examples from Wolof, the verb is not the first item in the dictionary entry, in contrast to the normal pattern. The ellipsis in the boldfaced entries signals the unusual word order (most inflectional clitics end up in that position, as the examples show).

(29) **cat...dugg** to suffer a downfall (after tempting fate through one's good fortune) (intr.: obj. id.) {Cat dugg na Yaasin. Yacine suffered as a result of her good fortune.}
˜ *cat* ('evil fate'), *dugg* 'to go into' :: 'for evil fate to go into...'
yaram...dab to put on a lot of weight (intr.: obj. id.) {Yaram dafa ma dab. I put on a lot of weight. | Ayda yaram dafa ko dab. Aïda put on a lot of weight.}
˜ *yaram* 'weight', *dab* 'to catch up with' :: 'for weight to catch up with ...'

The last example sentence in (29) could be translated 'Weight caught up with Aïda' or 'Aïda, weight caught up with her' (*Aïda* is topicalized, but not a syntactic subject): These idioms are identified in the dictionary as intransitive, since they "take" only one argument. They are called "object idioms" because the argument added to them is a syntactic object.

Here are some SLQZ examples:

(30) **rbih bxu'udy** is creased, is wrinkled, is pleated | *está doblado, está arrugado, es arrugado, está plisado, es plisado* {impers. id.; Rbih

bxu'udy x:casoonëng "His pants are pleated | *Sus pantalones son plisados*" }
˜ *rbih* 'sits', *bxu'udy* 'wrinkle' :: 'for pleats to sit on ...'
rnnàa'az x:lyiàa' gets a fever | *agarra fiebre (le da fiebre)* {impers. id.; Mnnàa'az x:lyiàa' la'anng "He got a fever | *Agarró fiebre, Le dio fiebre*" }
˜ *rnnàa'az* 'grabs', *x:lyiàa'* 'fever' :: 'for fever to grab ...'

The added argument appears at the end of the sentence here, in the SLQZ object position. (Thus, for instance, a post-verbal pronoun like *la'anng* 'him' (in the last example) can never be used to indicate a syntactic subject.) These idioms are called "impersonal" because they have a filled subject position and their most discourse-salient argument is a non-subject. As the examples show, there is a very English-centric slant to this whole analysis, which considers these expressions as non-standard because they do not have the same subject as the English translation.

Sometimes the SLQZ grammatical relations show a parallelism with Spanish. For instance, consider the parenthesized Spanish translation of the last SLQZ entry in (30). The natural way to say 'he gets a fever' in Spanish uses the verb 'give' with an indirect object 'him' (a non-subject), just as in the SLQZ expression. Other SLQZ "impersonal" expressions include Spanish loans and mimic Spanish syntax:

(31) **rdèèi'dy gaan** feels like (doing something relatively important) | *tiene ganas de (hacer algo relativamente importante)* {impers. id.; followed by irr.; Que'ity rdèi'dydi' gaan la'anng ycwàa'ng gyèe'ts lohrëb "He doesn't feel like writing a letter to them | *Él no tiene ganas de escribirles a ellos*" }
˜ *rdèèi'dy*: 'gives', *gaan* < Spanish *ganas* 'wishes'[14] :: 'for "wishes" to give ...' (cf. Spanish *le dan ganas*)
ruhny gwu'ast likes | *disfruta (le gusta)* {impers. id.; Ruhny gwu'ast nàa' biien "I like wine | *Disfruto el vino, Me gusta el vino*"; Ruhny gwu'ast lìu' biien "You like wine | *Disfrutas el vino, Te gusta el vino*" }
˜ *ruhny* 'does', *gwu'ast* < Spanish *gusta* 'pleases' :: 'for "pleases" to do ... (something)' (cf. Spanish *gustar*)

Although *ruhny gwu'ast* is influenced by Spanish *gustar* (*gwu'ast* is derived from the inflected form *gusta*) and more similar to it than to English *like,* there are differences between the SLQZ and Spanish impersonal structures: in SLQZ (32), the liked item is a second object (not a subject, as in Spanish); the liker is a first object (not an indirect object, as in Spanish); and *gwu'ast* is in subject position:

(32)	R-uhny	gwu'ast	naa'	biien.	'I like wine'
	hab-do	liking	pron.1s	wine	

Wolof "possessive idioms" and SLQZ "possessor idioms" are like Mohave (22), with an open possessor of the grammatical subject position. Here are some Wolof examples (again, the ellipsis in the entry words shows that the verb follows):

(33) **gémminñam...buur** to have all one's teeth (intr.: poss. id.)
˜ *gémmiñ* 'mouth', *buur* 'to be full' :: 'for ...'s mouth to be full'
làmmiñam...tar to have a sharp tongue; to always have an answer; to always have something to say (intr.: poss. id.) {Làmmiñam dafa tar. She

[14] *Gaan* (31), *gwu'ast* (31), and *nesesitaar* (43) are not used alone with the meanings indicated.

has a sharp tongue, She always has an answer, She always has something to say.}
˜ *làmmiñ* 'tongue', *tar* ('to be sharp'), 'to be oppressive' (Fal et al. 1990) :: 'for ...'s tongue to be sharp'

Here are some SLQZ examples:

(34) **dehts** back...
rcah dehtsëng he has to buy drinks for everyone (because of losing a bet or forgetting a commitment, for example) | *él les tiene que disparar las bebidas a todos (por haber perdido una apuesta o por olvidar un compromiso, por ejemplo)* {pssr. id.; Que'itydi' dehts Gye'eihlly ycah "Mike won't have to buy drinks | *Miguel no tendrá que disparar las bebidas*" }
˜ *rcah* 'hangs (on the tree, of fruit)' :: 'for ...'s back to hang'
gue'ehcy head...
bèi'ny cwe'eenn làa'iny guehs nu'uh gue'ehcyëng he doesn't understand, he's in the dark, it's as if he has his head in a pot | *él no entiende, es como si él tuviera la cabeza en una olla*{pssr. id.; Bèi'ny cwe'eenn làa'iny guehs nuuh' gue'ecya' "I don't understand | *No entiendo*" }
˜ *bèi'ny cwe'eenn* 'like', *làa'iny* 'in', *guehs* 'pot', *nu'uh* 'is inside' :: 'for it to be like ...'s head is inside a pot'

In these SLQZ expressions, the only syntactic reflex of the subject of the English translation is as the possessor of the grammatical subject of the expression, which appears at the very end of the entry sentence, as in this example from the second entry in (35):

(35)	B-èi'ny	cwe'eenn	làa'iny	guehs	n-u'uh	gue'ehcy=ëng.
	perf-do	account	in	pot	neut-be.in	head=3s.prox
	'It's like his head is in a pot'					

In the SLQZ dictionary, such "possessor idioms" are given as indented sub-entries under the main entry for these inflected words (whose definitions are ellipted in (34) and (37) below), the possessed subjects (with copious cross-references and a clear indication in the index of where the full entry is to be found). We felt that it wouldn't make sense to leave this possessed subject (especially a body part word such as the headwords in (34)) hanging uninflected at the end of such an expression, so these indented idioms are listed with arbitrary third person singular proximate pronominal inflection, forming a complete sentence (as is shown by the definitions in (19)).[15]

[15] There is an asymmetry between the English and Spanish translations of SLQZ verbs like those in (34). Thus, the bare third person verbs in English match the usage of the bare SLQZ verbs, since each requires a third person subject to form a complete sentence; however, a bare third person verb constitutes a complete sentence in Spanish (a pro-drop language). We include the subject pronoun *él* in Spanish translations like those of the first and second idioms in (34) to help suggest that these actually include an overt subject (the proximate clitic *=ëng*).

6. Dictionary Corpuses, Idioms, and Comparative Syntax

Where do the idioms in my dictionaries come from? You cannot elicit for idioms. Only rarely would a speaker be able to come up with answers to a question like, "Tell me some expressions that don't mean what you would think they would mean from the words they contain." Asking a speaker for other expressions that contain a particular verb or noun might produce some results, but it's a pretty hit-or-miss technique. The place for a fieldworker to find idioms—in the broad sense I use that term, of words and phrases whose meaning and/or syntax must be listed in the dictionary—is natural speech, hopefully supplemented by transcribed and analyzed texts. (Haas knew what she was talking about—both grammar and texts are necessary to ensure a good dictionary.) But a comprehensive dictionary can itself constitute a corpus of data about idioms. While preparing this paper, I found examples that help support some general observations concerning the syntax of idioms, and I discovered new evidence showing how idiomatic syntax like that I described in the preceding section can be restructured more conventionally.

6.1

Nunberg, Sag, & Wasow (1994) argue, for example, that the existence of families of related idioms is evidence for treating idioms componentially. Thus, in the two pairs of Wolof possessive idioms in (36), either the subject or the verb can be replaced by another with a similar meaning, indicating that speakers must be aware of the meaning of the idioms' component parts:

(36a) **coonoom...jéëx** to feel relieved (intr.: poss. id.)
coonoom...wàcc to be soothed, relieved; to be less tired (intr: poss. id.)
˜ *coono* 'tiredness', *jéëx* 'to be all gone' /*wàcc* 'to go down' :: 'for ...'s tiredness to be all gone / go down'

(36b) **fitam...tëf-tëfi** to have heart palpitations; to have heartburn (intr.: poss. id.)
xolam...tëf-tëfi to have heart palpitations; to have heartburn (intr.: poss. id.)
˜ *xol* 'heart' / *fit* 'courage' *tëf-tëfi* 'to jump up and down' :: 'for ...'s heart / courage to jump up and down'

Similarly, (37) is a variant of the second SLQZ idiom in (34) illustrating creative elaboration based on speakers' clear understanding of the meaning of the idiom's parts. As the note after entry (37) shows, this is an ongoing process (but one that could not happen if the idiom were treated by speakers as an unanalyzable unit):

(37) **gue'ehcy** head...
bèi'ny cwe'eenn te'ihby gami'izh gyìia' nu'uh gue'ehcyëng he doesn't understand, it's as if he has his head in a girl's blouse | *él no entiende, es como si él tuviera la cabeza en la blusa de una joven* {pssr. id. used by men}
§§ When a young man says, "Bèi'ny cwe'eenn te'ihby gami'izh gyìia' nu'uh gue'ecya'" ("I don't understand | *No entiendo*"), a friend may joke, "Bèi'ny cwe'eenn làa'iny casoon tyu'c me'eu a'ti'" ("No, it's like (it's) in dirty underpants | *No, es como (si tuvieras la cabeza) en calzones sucios*").

6.2
An extensive, analyzed dictionary corpus may reveal sets of related entries one of which has idiomatic syntax (as in section 5) and one of which does not.

Consider first the Mohave idiomatic expressions in (22) above. These idioms have an open possessor of subject position: to say 'I am happy', for instance, one says 'My heart is good', and so on. Mohave verbs with first person subjects have a glottal stop prefix, but no glottal stop is added to the verb in the example in the first entry (instead, the first person glottal stop appears on the possessed subject noun); the Mohave second person prefix *m-* appears on the possessed subject rather than the verb in the second entry example: thus, what we think of as the subject is not the grammatical subject of the expression. Is this simply an English-centric definition? Perhaps not, given that Mohave also has expressions like those in (38) and (39):

(38) **iiwany 'ahot=k** be happy § 'Iiwany 'a'ahotk. I am happy.
iiwany 'alay=k be sad, feel bad § 'Iiwany 'a'alayk. I'm sad.

(39) **wa 'ahot=k, wa 'ahoot=k** be happy, nice, generous § Wa ma'ahootk. You're nice.
wa 'alay=k be sad § Wa ka'alaymotm. Don't be sad.

In these verbs, the "logical" subject is treated as the grammatical subject of the verb of the expression (as shown by the appearance of the first person glottal stop prefix on the last words of the examples in (38) and the second person *m-* and imperative *k-* prefixes on the last words of the examples in (39)). The verbs in (38) are possessor raising structures: in these expressions, the possessor of the original subject 'heart' is the subject of the verb (though 'heart' is still marked for the same possessor). The verbs in (39) are even farther from the original structure, with *wa* 'heart' stripped of its first syllable and no longer marked for possessor.

The pattern of development seen in (22) > (38) > (39) (whose directionality seems clear, given the original semantics; cf. Munro (1976)) occurs in other languages as well. Thus, alongside the second Wolof entry in (33) we find (40):

(40) **tar làmmiñ** to have a sharp tongue; to always have an answer; to always have something to say (intr.) {Ayda dafa tar làmmiñ. Aïda has a sharp tongue.}

The first verb has the sharp-tongued person only as the possessor of the subject 'tongue'. In the second, however, the sharp-tongued person is the subject of the verb.

A comparable change affects Chickasaw multiple nominative constructions like those in (25). Speakers sometimes restructure such sentences to resemble ordinary dative object constructions like (24), as with (41) (cf. Munro (1999):

(41)	Zak-at	ofi'	im-illi-tok.	'Zak's dog died'
	Zak-nom	dog	dat-die-pt	

(41) is an alternative to the Possessor Raising structure (25) (itself derived from (26)). Every Chickasaw speaker I have surveyed uses this novel construction with some verbs.

The change from Chickasaw (26) to (25) to (41) is similar to the Mohave and Wolof examples since each has an apparent subject + verb + object pattern in which the "object" is the original subject of the verb and the syntactic subject originally played a different semantic role (possessor), as schematized in (42):

(42) *Chickasaw* (possessor + possession subject) + intransitive verb (26)
> derived possessor subject + old possession subject + verb (25)
> possessor subject + unpossessed possession + verb (41)
Mohave (possessor + possession subject) + intransitive verb (22)
> derived possessor subject + (possessor + possession) + verb (38)
> possessor subject + (unpossessed possession-clitic + verb) (39)
Wolof (possessor + possession subject) + intransitive verb (33)
> possessor subject + verb + unpossessed possession (40)

The best syntactic analysis of derived constructions like those in the examples in (39), (40), and (41) is far from clear. What is certain is that speakers restructure the original constructions in order to show that the possessor is more salient in the discourse than the possessed item. Usually the possessor is higher on an animacy scale than the possession and, since the possession is very often a part of the possessor's body, the change can be seen as a type of metonymy.

6.3
Many of these changes also remove a conflict between the target language structure and the English translation: thus, 'I am happy' has a subject 'I', and so do the derived Mohave structures (38) and (39); 'Aïda has a sharp tongue' has a subject 'Aïda', and so does the example in Wolof (40). The translation of Chickasaw Possessor Raising structures like (25) or the new structure in (41) often does not change from the original possessed subject structure, but 'have' translations sometimes work with these (e.g. 'Zak had his dog die'), and other such derived Chickasaw structures can express concepts like 'I got a divorce' (27, from 'my paper is clean').

The principal case where I earlier worried about English-centric definitions concerned the SLQZ constructions exemplified in (31), in which there was a clear mismatch between "impersonal" (quite likely Spanish influenced) syntax and English translations whose subjects corresponded to SLQZ non-subjects. But speakers can also express the semantic non-subjects of such constructions as syntactic subjects. (43) presents two SLQZ verbs that can be used in both "impersonal" and more English-like constructions (contrast the first and second definitions in each entry):

(43) **ràa'izynah dìi'zh** 1. gets hurt by words, gets offended by words | *se ofende, se siente (por lo que se dice)* {impers. id.; Ràa'izynah dìi'zh Gye'eihlly "Mike is hurt (by words) | *Miguel se siente (por lo que se dice)*"; Que'ity dìi'zh chàa̧'izynah la'anng "He won't be hurt (by words), Words won't hurt him | *Él no se sentirá (por lo que se dice), Las palabras (lo que se dice) no lo lastimarán*"}; 2. gets hurt by words, gets offended by words | *se ofende, se siente (por lo que se dice)* {Que'ityëng chàa'izynàa'ng dìi'zh "He won't be hurt (by words) | *Él no se sentirá (por lo que se dice)*"} > ràa'izy "hits | *golpea*"
˜ *ràa'izynah* 'hits hard', *dìi'zh* 'word' :: (1) 'for the word to hit .... hard'; (2) 'for ... to hit the word hard'
ruhny nesesitaar 1. needs | *necesita* {impers. id.; Ruhny nesesitaar nàa' mùuully "I need money | *Necesito dinero*"}; 2. needs | *necesita* {Runya'

nesesitaar mùuully "I need money | *Necesito dinero*"; te'ihby bùunny nih ruhny nesesitaarih "a person who needs it | *una persona que lo necesita*" }

˜ *ruhny* 'does', *nesesitar* 'to need' (infinitive) < Sp. *necesitar* :: (1) 'for "to need" to do ...'; (2) 'for ... to do "to need"'

Consider the examples in (44):

(44a)	R-uhny	nesesitaar	nàa'	mùuully.	'I need money'
	hab-do	to.need	1s.pron	money	
(44b)	R-uny=a'	nesesitaar	mùuully.	'I need money'	
	hab-do=1s	to.need	money		

In (44a), the syntactic subject appears to be the borrowed infinitive *nesesitaar.* 'I' is expressed here with a pronoun that can only be interpreted as a non-subject. In (44b), however, 'I' is the syntactic subject, indicated with a subject clitic on the verb. Using borrowed Spanish infinitives after *ruhny* 'does' is a productive process in SLQZ, but usually such expressions work like ordinary transitives, with the "logical" subject as their syntactic subject, as in (44b). (44a) is unusual in that the "logical" subject appears in object position and "impersonal" non-subject form.

7. Conclusion

The idiom restructurings I have surveyed in the preceding section seem to reflect a trend toward unity of syntax and semantics, with the actual subject of the discourse becoming the syntactic subject as well.

Such changes, however, may result in complications in the structure of the verb phrase. If MY TONGUE SHARP changes to something that looks as if it ought to mean I SHARP TONGUE, what does this mean? Does SHARP actually become a transitive verb, with TONGUE as its object? Perhaps SHARP and TONGUE are somehow compounded or incorporated together, forming a more complex but still intransitive verb with the syntactic role of the original subject even further diminished. I suggested in Munro (1976) that Mohave verbs like those in (38) and especially (39) might be best analyzed as involving incorporation, and such an idea has also been suggested for the Chickasaw and SLQZ constructions just surveyed.

However, an incorporation approach does not always work. Consider (45), which shows different uses of the Chickasaw verb *kaniya*, which means 'get lost', 'go away', or even (in its suppletive plural form *tamowa*) 'elope'. These sentences illustrate Possessor Raising, as in (26)-(25), with a semantic change when the Possessor Raising construction is used in (45b) to express the transitive verb 'lose'.

(45a)	Keeli	i̱-holiss-aat	kaniya-tok.	'Keeli's book got lost'
	Keeli	dat-book-nom	get.lost.s-pt	

(45b)	Keeli-at	holiss-aat	in-kaniya-tok.	'Keeli lost her book'
	Keeli-nom	dog-nom	dat-get.lost.s-pt	

The innovative structure (46) shows more changes. The lost item ('book'), which, like the "old subject" in (41), is not only not marked nominative, but is more individuated (marked for possessor) and, significantly, actually marked accusative. Further, a second object noun in the sentence (licensed by the locative applicative prefix *aa-*) intervenes between this erstwhile old subject 'book' and the verb.

(46) Keeli-at i̲-holisso-a̲ holissaapisa' aa-in-kaniya-tok.
Keeli-nom dat-book-acc school loc-dat-get.lost.s-pt
'Keeli lost her book at school'

The word order and the accusative marking argue against any assumption of incorporation in this case, suggesting rather that the verb *(in)kaniya* has actually acquired a new meaning in this construction, 'lose', whose grammatical relations are more like those we conceptualize in English.

Like many of the other idioms surveyed in this paper, the new uses of *kaniya* shown in (45b) and (46) represent dead metaphors, idioms whose "literal" meaning is certainly not accessed by speakers every time they use the idiomatic structure, although that meaning may be fully or partially accessible with conscious thought, and although certain parts of that meaning continue to be felt. (Thus, in its original intransitive use *kaniya* refers only to a single entity getting lost or going away, a feature retained even in the 'lose' sentences (45b) and (46)—although Chickasaw does not distinguish nouns like 'dog' and 'book' for number, such sentences can only be used to talk about losing a single dog or book.)

Noticing idioms in the languages you do fieldwork on is important, and putting them in the dictionary can be an interesting challenge. In this paper, I have very briefly surveyed a variety of idiomatic verbal expressions from four very different languages, and described some of the decisions I made in deciding how to put these into dictionaries. Dictionaries are often unappreciated as contributions to linguistic analysis, so I have also tried to illustrate some of the ways in which they can be more than simply lists of words. Diverse languages may have quite similar types of idiomatic verbs which pose problems for the dictionary maker, but which may contribute to a general understanding of how idioms work and are analyzed. Finally, the restructuring of idiomatic, nonstandard syntax often produces a striking parallelism with English, a language in which grammatical subjecthood is far more strongly correlated with discourse salience than is true in many other languages.

References

Albright, Adam. In preparation. The Identification of Bases in Morphological Paradigms. PhD dissertation, UCLA.

American Heritage College Dictionary. 1993. Third Edition. Boston/New York: Houghton Mifflin.

Fal, Arame, Rosine Santos, & Jean Léonce Doneux. 1990. *Dictionnaire wolof-français.* Paris: Éditions Karthala.

Gordon, Lynn. 1986. *Maricopa Morphology and Syntax. University of California Publications in Linguistics* 108. Berkeley, California: University of California Press.

Gordon, Lynn, Pollyanna Heath, & Pamela Munro. In preparation. *A Maricopa Dictionary.*

Manaster-Ramer, Alexis. 28 Jan. 1993. "Subject-Object Asymmetry and Idioms: A Query." http://linguistlist.org/issues/4/4-43.html#4

Marantz, Alec P. 1984. *On the Nature of Grammatical Relations.* Cambridge and London: MIT Press.

Martin, Jack B., & Margaret McKane Mauldin. 2000. *A Dictionary of Creek / Muskogee, with notes on the Florida and Oklahoma Seminole dialects of Creek.* Lincoln / London: University of Nebraska Press.

Munro, Pamela. 1976. Subject Copying, Predicate Raising, and Auxiliarization: The Mojave Evidence. *IJAL* 42: 99-113.
Munro, Pamela. 1996. Making a Zapotec Dictionary. *Dictionaries* **17**: 131-55.
Munro, Pamela. 1999. Chickasaw Subjecthood. In Payne, Doris L. & Immanuel Barshi (eds.) *External Possession,* 251-89. Amsterdam/Philadelphia: John Benjamins.
Munro, Pamela. To appear. Dictionary Entries for Verbs. In Frawley, William, Kenneth C. Hill, & Pamela Munro (eds.) *Making Dictionaries.* UC Press.
Munro, Pamela, Nellie Brown, & Judith G. Crawford. 1992. *A Mojave Dictionary.* UCLA Occasional Papers in Linguistics 10.
Munro, Pamela & Dieynaba Gaye, 1997. *Ay Baati Wolof: A Wolof Dictionary (Revised Edition).* UCLA Occasional Papers in Linguistics 19.
Munro, Pamela & Lynn Gordon. 1982. Syntactic Relations in Western Muskogean: A Typological Perspective. *Language* **58**: 81-115.
Munro, Pamela & Felipe H. Lopez, with Olivia V. Méndez, Rodrigo Garcia, & Michael R. Galant. 1999. *Di'csyonaary X:tèe'n Dìi'zh Sah Sann Lu'uc (Dictionary of San Lucas Quiaviní Zapotec / Diccionario Zapoteco de San Lucas Quiaviní)* (two volumes). Los Angeles: Chicano Studies Research Center Publications, UCLA.
Munro, Pamela & Catherine Willmond. 1994. *Chickasaw: An Analytical Dictionary.* Norman/London: University of Oklahoma Press.
Nunberg, Geoffrey, Ivan A. Sag, & Thomas Wasow. 1994. Idioms. *Language* **70**: 491-538.
Payne, Doris L. 1982. Chickasaw Agreement Morphology: A Functional Explanation. In Hopper, P. & S. A. Thompson (eds.) *Studies in Transitivity* (*Syntax and Semantics* vol. 15), 351-378. New York: Academic Press.
Sylestine, Cora, Heather Hardy, & Timothy Montler. 1993. *A Dictionary of the Alabama Language.* Austin: University of Texas Press.

Department of Linguistics, UCLA
UCLA Box 951543
Los Angeles, CA 90095-1543

munro@ucla.edu

Interrogatives in the field: Imaginary questions, real answers, and the law[1]

ESTHER PASCUAL
University of California, San Diego

0. Introduction

On an otherwise ordinary day, we may listen to a contemporary philosopher *question* Immanuel Kant's ideas and *respond* to them. We may see an ad with a picture of a missing pet that *asks* us *"Have you seen me?"*. Or we may lose our keys and ask ourselves *"Where did I just put them?"*. In actuality, a modern-day philosopher cannot bring Kant to the modern era to engage in simultaneous debate with him. A missing pet cannot talk to those who are looking for him. And nobody can be split in two parts and have the two selves interact with each other. However, cognitive linguists, and more specifically mental space and conceptual blending researchers, have shown that such apparent impossibilities do occur and are even quite common at the conceptual level (Fauconnier & Turner 1996, 2002; Lakoff 1996). This paper deals with the use of *imaginary* or *fictive* questions of the sort exemplified above as argumentative devices that can structure on-line discourse and reasoning in legal settings. The theoretical perspective used is the theory of *conceptual integration networks* or *conceptual blending* (Fauconnier & Turner 1996, 1998, 2002).

1. Conceptual integration networks

Conceptual blending is a theoretical framework of on-line dynamic construction of meaning. The basic cognitive operation involved in *blending* is the combination of two or more *input spaces* to produce another space, the blend. That space inherits partial structure from the input spaces and has emergent

[1] Field research and data analysis were supported by a predoctoral grant from the Fulbright Foundation. I have benefit from general comments and brainstorming from Aaron V. Cicourel, Gilles Fauconnier, Ed Hutchins, Theo Janssen and Frederike van der Leek during this research. I am also thankful to all the informants in the project, especially to the district attorney in this trial, who offered insightful introspective views on his own discourse. I am grateful to Frederike van der Leek and Line Brandt for reading and commenting on earlier drafts of this paper. Whatever errors the paper contains are of course my own.

structure of its own. Conceptual integration networks are constituted by *mental spaces* (Fauconnier 1994). These are mental constructs of potential realities, which are being set up as discourse unfolds. Spaces are structured by *elements*, which represent conceptual entities, and they are enriched by *cultural models* and *frames*. Frames are structures of role-value pairs like the family, a debate or a conversation.[2] The integration of a frame as an input space with some roles that are mapped onto certain kinds of elements as values in another input space constitutes a so-called *simplex network* (Fauconnier and Turner 2002). A theory of mental spaces and conceptual blending is extremely useful for an account of legal argumentation, since, as Tannen (1998) points out, the very kernel of the adversary system is the different meaning constructions of the same objective reality by the counsels of each side (see also Goodwin 1994).

2. Analysis

The data discussed below comes from an ethnographic study of a lawyer's closing argument to the jury in a high-profile murder trial which took place in a California court in the fall of 2000.[3] The defendant was accused of brutally killing his wife with a fire poker in the couple's home. Two examples will be discussed, in which the lawyer produces a question that sets up an imagined communicative interaction. These are: a) the response to a rhetorical question that the counsel of the opposite side was supposed to have uttered; and b) the predicative use of an interrogative sentence type as a definition.

2.1. Rhetorical questions, asked and answered

Rhetorical questions challenge us with an interesting paradox. They are clearly not information-seeking, and yet they are produced as interrogatives, which are conventionally associated with question asking. This paradox becomes particularly intriguing if one considers the frequent use of rhetorical questions in communicative contexts like litigation, in which the compulsive force of the argument does not leave much room for unnecessary embellishment. Indeed, rhetorical questions are extremely common argumentative devices in the courtroom (cf. Ilie 1994), and they are actually recommended by legal tacticians to both lawyers and witnesses alike (cf. Casanovas 1998; O'Barr 1982;

---

[2] See Coulson (2001) for a detailed explanation of frames, and the operation she calls *frame-shifting*.

[3] The data was collected within the project "An ethnographic approach to discourse processes in court" (# 020033S), reviewed and approved by the UCSD Human Subjects office. Ethnography included: i) participant observation, ii) official transcripts, iii) written and visual media coverage, v) official documents, and vi) interviews with the main legal professionals and lay participants. Written informed consent was obtained from all informants, whose participation in the study was entirely voluntary. In order to protect the confidentiality of both the informants and the persons referred to in the transcript, all names have been changed.

I.C.D.A.'10[4]). I assume the following characterization of rhetorical questions (Pascual 2001, to appear): i) they constitute a *pragmatic* rather than a *grammatical* category (Ilie 1994); ii) they are to be heard as *questions* but understood as *statements* (Ilie 1994); iii) they suggest a *clash* between two presupposed scenarios; and iv) they set up an *imaginary trialogue* in which a communicator *a* attempts to convince *b* that *c*'s argument is incorrect. In this section I explore the conceptual blends behind a *fictive* rhetorical question asked and answered by a lawyer. The example to be discussed, extracted from the prosecutor's closing argument, is the following (my italics):

(1) Now, Mr. Loeber [defense counsel] questions, *"Well, how could the blood get on the end of the poker, because the poker is not hitting her in the head?"* [...] The reason why blood gets on the end of the poker [...] is centrifugal force.

In this piece of discourse the lawyer appears to respond to an actual question previously raised by his adversary. If one looks at the argument referred to, however, it turns out that such a question was actually never produced. Instead, what had been said was (my italics):

(2) And we know from Dr. Stone's [spatter expert] testimony and from our own common sense, when we look at these unfortunate, sad photographs of Rachel [victim] from the coroner, that *there were no wounds there that correspond to the end of a fire poker.* They're linear wounds. That's why we have linear, linear, linear. But to get that castoff spatter *we have to have blood on the end of the poker*, and that would get there most likely -- we've had *no other explanation* -- by the end of the poker hitting Rachel's head.

In (1) the utterer presents the whole passage above condensed in a single sentence, a rhetorical question that he then proceeds to answer. When interviewed after the trial and asked to comment on that passage, the prosecutor explained that his adversary "was saying that, you know, if it was a poker, *why aren't there poker marks in her head?*". He then concluded: "so, I was just *responding*". Note that not only did the defense counsel *not* produce any of those questions, but that the argument to which the prosecutor is *responding* occurred a good forty-five minutes prior to his response. Moreover, they both appear embedded in long monologues addressed to the jury. In order for the question-answer operation in (1) to occur at all, the whole interactional sequence, that is the discourse by the prosecutor and the defense counsel respectively, needs to be merged with a *debate*

[4] The 10th annual I.C.D.A. Trial Skills Academy, held at the California Western School of Law. April 21st-28th, 2001. Ethnographic notes.

frame into a *simplex network* (Fauconnier & Turner 2002). One input contains the debate frame with turntaking and participant roles and no values, and the other input contains unframed elements, namely the actual arguments produced and the two lawyers. The inputs are then matched by a Frame-to-value connection.

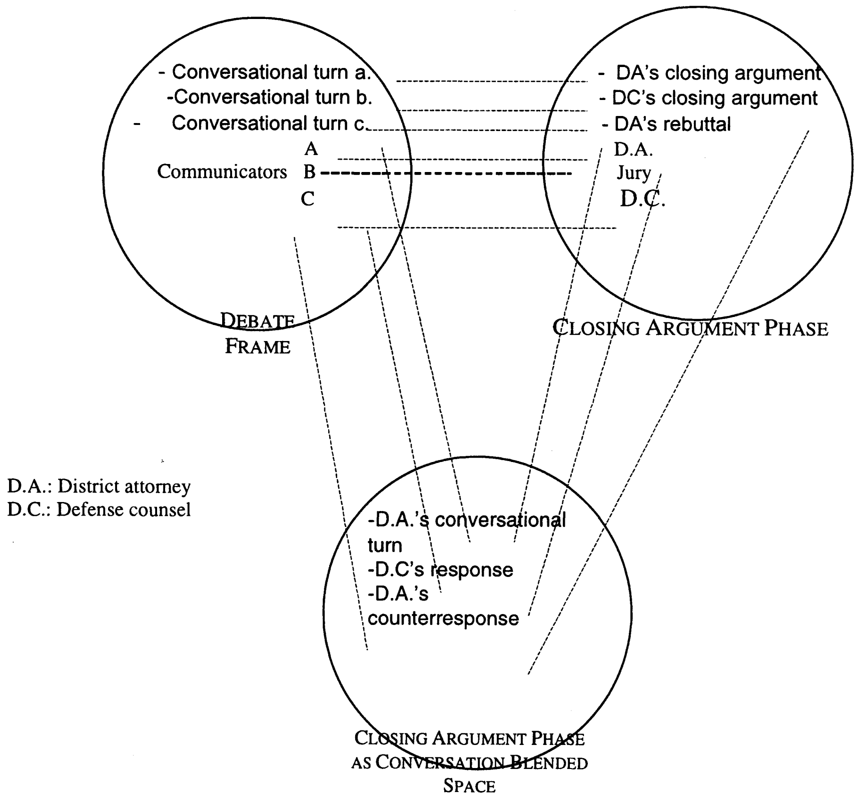


Fig. 1. Closing arguments phase as debate blend.

Each communicative phase in a trial (e.g. closing argument, direct- and cross-examination) seems to be regarded by participants as a single turn in an imagined debate with the opposite side. Thus, each communicative *turn* is to be *responded* to by the adversary's subsequent turn, which may then be further countered by yet another turn. The "closing argument~debate" network seems to be a conventional blend in the courtroom. In the American legal system this is overtly manifested in different interactional phases: "direct" vs. "*re*direct" and "cross" vs. "*re*-cross" examination; and "closing argument" vs. "closing argument *rebuttal*". Significantly, in the redirect, re-cross and rebuttal phases, one can only discuss issues that were brought up in the prior direct-, cross- and closing argument respectively. This feature is directly projected from the debate frame. It should be noted that the conceptualization of subsequent monologues as a simultaneous interaction is not restricted to the legal setting. Communicative participants often seem to conceptualize subsequent communicative performances, both written and oral, as different turns in a larger conversation-like structure. A well-known example from the conceptual blending literature which seems to undergo that process is Fauconnier & Turner's Debate-With-Kant (1996, 2002), mentioned in the introduction. A discussion between two philosophers of different centuries presents no problems at the conceptual level, couched as it is in terms of the turn-taking pattern of the ordinary face-to-face conversation.

Once the overall closing argument phase of the trial is condensed into a sequence of turns in a debate, the projection of a question-answer pattern as in (1) becomes almost self-explanatory. What seems less obvious is the fact that the rhetorical question in (1) should be immediately followed by a response.[5] However, under a non-derivational paradigm like the one adopted here, the very possibility of responding to a rhetorical question – even a fictive one – is easily explained by the hypothesis that rhetorical questions *are* questions, and that they are question *uses* rather than *kinds* (Ilie 1994). Notice that no linguistic means indicates that the question in (1) is to be understood as rhetorical. At the same time, the power of responding to a rhetorical question seems to derive from the fact that these are still to be understood as *statements*, albeit heard as questions.[6]

---

[5] The production of a rhetorical question only to be subsequently answered by the same utterer seems to contradict the definition of rhetorical question as those that "do not expect an answer," given by some some scholars and the 1988 edition of the Longman Dictionary of the English language rhetorical (see Ilie 1994: 42).

[6] A similar example of a response to a rhetorical question raised by the adversarial counsel in a prior closing argument comes from a criminal trial in a Spanish court. The defense was arguing that since the defendant and the victim were suffering from deep financial problems, the defendant had aided the victim in committing suicide. The private prosecutor confronted this version of the facts with a rhetorical question: *"Quién no sufre por problemas económicos?"* ["Who doesn't have financial problems?"]. When his turn came, the defense counsel quoted this question – as rhetorical - and responded to it, in a way that was consistent with his explanation of the facts: *"Quién no sufre por problemas económicos?, nos dice la acusación particular. Yo se lo diré. No sufre por problemas económicos una niña de diecisiete años. No tiene que sufrir por problemas económicos"* [who doesn't have financial problems? Says the private prosecutor. I will tell you who. A seventeen-year old kid does not have financial problems. She shouldn't suffer from financial problems.] Data from GRES (UAB Sociolegal Studies Group). 1997. El Informe del Abogado de la Defensa ante el Jurado. Juicio por Asesinato. Video code: T25.jj.4.

Let us now zoom in and focus on the particular elements involved in the blend prompted in (1). The prosecutor's memory of the argument he attempts to attack, which runs through many sentences, is compressed into one. This *compression* must be allowed by some commonality between the two elements that get mapped and blended, namely the counsel's argument in input one and a rhetorical question frame in input two. First, if rhetorical questions are truly pragmatic categories and to be "heard as questions but understood as statements", then it follows that for the argument in (2) to get compressed into a rhetorical question it must be readable as both a question and a statement. Second, if the argument of the defense counsel is to be fully integrated with a rhetorical question frame, then it should also "suggest a *clash* between two presupposed scenarios". Lastly, the argument in (2) should prompt an "imaginary trialogue in which a communicator *a* attempts to convince *b* that *c*'s argument is incorrect". The basic categorization of rhetorical questions does seem to match with the skeletal structure of the argument in (2). The counsel is making two explicit *statements*: a) "there were no wounds there that correspond to the end of a fire poker", and b) "we have to have blood on the end of the poker". By so doing, the counsel is pointing out the *contrast* between two mental spaces: a) a reality space in which the victim's wounds do not seem to correspond to the end of a fire poker, and b) another reality space in which blood spatter suggests that there must have been blood at the end of the poker. Both reality spaces are structured by frames and cultural models ("our own common sense"). The latter scenario is structured by a cultural model of force-dynamics. In the blend, there is an inner space cause-effect mapping linking the end of the poker hitting the victim's head with blood at the end of the poker (and subsequently blood spatter found by the investigators and experts). By presenting these two contradictory scenarios, the defense counsel challenges the prosecution's theory of the case. According to the prosecution, the fire poker that is missing from the couple's home corresponds to the murder weapon that was never found, a claim the defense wishes to cast reasonable doubt upon. If the blood spatter shows that there must have been blood at the end of the poker, and the victim was not hit with the end of a poker, then the question surely arises: *how could the weapon have been a poker?*

Lastly, the *space-builder* that introduces the question in (3), the verb "to question" rather than "to ask" indicates a rhetorical rather than an answer-seeking use of the question. This verb choice frames the defense counsel's verbal space within a confrontation rather than an information-seeking frame, which is consistent with the confrontational feature of rhetorical questions assumed here. Notice that in the adversary system, the role of the defense is not to prove the defendant's innocence, but rather to convince the jury that there is reasonable doubt as to guilt. At the same time, the prosecution will argue the defendant's guilt before the jury, beyond the possible doubt of which the defense may have convinced the jury. The basic communicative structure of confrontational argumentation, and then surely of the courtroom, fits perfectly with the skeletal structure that characterizes rhetorical question use: "communicator *a* attempts to convince *b* that *c*'s argument is incorrect.

To conclude, it seems that producing an imaginary rhetorical question that one ascribed to one's adversary, and subsequently responding to it constitutes a powerful means of persuasion. It brings the issue under dispute and the legal theory of the case to human scale, while at the same time it highlights the very adversarial structure of confrontational communication. Moreover, the whole network is supported by culturally meaningful conceptual blends.

2.2. A how-to definition: The law vs. the application of the law

In this section I focus on a definition with interrogative syntax. In particular, I deal with the predicative use of a question in the litigator's definition of a legal term to the jury. The analysis will be in terms of an imaginary or fictive question that reproduces the nature of jury deliberation. The piece of discourse to be discussed is:

(3) Express malice means, simply, *was it an intentional killing*, okay?

This utterance was produced early in the prosecutor's closing argument. At this point, the district attorney is telling the jury about the laws that govern jury deliberation, and is giving definitions of the charges that they will have to accept as proven or not. What is most striking about (3) is of course the use of an interrogative sentence type as a definition. The structure used is basically: NP + *means* + YES/NO INTERROGATIVE, whose semantics can hardly be looked upon as informative. In contrast, the judge's definition of that same term, "express malice" in his instructions for the jury in that same case was simply: "the unlawful intention to kill a human being". Let us have a look at the discourse immediately surrounding the utterance in (3):

(4) Express malice means, simply, *was it an intentional killing*, okay? [...] *Did the person who killed think about it? Did they have a choice?* [...] But what premeditation and deliberation really mean is, *was there weighing? Did the person doing the killing consider what it would do to the victim, what it would do for him?* [...] Well, let's apply this. If you apply it to this case, *was there planning? Of course there was planning.*

First, a set of questions are produced in order to explain the meaning of "express malice" and its related terms "premeditation and deliberation". Then, the definitions are "applied" to the case at hand. Such an application is through an introductory question, "*was there planning?*", which is subsequently answered in the affirmative by the utterer himself, "*Of course there was planning*". Even before the attorney gets to this final question-answer pair, it seems that by merely explaining the meaning of legal terms through questions, he is already simultaneously teaching the jury on the law and reasoning with them on the kinds of decisions, the kinds of *questions* that they will be confronted with in deliberation. There seems to be a conceptual blend involved, in which the definition of the term gets fused with its application. A similar example is provided by the defense counsel in that same case:

(5) What beyond a reasonable doubt *means* is that when you look at the case in totality, *do you know, do you have any reasonable doubt as to whether or not the defendant inflicted these injuries on Rachel.*

In that case too, the definition of a legal term, "reasonable doubt", is presented through the supposed reasoning process that the jury will have to go through in deliberation when deciding whether the term does or does not apply to the case at hand. Just like we may ask ourselves *"Where did I just put them?"* when we lose our keys, jurors are also presented by lawyers in their discourses as asking questions to themselves and each other when trying to come up with the right verdict.[7] It seems that the "definition~application" blend also recruits structure from the question-answer conversation frame. Indeed, when asked about his frequent use of questions in his definition or introduction of the legal charges, the prosecutor in the case explained:

> it may be a question that the **jury** might have! you know, that . that even though the jury hasn't got to **ask me** that question I think they **may** ask that question, so I'm gonna **ask** it **for** them and then I'm gonna respond to it! [...] [I am] answering questions that I think the jury will be asking in the jury room.

Since the jurors remain silent in their seats, and are not supposed to interrupt the counsel's argumentation with questions, counsels need to attempt to put themselves in the jurors' minds and create the dialogue the jury may otherwise engage in. In that sense, the counsel speaks for the jury in a similar way that the pet owner speaks for the pet in the lost pet ad in the introduction. The same individual in actuality, the public prosecutor uttering the discourse in (3) and (4) needs to split himself into multiple identities simultaneously. He needs to be the questioner as much as the provider of the answers. At the same time, he is not merely splitting himself into a double questioner-answerer role, but he is also blending his identity with the jury as audience, and with the different jurors in deliberation in the jury room.

Interestingly, both the definition~application integration and the identity blends seem to be grounded on a stabilizing aid or *material anchor* (Hutchins 2002). Notice the verdict form that he jury in this case had to fill in:

---

[7] In the popular O.J. Simpson trial, for instance, the district attorney told the jury: "all you have to do is decide *is it more probable than not, did he probably do it?*". Data from the free on-line transcripts of Sharon Rufo, et al. vs. Orenthal James Simpson, et al. case, Los Angeles, CA, 1996. Vol.50. Court TV: ww.courttv.com/casefiles/simpson.

> We, the jury in the above entiled cause, find the defendant, X, __________ [...]
> (GUILTY) (NOT GUILTY)
> And we further find that the above offense _______ willful, deliberate, and
> *(was) (was not)*
> premeditated, within the meaning of Penal Code section 189

The jury needs to fill in the blanks with either of the two possibilities that are specified for them. The cognitive task assigned to the jury is guided through an observable object, the verdict form, with which they need to interact during deliberation. This piece of paper undoubtably models the jury's reasoning process and the use they make of the definitions of the legal terms they have been told about during the trial. As it is, this material anchor contains the specification of the precise decisions to be made, the relevant questions to be answered. That is where their recollection of the meanings of legal terms needs to be applied. And a good litigator knows that, and will try to come up with useful *how-to* definitions, blends of the law and its application.

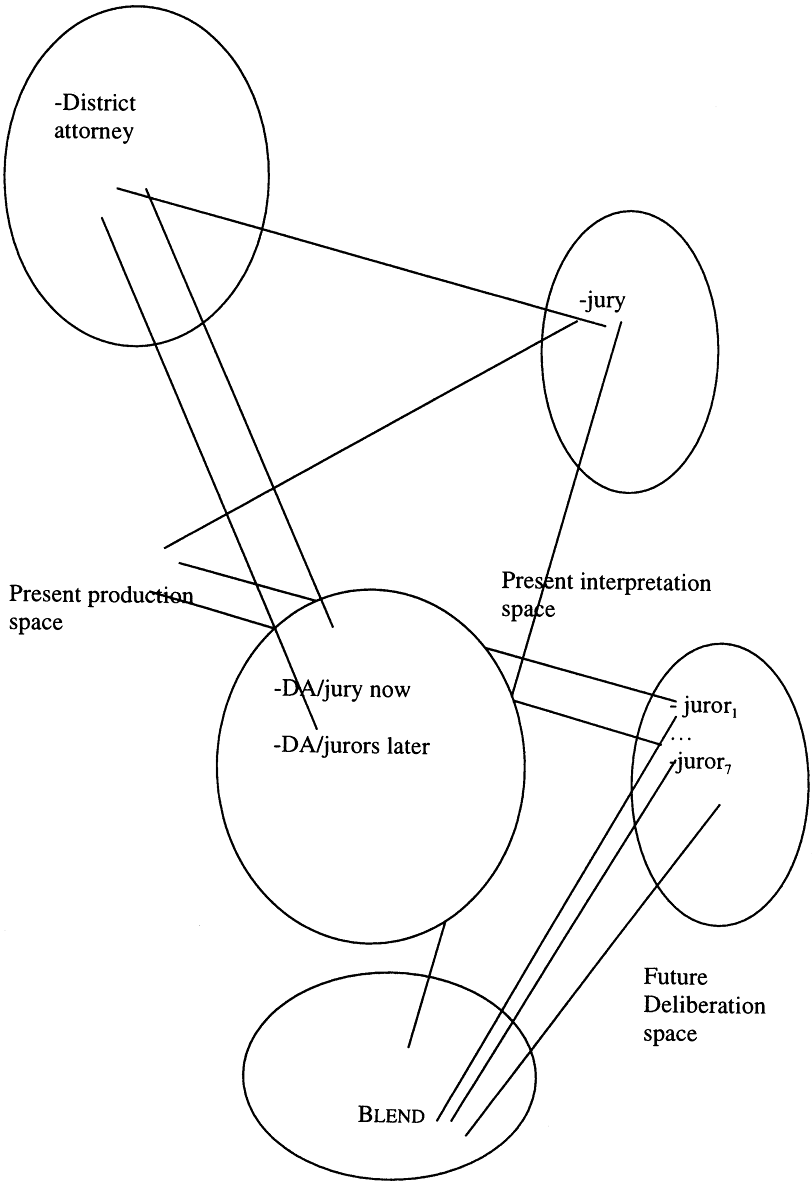


Fig. 2. Jury's deliberation: "Express malice means, *was it an intentional killing, okay?*"

In sum, it is clear that the ultimate goal of the litigator giving a definition of a legal term to the jury is not to educate them in the law. Rather the ultimate aim is to make sure the jury comes to the desired verdict. Thus, the counsel blends himself with the different jurors in a single-multiple identity blend, which conceptually integrates the present closing argument and the future deliberation. At the same time, the imaginary interaction the litigator sets up, draws the jury into the conversation, as those that will eventually answer the relevant question with him. This integration network has the persuasive power of turning addressees into co-constructors of discourse.

3. Final remarks

In this paper I have dealt with cases in which a question was produced to prompt an imagined situation of communication. In those cases, the schematic conceptual structure of situated question-answer interaction served as a frame to structure discourse flow and reasoning. I hope to have shown that there exists an interrelation between the production of particular question uses and the global socio-cultural configuration of the setting in which they occur. In contextualized settings, understanding a particùlar question use and a potential subsequent response to it may require more than a look at the semantics of the words or construction(s) in which they appear. Lastly, it seems that Conceptual Integration theory can be used to account for on-going interaction in institutional contexts.

References

Casanovas, Pompeu. 1998. Pragmática de las decisiones judiciales: Diferencias de razonamiento y lenguaje ante el jurado. *Cuadernos de Derecho Judicial.* Consejo General del Poder Judicial.

Coulson, Seana. 2001. *Semantic Leaps: Frame-Shifting and Conceptual Blending in Meaning Construction.* CUP.

Coulson, Seana & Todd Oakley (eds.) 2000. Blending. *Cognitive Linguistics* **11**.

Fauconnier, Gilles. 1994. *Mental Spaces.* CUP.

Fauconnier, Gilles, & Mark Turner. 1996. Blending as a central process of grammar. In A. Goldberg (ed.) *Conceptual Structure, Discourse and Language.* Standford: CSLI.

Fauconnier, Gilles & Mark Turner. 1998. Conceptual integration networks. *Cognitive Science* **2**(1): 133-187.

Fauconnier, Gilles & Mark Turner. 2002. *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities.* London: Basic Books.

Givón, Talmy. (ed.) 1997. *Conversation: Cognitive, Communicative and Social Perspectives.* Amsterdam/Philadelphia: John Benjamins.

Goodwin, Charles. 1994. Professional vision. *American Anthropology*, **96**(3): 606- 633.

Hutchins, Edwin. 2002. Material anchors for conceptual blends. Draft.

Ilie, Cornelia. 1994. *What else can I tell you? A Pragmatic Study of English Rhetorical questions as discursive and argumentative acts.* Stockholm: Almqvist & Wiksell International.

Lakoff, George. 1996. Sorry, I'm not myself today: The metaphor system for conceptualizing the self. In G. Fauconnier & E. Sweetser (eds.) *Spaces, Worlds, and Grammar.* Chicago and London: University of Chicago Press.

O'Barr, William.M. 1982. *Linguistic Evidence:Language, Power, and Strategy in the courtroom.* New York: Academic Press.

Pascual, Esther. 2001. Why bother to ask rhetorical questions (if they are already answered): A conceptual blending account of argumentation in legal settings. Paper presented at the 7th International Cognitive Linguistics Conference.

Pascual, Esther. To appear. *Imaginary trialogues: Conceptual blending and fictive interaction in the courtroom* . The Hague: LOT.

Tannen, D. 1998. Litigation Is War. In *The argument culture: Stopping America's war of words.* New York: Ballantine Books.

Dept. Cognitive Science
9500, Gilman drive Dept. 0515
University of California, San Diego
La Jolla, CA 92093

Dilemmas and Paradoxes in Linguistic Fieldwork

JØRGEN RISCHEL
University of Copenhagen, Denmark

0. Introduction

Linguistic fieldwork has a long and glorious past in the United States, first and foremost because of the devotion of Boas, Kroeber, Sapir and their followers to the documentation of the numerous native languages of North America. I feel very humble in this context, and it is a great honor for me to have been invited by the Berkeley Linguistics Society to address this topic.

Linguistic fieldwork involves three essential components: native speakers, the target language, and the technical outfit and skills of the field worker. This can be observed, with large variations in emphasis, in the traditional readings of the sixties such as Healey (1964), Longacre (1964), and Samarin (1967), or in the brief statements in Gudschinsky (1967: 4-7) about "Field Procedures" and "Use of Informants." Issues to do with native speakers as individuals and as a speech community are, however, much more in focus nowadays, and personally I feel they should be given the highest priority. The present paper is very broad in scope and may seem almost redundant since there has been an recent upsurge of interest in fieldwork methodology, as witnessed by publications such as Vaux & Cooper (1999), Newman and Ratliff (2001). Still, I hope to raise some live issues.

There is a general prevalence in linguistic fieldwork of settings involving more or less fluently bilingual speakers. In this presentation, on the other hand, my focus is on the interaction with speakers of oral languages who are monolingual or have at most a very restricted competence in other languages, who are absolutely non-literate (in any language), and who have had limited exposure (or virtually no exposure) to non-native behaviour such as the ways in which they may be approached by field workers. They are the prototype of so-called "unsophisticated informants" (Healey 1964). My reason for this particular focus is that I happen to have been working in such speech communities over the last twenty years.

There used to be one school which explicitly emphasized the necessity of formalizing discovery procedures, namely post-Bloomfieldian structuralism, but in fact its methodology did not specify how fieldwork was to be performed in practice, with the more Sapir-inspired Pike-Nida trend as a notable exception (Nida 1946).

The main point, as I see it, was a warning against involving native speakers in metalinguistic discussions about the meanings of linguistic forms.

My own experience is that it is difficult and hardly fruitful to steer clear of discussions of meaning. Native speakers are often little interested in whether linguistic tokens are alike or different whereas they are extremely interested in helping the ignorant linguist to understand what an expression means. After all, that is what language is about. When working on a little explored language it is a source of potential error that the field worker and the native speaker misunderstand each other's intentions. The native speaker expects to be engaged in meaningful communication in which the spoken language is the medium rather than the object of study. If eventually it becomes clear that the field worker wants to focus on the language itself, native speakers who have little or no experience with researchers tend to think in terms of synonymy rather than phonetic similarity. There is a priori little inclination to engage in seemingly pointless comparisons of phonetically related words or utterances. Personally I am often in doubt about responses to questions about *similar or different?* unless I have specified that I want to know whether things *have the same "name" or not.*

The retrieval of data was never an issue that loomed large in standard textbooks on linguistics. It remained rather peripheral, even to scholars with a long experience in approaching linguistic data. Introductory textbooks teaching morphological or syntactic analysis are often very bad in this respect: they tend to depict raw data as a set of carefully selected and flawlessly transcribed forms. Thus they create a dangerous illusion in the minds of students unless the teacher provides a counterbalance by explaining and demonstrating how crucially important it is for the whole field of linguistics how the raw material used as input was retrieved in the first place.

1. Phonetic issues

I myself have attempted to say in various papers over the years (cf. Rischel 1983, 1987, 1989) that one cannot overestimate the importance of the initial confrontation with phonic data and the initial generalizations over data made in the field. Errors of perception or sloppiness in transcription may cast a long shadow over all the ensuing work at home, and even worse: over the work of other linguists who trust the information. Communication breakdown between field worker and native speaker is a more tangible source of error, and the field worker may not detect such breakdowns. There is an additional danger in that one's data may be skewed in some way from the start because of the conditions under which one works. I shall discuss some of these sources of error in the present paper.

One of the most urgent questions in fieldwork is how to achieve phonetic adequacy. I wish to highlight one issue, namely the status of field notes versus electronic recordings. First some words about the choice of electronic recording media: I have found it increasingly rewarding to use video along with or even instead of audio tapes. Especially if one is recording narratives, it is a matter of so-called total communication involving significant mimicry and gestures as, for example, when the narrator uses different points in space deictically to refer to the actants in the story, or uses his hands to show trajectories of movements which are significant in the story. This is an essential part of the encoding of its contents, and audio recordings of narratives may be defective or at least much harder to understand without such visual support.

The importance of video recording is a reflection of the accessibility, in the field situation, of other important sources of information than the sound of the speech itself. Pike spoke of phonemics as a way of *reducing* speech to writing. I would say that tape recording is in itself a way of reducing oral language: one peels a lot of information off, and this reductionism is of course carried a very big step further if one transcribes the data linearly by leaving out such paralinguistic features as expressive voice quality and variable speech tempo. These phenomena are communicatively crucial and play a paramount role in daily conversation, but unfortunately they do not fit easily into the format of conventional phonetic transcription, and as linguists we have generally not been taught to pay much notice to them. Still, it is easy to point to instances where paralinguistic or extralinguistic phenomena such as voice quality or the direction of the speaker's glances are cues to actants in the story like pronouns, for example.

This paper is *not* about instrumental phonetic field techniques. Still, I wish to mention that although the access to instrumental analysis either in the field or at home is becoming increasingly rewarding for the purpose of analysis and important for the purpose of documentation, there is an inherent dilemma in modern phonetic fieldwork. It is part and parcel of the professional phonetician's education that much instrumental phonetic analysis requires averaging over several tokens recorded under so similar conditions that it makes sense to process them statistically. In the field, on the contrary, it may be well nigh impossible to monitor the variability; everything changes all the time. What one can hope to get out of a natural field situation, then, is a selection of *typical* tokens rather than average tokens.

Even as a phonetician I would claim that the most indispensable accessories in the early field situation are not electronic devices but ears, eyes, a pencil and a notebook with good quality paper. I have always found that the notes made in the field have a first-rank status as data, no matter how much electronic recording one has made. Listening in the field and listening to a tape at home are two very different things. In the field one may hear the same word repeated several times and form *an impression which integrates the variability occurring over the various tokens*. One can listen at a distance and observe the movements of the lips and the jaw at the same time, or one can have one ear close to the speaker's mouth to hear such things as the degree of voicing in stops more clearly. One can then combine these impressions into a faithful and meaningful transcription. On the other hand, one cannot manage to write speech down in real time with the use of phonetic notation, so if the speaker volunteers a lot of data spontaneously but is reluctant or unable to repeat it later, one is at a loss without electronic recording. Especially in the case of narratives I have had to rely very much on listening to tapes or viewing videos at home in order to make raw transcriptions which can then be rechecked in the field (hopefully). The electronic recording is also indispensable as a means of lasting scientific documentation, but according to my experience tapes are of very little use to posterity if the field worker does not manage to supply competent transcriptions before handing them over to others.

I shall now switch to a related issue: how to combine accuracy and consistency in the narrow or broad phonetic notation. I have written a paper about this issue several years ago, and I shall just recapitulate a couple of points and add a new comment. As a retired phonetician I ought not to say anything negative about the IPA notation, but I cannot help stating that I think it has inherent problems which appear in fieldwork. One problem is the decision whether to choose a fairly narrow or a fairly broad notation. It makes sense to stick to a very narrow transcription if one makes sure that the phonemic status of each wordform is under control; in that way one captures all the phonetic variation occurring over the data, but a major difficulty is the faithful notation of vowel qualities. To devote a lot of attention all the time to the proper use of diacritics on vowel symbols, may be cumbersome and time consuming in the sometimes stressful field situation. It may even become an unnecessary source of distraction from the issue one is concerned with, especially if the words in question have already been carefully phonemicized and entered into the lexicon.

I have felt this problem particularly when I was taking down phrases or sentences for the purpose of grammatical analysis. If new words occur in the process, I of course transcribe them in narrow IPA, but I have found that for the words that occur all the time I tend to jump immediately to my phonemic image of them which means skipping to a broad transcription so that the overall notation becomes a mixture of narrow and broad transcriptions.

Using a hybrid transcription is all right if one can keep track of the two modes of notation, but there is a serious source of error here because of the nature of the IPA. In this phonetic alphabet a vowel symbol without diacritics has two quite different roles: if used in broad notation it can function as a cover symbol for a variety of vowel qualities within a certain range; in narrow transcription, on the other hand, it denotes a fairly specific vowel quality by virtue of the cardinal vowel convention. If I enter a word from my field notes into my lexicon I must all the time recheck whether vowel symbols without diacritics are meant in the broad or the narrow sense. The problem is particularly serious in the range of open vowels: I honestly think there are too few *a*-symbols in the IPA alphabet. My private solution now is to switch between a fairly narrow IPA and a romanization which I cannot possibly mistake for IPA, and to use the IPA whenever it is relevant in order to capture unpredictable phonetic detail.

2. Inadequate coverage of linguistic usage

The data at one's disposal may be incomplete or skewed in numerous ways, and that may have dire consequences at all levels of linguistic analysis and documentation.

That is a particularly serious issue if one is working with a language which is moribund or "terminal" (Dorian 1986), so that there may be little chance of repairing inadequacies later. Since it is a vast and confusing topic, I must limit myself to a few types of significant skewedness or non-representativity:

One problem is the *choice of speaker*. The members of any speech community, even a very small tribe, may be as different as people in the street: some are eloquent, some say less but think more conscientiously about the information they give; others again refuse to perform in a fieldwork setting, because they consider themselves dumb or rustic or because they find (with some justification) that the whole set up is phony. But these same people may thaw up if one manages to get involved in genuine communication about real issues. The most willing and eloquent speakers, or the ones with the best voices, are certainly not always the ones that yield the most genuine information about the language one is studying. I myself once recorded quite a bit of word elicitation with a young boy because he had a clear voice and could control himself so as to speak at a moderate speed, but eventually the material must be scrapped because it turned out that he imitated my own speaking style (with all of its speech errors) out of veneration for me, whom he obviously regarded as a close senior relative. Then again some speakers may speak with a more or less slurred articulation because of missing teeth or other impairments, but these (typically old) people may be a gold mine of information e.g. about lexicon or conservative morphology if one can find out what they are saying. Some of my best field sessions were made with people sitting at leisure, e.g. with a pipe between the teeth. This sometimes brought me—as a phonetician—into despair, but it was a dilemma since attempts to force the speaker to abandon his or her pipe might ruin the mood which was a prerequisite to the volunteering of precious information. There is often a conflict between what satisfies the phonetician and what satisfies the syntactician or the cultural researcher. This is an unfortunate dilemma if the field worker performs in several roles at the same time.

A related challenge is genuine *variation across the language community*. Although it is convenient to work with just one native speaker and define one's goal as the description of an idiolect rather than a dialect or language in its totality, the speaker may have very idiosyncratic features of phonology and even of grammar or lexicon. Others in the group may speak differently.

I have met with impressive lexical variation in an oral language in Southeast Asia which I am working with: the Austroasiatic hunter-gatherer language Mlabri. It is obviously highly relevant in such cases to know as much as possible about each speaker's family and experiences in life. Some such inter-speaker variation may persist for a long time as idiosyncratic characteristics of individuals or families. In Mlabri this is true of certain strange phonological fluctuations which I observed in a language variety spoken by a tiny and fairly isolated group of twenty-odd persons living in the forest close to the wild border between Northern Thailand and Laos. One of their variational features is an idiosyncratic preference for either a final dental or a final velar stop in the definite article, another is the idiosyncratic substitution of a palatal glide for an alveolar lateral. If one looks at the only early recording of the language: an otherwise rather inadequate word list from 1938, it most remarkably shows these very fluctuations. Other inter-speaker differences are more important for an overall characterization of the language since they are associated with socially defined groups; it is noteworthy in particular that a language such as Mlabri exhibits several lexical doublets such that one word of a synonym pair is dubbed as male speech, the other as female speech.

More or less orthogonally to the dimensions of sociolectal and idiolectal variation there is in probably every speech community a variation over *speaking styles*. This is a challenge to fieldwork. In order to understand the phonological mechanisms of a language it is indispensable to have access to several renderings of the same forms in distinct, normal, and casual style, or whatever style distinctions seem appropriate to make. This is a very important property of language, and it is exploited communicatively. For diachronic-comparative purposes highly distinct forms are often of crucial importance, but speakers may construe artificial pronunciations or make spurious etymological associations if asked to speak over distinctly.

Narratives give access to natural connected speech, to genuine syntax, and to a rich lexicon, but how do we approach narratives in a hitherto unexplored language? And how do we find the ideal speaker? I have always looked with envy at those classic editions of folklore which look as if everything worked smoothly, as if everything was delivered in dictation style by story tellers or by those reciting ritual texts or traditional songs. Like many others I have been recording narratives quite extensively, and that works fine in the recording phase, but I think we all have the experience that the interpretation requires much more extensive fieldwork with the same speaker or with other speakers afterwards. In the case of Mlabri my best material stems from a speaker who is a narrative genius but is half deaf and has now also developed a disease so that he has difficulty phonating. To make things worse he speaks incredibly fast.

The approach I have chosen is to play the recordings back several times to his children who can perceive almost all he says. Their task is to repeat the words *exactly* so that I get a supplementary, distinct tape recording. The session is monitored by the father who endorses or corrects their rendering. That may sound easy but it did not work smoothly from the start. These youngsters were convinced that I only wanted the contents of each story and that my only problem was lexical incompetence. So they kept paraphrasing the paragraphs I was supposed not to understand, instead of repeating the words more distinctly. As they obviously know how to speak to a half deaf person, my strategy has been to tell them that I myself am handicapped by poor hearing but I treasure their father's *exact* words. Eventually that helped. Still, all paraphrasing is of value. In fact my Mlabri narrator himself exhibits intriguing verbal variation if he tells a story a second or third time.

To avoid too much skewedness in the representation of speaking styles and of genres I have always found it indispensable to record a lot of small talk beside narratives and lexical material. One way to get the small talk is to work monolingually, as a kind of participant observer. In this way one also avoids interference from another language, but then the problem arises that many native speakers change their way of speaking when speaking to outsiders. I shall return to this issue presently.

3. The dangers of questionnaires and of interpreters

A quite different kind of skewedness or downright distortion in the data stems from the field worker asking leading questions to the native speaker or—in a bilingual setting—asking him or her to translate words or sentences from a questionnaire into the native language. The use of a standard questionnaire containing nouns and verbs and phrases which "one can be pretty sure to find everywhere" as a first approach has a long tradition, and it was endorsed by the publication of a field questionnaire in the early years of the CIPL (Cohen 1931).

The success of such an approach certainly depends on the setting. My own experience with forced translation of sentences from a peer language into a hunter-gatherer language is dismal. I am not happy to make theoretical morphosyntactic points that way. If, for instance, a bilingual speaker is to translate from a peer language with extensive use of particles to the native language which perhaps has lesser use of such particles in certain syntactic structures, the speaker may feel a pressure to come up with equally explicit constructions in his or her mother tongue and thus to produce calques with an atypical or even spurious use of particles. Each language has its typical and more or less unique profile, and if one looks at natural speech it is likely to exhibit a repertoire of truly elegant ways of encoding messages, though these may well be poor in quantifiers, or comparison, or relative constructions, or whatever else the general linguist wants to cover in his or her syntactic elicitation.

In this respect I see no principled difference between heavily morphological languages such as Eskimo and almost isolating languages such as Mlabri, just to mention two languages of traditional hunter cultures with which I myself happen to be familiar. I belong to those who find it more rewarding to attempt to understand the ethos of the language rather than search for reflexes of alleged syntactic universals. The hope, then, is that the non-squinting approach may ensure solid data to be used in typological generalizations.

To be blunt, I never believed in questionnaires at all, and when it comes to the languages of peoples who were traditionally hunter-gatherers such as the Eskimos and the forest people in Southeast Asia, I have found the results of this approach to be sometimes slightly grotesque because of the tendency to cultural bias. One can be sure that the resulting lexical data form no representative sample of genuine, traditional vocabulary. In addition, phrasal data obtained via a questionnaire may not permit valid syntactic generalizations.

Item-per-item elicitation of data from a list may nevertheless be a meaningful first aid, especially if done by a competent linguist. The real danger is that the use of questionnaires invites a brute force approach in which minority languages are screened, as it were, by research assistants with little training in linguistic documentation and possibly with no experience in interacting with persons of a quite alien group or in monitoring the performance of hired interpreters.

In what I have already said, I have made some reference to the linguistic vehicle of communication between field worker and native speaker. This leads over to my next issue: how to ensure that interesting information is efficiently transferred without too much intermittent communication breakdown.

I want to say a couple of bad things about the use of interpreters. Firstly, an interpreter may, like any person with real or apparent authority, intimidate the native speaker and compromise the genuineness of the data. A field worker performing as a half proficient participant observer stands a better chance of not having this influence (but then he or she may influence the native speaker's language use in other ways; see above and below). Secondly, if the interpreter and the native speaker communicate via a common lingua franca which the fieldworker may not master, there are sources of error at all stages when a question is transmitted successively from the field worker's language via the interpreter's use of the lingua franca to the native speaker who attempts to understand what the interpreter said, or when the answer is transmitted back.

4. Monolingual fieldwork

The very opposite procedure of those mentioned so far is to work directly with a native speaker, without any interpreter and perhaps even with no shared language of communication, especially if the target language is little known or hitherto unexplored and is spoken by a relatively isolated group. Most people probably start such a session by showing objects or actions, directly or via pictures. One can indeed get a basic vocabulary that way (although pictures easily confuse people who are not accustomed to visual representations of things, cf. below). But eventually the range of easily accessible items to be presented and identified is exhausted, and I know by personal experience that this may lure the field worker into falsely believing that the field notes now give a fair coverage of the lexicon. After all this is a tribal language! The truth, of course, is that so-called "small" languages are extremely rich and complex when it comes to semantic fields which are culturally or socially important.

I greatly prefer a setting in which one can *communicate* monolingually with the native speaker (also cf. Everett 2001). It is true that this approach may be frustrating if one does not understand enough of what is said, so that a lot of information is lost. Also misunderstandings may arise e.g. if the field worker's questions are too poorly phrased.

On the other hand native speakers are both highly motivated and free of anxiety so they are at leisure to explain carefully and instructively provided that the main vehicle of communication is their own language. This means that whatever the field worker really understands tend to be fairly reliable. Practically all I know about the rich Mlabri vocabulary of verbs that refer to body postures and body movements was communicated to me monolingually and explained visually. It did not take too much speaking proficiency to ask simple questions such as: "If you want me to sit like this, or to move my hand like this, how then do you say that?"

In the case of nouns, on the other hand, the lack of a fluent reference language may create difficulties with the retrieval of lexical information as soon as one proceeds beyond objects within easy reach. As for animals (including insects and worms) and plants in the forest, I have had disappointingly little success in attempting to cover the native Mlabri lexicon adequately.

Even showing charts of different species may create confusion with "unsophisticated" native speakers if consecutive pictures are not to the same absolute scale. I have been using a picture book with local birds in which each colour chart typically included two or more different species as if they lived close together, and then a different problem arose: birds in such a chart were often identified as tokens of the same species, e.g. as 'fatherbird', 'motherbird' and 'offspring'. It is also my experience that some of those consulted take pride in identifying as many species as possible, at the expense of accuracy, whereas others takes the opposite stand and hardly make any wholehearted identifications. It goes without saying that there are much more professional approaches to such elicitation by means of photographs, but in the field situation, especially inside a mountain forest, it tends to be a matter of what is practical.

For many lexical entries I just have more or less vague descriptions or sometimes broken, incomplete or decaying parts. I have had to consider it impractical to make extensive zoological and botanical safaris together with experts in the natural sciences in order to find the objects *in situ*, both because the presence of several outsiders easily ruins the hard-won, relaxed atmosphere of monolingual fieldwork with shy native speakers, and also because of practical problems (the Mlabri suggest that some information retrieval is best accomplished by climbing trees, which even to them themselves is sometimes dangerous). Therefore, many such entries in my notebooks have pseudogeneric, descriptive translations like 'species of tree described as tall, with small leaves and inedible fruits, growing high up in the forest' or 'inedible tuber species' or '(presumably:) cicada species, not eaten'. This is not satisfactory, but I think it is potentially more misleading for users of the data if one makes *a posteriori* identifications and supplies a Latin name for such a species on shaky evidence. The most important thing is to strive not to supply ill-founded information on the basis of guesswork.

The alpha and omega of lexical fieldwork is to recheck one's data over and over again in different situational contexts or at least in different discourse contexts. But that may be difficult when it comes to words for things which are outside one's own range of experience. As for animals and plants in the deep forest the Mlabri sometimes volunteer a Thai or Lao name of such an animal or tree, but their use of Thai or Lao is not reliable. I once recorded a mythological narrative of The-Great-Flood-Myth type, and when I went over it with the story teller he explained about a white bird flying over the ground after the natural disaster with fresh seeds in its crop. I was in doubt about the Mlabri word for the bird, and he then explained in the local lingua franca that it was an owl. This sounded spurious, so I later showed him pictures of an owl and a pigeon. Unsurprisingly, he pointed at the pigeon but said in the lingua franca: "Yes, it was an owl!"

We all know that one of the best ways of getting genuine everyday speech is to spend time on engaging in conversations about relevant topics. In this way one also stands a much better chance of recording vocabulary with correct glossing. The ideal case is for a situation to arise which invites such a conversation. Once, for example, a young Mlabri woman suddenly felt an urge to explain in very much detail about her recent troubles giving childbirth. The child's grandmother afterwards gave her version of the whole event, and others present elaborated on the story so as to make sure that I had understood every part of this drama. The focus was not at all on the language but on the components of such an event, but since they were eager to make me understand, I got a lot of new vocabulary, and it became an exemplaric case of interactive linguistic fieldwork.

Still, a monolingual approach has its shortcomings. The output may be biased by the nature of the communication. This problem is shared with sociolinguists of the Labov brand making a so-called "sociolinguistic interview." One should be prepared for a shortage of reliable data on *interrogative constructions*, for example, if one uses an interactive approach. There are several reasons for that. Firstly, the situation itself may not invite a lot of real questions put *by* the native speaker *to* the field linguist. In addition, it may be awkward for the tribesperson to address a quest for information to a seemingly more prestigious outsider. The field worker may then attempt to provoke samples of question types, but the sensible native speaker is likely to find it weird to be asked to produce a pointless question. It is obviously preferable for the linguist to stay for a long time in the community and to pick up questions from conversations between tribespeople. It should not be overlooked, however, that in-depth processing of such material actually requires very thorough familiarity not only with the language as such but with the whole communicative context and with the communicative strategies used in the speech community, with expected presuppositions and with speech act types. It may be easier and in the initial phase more rewarding to look at data for which the contextual setting is much simpler than in natural conversation. I have found it useful to look at questions imbedded in folk tales in which the characters talk together. If at times one does not understand the reason why this or that question occurs in the tale, one can be sure that the native speakers will gladly volunteer to explain that in accordance with their interpretation of the narrative.

A further danger in the monolingual setting is that the unfortunate field worker may happen to be retrieving *spurious* data on connected speech. I mentioned earlier that native speakers may choose, perhaps inadvertently, to adjust or even distort their speech so as to accomodate to the field worker. Such distortion, particularly in the form of excessive simplification in morphosyntax and lexicon, may also occur if the native speaker detects that the outsider's linguistic proficiency is poor. Predecessors may cast a shadow of this kind. One elderly couple in Northern West Greenland spoke to me in an almost unintelligible kind of pidgin Greenlandic, largely without inflections, which is serious in a polysynthetic language. The reason turned out to be that they had been hired for practical work by the local Danish school teacher who had arrived several years before with little command of the language, so they had chosen that strategy rather than confuse him with inflections. This was an easily detectable pidgin situation, but in less extreme cases the field worker may remain for a while in happy ignorance about the spuriousness of the linguistic input, especially if it is his or her first encounter with a little-known language. Hopefully, the uncomfortable truth reveals itself sooner or later.

5. The ethics of fieldwork

I wish to round off by emphasizing that some of the most serious dilemmas have to do with the basic ethics of fieldwork. Speakers are not automata but members of a speech community. We must all the time ask ourselves: do we understand their attitude, and do we know what is the appropriate way of showing our appreciation of their help? Some native speakers offer their help freely and without limits if they feel that the endeavour may assist the survival or at least the memory of their language and culture. They may react very differently if they assume that the purpose of the fieldwork is to gain a personal advantage on the basis of shallow knowledge. Sometimes the quality of our field data depends on the extent to which the native speaker can see the point in painstaking accumulation of seemingly trivial information.

In the field we are forced to acknowledge that a language is not just an autonomous object but the backbone of the cultural make up of its speakers. Access to a language also means access to facets of spiritual culture which would otherwise remain inaccessible. But the retrieval of information always comes at a cost. In one sense we serve the cause of humanity by struggling to document a vanishing language and culture, but in doing so we may violate the integrity of its bearers if we ask persistent and seemingly silly questions about matters which to them are sensitive. Researchers in the field must take a stand on such issues personally, and I hope we all do.

References

Cohen, Marcel. 1931. *Questionnaire Linguistique*. Publications de la Commission d'Enquète Linguistique. Comité International Permanent de Linguistes.

Dorian, Nancy C. 1986. Gathering language data in terminal speech communities. In Jushua A. Fishman et al. (eds.) *The Fergusonian Impact. In Honor of Charles A. Ferguson*. Vol. 2: *Sociolinguistics and the Sociology of Language*, 555-75.

Everett, Daniel L. 2001. Monolingual field research. In Newman, Paul & Martha Ratliff (eds.) *Linguistic Fieldwork*, 166-88.

Gudschinsky, Sarah C. 1967. *How to Learn an Unwritten Language*. New York: Holt, Rinehart and Winston.

Healey, Alan. 1964. *Handling Unsophisticated Linguistic Informants*. Linguistic Circle of Canberra Publications, Series A: Occasional Papers 2. Canberra: Linguistic Circle of Canberra.

Longacre, Robert E. 1964. *Grammar Discovery Procedures: A Field Manual*. The Hague: Mouton & Co.

Newman, Paul & Martha Ratliff (eds.). 2001. *Linguistic Fieldwork*. Cambridge: Cambridge University Press.

Nida, Eugene. 1946. Field techniques in descriptive linguistics. *International Journal of American Linguistics* 13.3: 138-46.

Rischel, Jørgen. 1983. Fieldwork data as input to instrumental analysis: A dilemma. In Pieper, Ursula & Gerhard Stickel (eds.) *Studia Linguistica Diachronica et Synchronica*, 709-22. Berlin: Mouton de Gruyter.

Rischel, Jørgen. 1987. Phonetic transcription in fieldwork. In Almeida, António & Angelika Braun (eds.) *Probleme der phonetischen Transskription*. Zeitschrift für Dialektologie und Linguistik, Beiheft 54, 57-77. Stuttgart: Franz Steiner Verlag.

Rischel, Jørgen. 1989. Fifty years of research on the Mlabri language. A re-appraisal of old and recent fieldwork data. *Acta Orientalia* 50: 49-78.

Samarin, William J. 1967. *Field Linguistics*. New York: Holt, Rinehart and Winston.

Vaux, Bert & Justin Cooper. 1999. *Introduction to Linguistic Field Methods*. München, Newcastle: Lincom Europa.

Stenhøjgaardsvej 57
DK-3460 Birkerød
Denmark

jr@cphling.dk

Directionality in Vowel Harmony: The Case of Karajá (Macro-Jê)

Eduardo Rivail Ribeiro[1]
University of Chicago and Museu Antropológico (Universidade Federal de Goiás, Brazil)

0. Introduction

This paper describes some unusual features of vowel harmony in Karajá, a Macro-Jê language from Central Brazil, outlining some of the implications they may have for a broader typological characterization of vowel harmony systems.[2] Karajá presents a system of vowel harmony in terms of the feature [ATR] 'advanced tongue root'—apparently, the first documented case of [ATR] vowel harmony in a South American language (Ribeiro 2000). However, when compared with more typical cases of tongue-root harmony systems, such as the ones found in the African languages Yoruba (Archangeli & Pulleyblank 1989) and Turkana (Noske 1995), Karajá presents some interesting peculiarities. For example, in Karajá, [ATR] alternations involve not only front and back vowels (as is typically the case with well-known ATR harmony systems), but high central

[1] The fieldwork on which the present paper is based was conducted under the auspices of a grant from the Tinker Foundation, administered by the Center for Latin American Studies of the University of Chicago. I would like to thank these institutions, as well as CNPq (the Brazilian National Council for Scientific Development) and the Wenner-Gren Foundation for Anthropological Research, for their financial support. My understanding of this topic has benefited greatly from discussions with Gunnar Hansson, to whom I would like to extend my acknowledgments. Any remaining shortcomings are, of course, my entire responsibility. The research reported in this paper is still in a preliminary phase, and is part of a larger project of description of the Karajá language. Any comments, suggestions, and/or criticisms would be very much appreciated.

[2] Karajá is spoken along the Araguaia River (an indirect tributary of the Amazon River), in the states of Goiás, Mato Grosso, Tocantins, and Pará. The phenomena described in this paper are common to all of the four Karajá dialects (Southern Karajá, Northern Karajá, Javaé, and Xambioá). The language presents systematic differences between male and female speech (Ribeiro 2001), but such distinctions bear no relevance for the discussion of the facts described in this paper. The data on which this paper is based are mainly from the female speech of the Southern and Northern Karajá dialects. Abbreviations: ANTI 'antipassive', CTFG 'centrifugal direction', EMPH 'emphatic', FUT 'future', IMPER 'imperative', IMPERF 'imperfective', INTR 'intransitive', LOC 'locative postposition', POT 'potential', PERF 'perfective', PROGR 'progressive', REL 'relational prefix', TRANS 'transitive'.

vowels as well.[3] Another interesting characteristic of Karajá vowel harmony is its extreme pervasiveness and productivity. It applies not only accross the different constituents of a compound, but across word boundaries as well.

Vowel harmony in Karajá can be roughly described as a process of regressive propagation of the feature value [+ATR] to vowels that would otherwise surface as [-ATR], such as represented schematically in (1) below. Any morpheme containing a [+ATR] vowel can trigger harmony, including stems, affixes, and clitics, a property which characterizes Karajá vowel harmony as a *dominant-recessive* system. However, unlike other well-known dominant-recessive vowel harmony systems, such as Turkana (Noske 1995) and Nez Perce (Rigsby & Silverstein 1968), vowel harmony in Karajá is strictly directional, applying exclusively from right to left. Thus, in the example below, vowel harmony is triggered by the [+ATR] vowel of the imperative particle, turning preceding [-ATR] vowels into [+ATR]; notice that the [-ATR] vowel of the emphatic particle remains intact:

(1) V V

[+ATR]

(2) *b-ɛ-ɗɛhɛ=ikuɗĩ=hɛ* [beɗeˈheikunihɛ]
2-INTR-look=IMPER=EMPH
'Look!'

The fact that [-ATR] vowels can follow, but not precede [+ATR] vowels clearly shows that vowel harmony in Karajá is strictly a right-to-left process. The straightforward relevance of directionality for the description of vowel harmony in Karajá challenges theories that discard directionality as an independent parameter of assimilation, such as the one proposed by Bakovič (2000), among others (Beckman 1995, 1997, 1998; Lombardi 1996, 1999). In this view, directionality is an epiphenomenon dependent mostly on the morphological structure of the language. As I intend to show in this paper, although such an account seems to be especially appealing for stem-controlled harmony systems, as well as for more familiar examples of dominant-recessive systems, the Karajá

[3] In fact, some languages indeed present an [ATR] contrast between central vowels, although less commonly found than the one involving front and back vowels. That is the case of Degema (Niger-Congo), for example, which is described as having a 'complete' 10-vowel [ATR] harmony system (Fulop, Kari, & Ladefoged 1998). In this language, the [-ATR] low vowel [a] constitutes a harmonic pair with the [+ATR] vowel [ə], in addition to the contrasts between [ɪ, ɛ, ɔ, ʊ] and [i, e, o, u], more commonly attested. [ATR] contrast between high central vowels, such as displayed by Karajá, seems to be a rather less common phenomenon.

data provide a strong counterexample to such claims, suggesting that such theories are inadequate as a universal characterization of vowel harmony phenomena.

1. Karajá [ATR] harmony

In previous phonological descriptions of Karajá (Fortune & Fortune 1963; Cavalcante 1992), vowel harmony is treated as a matter of height assimilation, being briefly mentioned as a process by which a high or close-mid vowel "closes" an open-mid vowel in a preceding syllable. However, this formulation would not account for a number of cases in which a high vowel would 'fail' to trigger vowel harmony (Ribeiro 2000: 80-81). This is a direct consequence of the fact that both accounts did not consider phonological contrasts such as the ones illustrated by the minimal pairs below:

(3)	a.	*wɪ*	'good'	(4)	a.	*wi*	'reciprocal'
	b.	*kɨ*	'bark fiber'		b.	*kɨ̝*	'inside'
	c.	*rʊ*	'thigh; skirt'		c.	*ru*	'eye, sight'

While the high vowels in (4) behave like the close-mid vowels, triggering vowel harmony, the high vowels in (3) behave like the open-mid vowels, undergoing it. This fact demonstrates that vowel harmony in Karajá is of the 'cross-height' type, involving not height, but the [ATR] feature. According to their behavior in triggering, undergoing, or blocking vowel harmony, the vowels of Karajá can be grouped as in the table below:

(5) Vowels according to their behavior regarding vowel harmony (*apud* Ribeiro 2000, revised)[4]

Oral

[+ATR]			**opaque**	**[-ATR]**		
i	ɨ̝	u		ɪ	ɨ	ʊ
e	ɵ	o		ɛ		ɔ
			a			

[4] One of the main phonological differences among the four dialects is that in Southern and Northern Karajá there occurs a schwa /ə/ in unstressed positions, corresponding to environments in which Xambioá and Javaé present a vowel identical to the one occurring in the following syllable: Southern/Northern Karajá *d-ɔrəɗɔ* 'tongue', *bədɛ* 'land, ground'; Javaé/Xambioá *d-ɔrɔɗɔ*, *bɛdɛ*. Although its occurrence in Southern and Northern Karajá is limited to unstressed syllables, the schwa can clearly be reconstructed for Proto-Karajá. Regardless of its phonemic status, however, the existence of the schwa does not alter the discussion of vowel harmony presented above, since the schwa is transparent to vowel harmony: *d-ɔrəɗɔ=le* (REL-tongue=EMPH) [dorəɗo'le] 'just the tongue' (Javaé/Xambioá [doroɗo'le]).

Nasal

[+ATR]	**opaque**
ĩ	ə̃, õ
	ã

As shown in (5), the only oral vowels that do not have a harmonic counterpart are the low vowel /a/, which is opaque (9a), and the mid-close central vowel /ɘ/, which is dominant (12a). As the examples in (6) below demonstrate, all combinations of vowels in a phonological word are possible, except [-ATR] vowels preceding [+ATR] vowels. When a [-ATR] vowel precedes a [+ATR] vowel, vowel harmony takes place (6d).

(6) a. [+ATR] [+ATR]
bekɘ [be'kɘ] 'harbor'
woku [wo'ku] 'inside'
-obi [o'bi] 'to see'

b. [+ATR] [-ATR]
buɗɛ [bu'ɗɛ] 'little, few'
iθɛ [i'ʃɛ] 'dance'
kɨθɛ [kɨ'ʃɛ] 'grassy, bushy'

c. [-ATR] [-ATR]
dɔrɛ [dɔ'rɛ] 'parrot'
dɛbɔ [dɛ'bɔ] 'hand'
hɛɗɨ [hɛ'ɗɨ] 'blanket'

d. *[-ATR] [+ATR]
dɔrɛ d-e [dore'de] *d-ɛbɔ kube* [deboku'be]
parrot REL-wing REL-hand palm
'parrot's wing' 'palm (of hand)'

The process of vowel harmony in Karajá is further illustrated below by examples involving the imperfective auxiliary *=r-e*, a clitic (7), and the derivational suffix *-ɗĩ* 'similar to' (8). As shown in (5) above, both high and mid [-ATR] vowels undergo vowel harmony:

(7) a. *Ø-r-a-kɔhɔɗɛ=r-e* [rakoho'ɗere]
3-CTFG-INTR-hit=CTFG-IMPERF
'He/she hit.'

b. *∅-r-ʊ-bɛhɛ=r-e* [rubeˈhere]
3-CTFG-INTR-go.down=CTFG-IMPERF
'He/she went down.'

c. *∅-r-a-rɪ=r-e* [raˈrire]
3-CTFG-INTR-leave=CTFG-IMPERF
'He/she was left.'

(8) a. *brɔrɛ-dĩ* [broreˈni] 'cow'
deer-similar.to

b. *bɛdɔ-dĩ* [bedoˈni] 'a type of *filhote*'
filhote (fish sp.)-similar.to

c. *rəkʊ-dĩ* [rəkuˈni] 'watermelon'
gourd-similar.to

d. *hãdɪ-dĩ* [hãɗiˈni] 'a type of *jacu*'
jacu (fish sp.)-similar.to

e. *rɨ-dĩ* [rɨ̞ˈni] 'a type of *cari*'
cari (fish sp.)-similar.to

On the other hand, the vowels /a/, /ã/, /õ/, and /ə̃/ are opaque, systematically blocking harmonization:

(9) a. *∅-r-∅-ɛra̱=r-e* [rɛˈrare]
3-CTFG-INTR-copulate=CTFG-IMPERF
'He copulated.'

b. *r-ɛ-hã̱dɛ=r-e* [rɛhãˈɗere]
CTFG-1+TRANS-hit=CTFG-IMPERF
'I hit (it).'

c. *∅-r-a-kɔhɔdɛ=kõ̱=r-e* [rakɔhɔɗɛˈkõre]
3-CTFG-INTR-hit=NEG=CTFG-IMPERF
'He/she didn't hit.'

d. *r-ɛ-bə̱̃=r-e* [rɛˈmə̃re]
CTFG-1+TRANS-catch/take=CTFG-IMPERF
'I caught (it).'

1.1 Mid *versus* high vowels

As the examples above show, both high and mid [-ATR] vowels are recessive, undergoing vowel harmony. However, there is a crucial difference between both categories: while mid vowels undergo vowel harmony iteratively (7a-b, 8a-b), high vowels undergo harmony non-iteratively. That is, although high vowels do undergo vowel harmony, they fail to transmit the harmonic feature to preceding vowels:

(10) a. *krɔbɪ-dĩ* [krɔbi'ni] 'a type of monkey'
monkey-similar.to

b. *kɔɗʊ-dĩ* [kɔɗu'ni] 'a type of turtle'
turtle-similar.to

c. *hɛkɔɗɨ-dĩ* [hɛkɔɗɨ̞'ni] 'oven'
fire-similar.to

d. *kɔlʊkɔ-dĩ* [kɔluko'ni] 'cedar'
cajá (tree sp.)-similar.to

Also striking is the fact that the behavior of high [-ATR] vowels seems to be sensitive to morphological considerations. While in examples such as the ones above, with the suffix *-dĩ*, high [-ATR] vowels apparently only undergo vowel harmony non-iteratively, examples involving clitics such as the perfective *=re* or the future particle *=kəre* can optionally harmonize iteratively:

(11) a. *∅-r-a-hɨlɔɪ=kəre* [rahɨlɔ'ikre] ~ [rahɨ̞lo'ikre]
3-CTFG-INTR-vomit=FUT
'He/she will vomit.'

b. *∅-r-∅-ɛbʊrɛ=r-e* [rɛbu'rere] ~ [rebu'rere]
3-CTFG-INTR-get.angry=CTFG-IMPERF
'He got angry.'

c. *∅-r-a-brɔɗɨrɛ=kəre* [rabrɔɗɨ̞'rekre] ~ [rabroɗɨ̞'rekre]
3-CTFG-INTR-bròtyre=FUT
'He/she will become *bròtyre*.'[5]

This 'semi-opacity'—that is, the fact that high [-ATR] vowels both *undergo* and *block* vowel harmony—may have interesting theoretical implications for output-oriented frameworks, since it prompts the need for distinguishing

[5] *Bròtyre* is a kind of ceremonial relationship (roughly similar to godparenthood).

underlying [high, +ATR] vowels, which trigger vowel harmony, from derived ones, which block it. This is a question to be further discussed in a future work.

1.2 Domain

As the data presented above suggest, vowel harmony in Karajá is extremely pervasive, applying not only in and across compounds (12), but also across word boundaries (13):

(12) a. *Ø-r-a-ruɛ-ɗə=r-a* [rarue'ɗəra]
3-CTFG-INTR-eye-close=CTFG-PERF
'He/she became blind.'

b. *Ø-r-a-wɛ-bɔhɔ=r-e* [rawebo'hore]
3-CTFG-INTR-belly-break=CTFG-IMPERF
'[They] had their bellies broken.'

(13) a. *hãlɔkɔɛ kɨ̞la* [hãloko,ekɨ̞'dʒa]
jaguar small
'small jaguar'

b. *wa-rikɔrɛ d-u* [waritʃo,re'dʒu]
1-offspring REL-tooth
'my child's tooth'

The domain of vowel harmony seems to be the phonological word, characterized by a single primary stress. As the example below shows, vowel harmony does not seem to apply across phonological words (14). Although stress seems to be useful in determining the domain of vowel harmony, it is irrelevant in characterizing triggers, since, as we have seen, vowel harmony can be triggered not only by stems and derivational suffixes (which are intrinsically tonic), but by clitics (which are intrinsically unstressed) as well.[6]

---

[6] Vowel harmony languages commonly present disharmonic roots—mostly loanwords that 'refuse' to follow the harmonic pattern of the borrowing language. As of yet, I have not found any example of disharmonic roots in Karajá. Potential sources of disharmony seem to be systematically 'fixed up', as illustrated by the examples below. In the likely source of these loanwords (the dialects of Portuguese spoken around the Karajá territory), alveolar stops are palatalized when followed by the high front vowel [i]. As I have shown elsewhere (Ribeiro 2000: 86-88), palatal consonants in Karajá only occur in contiguity to [high, +ATR] vowels. Therefore, the loanwords below pose a conflict to Karajá phonotactic patterns, since a syllable containing an alveopalatal fricative is preceded by a [-ATR] vowel. The conflict is solved by substituting the alveolar implosive /ɗ/ for the original alveopalatal consonant:

i. *kaɗĩbɛɗɪ* [kanibɛ'ɗɪ] 'pocket knife'
(from Portuguese *canivete* [kani'vɛtʃɪ])

(14) *∅-r-∅-ɛlɛ=kɛ* *∅-r-∅-ɛlɛ=kɛ=le* [rɛˈlɛkɛ reˌlekeˈle]
3-CTFG-INTR-become=POT 3-CTFG-INTR-become=POT=EMPH
'[He] was in the process of becoming [a dolphin].'

2. Directionality

As we have seen, the fact that [-ATR] vowels can follow, but not precede [+ATR] vowels clearly shows that vowel harmony in Karajá is strictly a right-to-left process. This is further illustrated by the examples below, involving the stems *buɗɛ* 'few, little', *rikɔrɛ* 'offspring', *ɗuhɔ* 'to curse', and *kɨθɛ* 'grassy'. Since these stems contain both dominant and recessive vowels, they can either trigger (a) or undergo (b) vowel harmony:

(15) a. *bədɛ-buɗɛ* [bəˌdebuˈɗɛ]
land-few
'island'

b. *i-buɗɛ=r-e* [ibuˈɗere]
3-few=CTFG-IMPERF
'It is little.'

(16) a. *wa-θɛ-rikɔrɛ* [waθeritʃɔˈrɛ]
1-mother-offspring
'my sibling'

b. *wa-rikɔrɛ boho* [waritʃoˌreboˈho]
1-offspring PLURAL
'my children'

(17) a. *∅-r-ɔ-ɗuhɔ=rɛrɪ* [rotʃuˈhɔrɛrɪ]
3-CTFG-ANTI-curse=CTFG-PROGR
'He is cursing.'

b. *∅-r-ɔ-ɗuhɔ=r-e* [rotʃuˈhore]
3-CTFG-ANTI-curse=CTFG-IMPERF
'He cursed.'

---

ii. *bɔɗɪ* [bɔˈɗɪ] 'pot'
(from Portuguese *pote* [ˈpɔtʃɪ])

(18) a. *bədɛ kɨθɛ* [bəˌdekɨˈʃɛ]
land grassy
'grassy land'

b. *i-kɨθɛ=r-e* [ikɨˈʃere]
3-grassy=CTFG-IMPERF
'It is grassy.'

As mentioned above, examples such as these, in which directionality is clearly at play, pose an interesting challenge to theories which reject directionality as an independent parameter of assimilation, such as the one proposed by Bakovič (2000). Bakovič claims that "agreement constraints are *left-right symmetrical*" (p. 6), and that directionality is in fact an epiphenomenon derived from morphological considerations. This claim seems to be rather plausible in the cases of languages presenting stem-controlled vowel harmony. As he states, the majority of languages with vowel harmony (such as Turkish and Hungarian) are strictly suffixing, and present *stem-controlled* vowel harmony. Thus, despite the appearances that vowel harmony in these languages is unidirectional, left-to-right, this directionality would be merely a consequence of the morphological structure of the language (p. 7). In other vowel harmony languages, such as Yoruba, "morphology is strictly prefixal; the apparent right-to-left directionality of [ATR] harmony is thus a reflection of stem control" (p. 61).

As for dominant-recessive harmony systems, Bakovič's proposal seems to be based on the assumption, tacitly or explicitly stated in the literature on vowel harmony, that dominant-recessive harmony systems are inherently bidirectional. Examples such as *budɛ* 'few' and *rikɔrɛ* 'offspring', presented above, in which directionality is clearly at play, are, according to Bakovič, 'unattested':

> "If dominant-recessive harmony could in principle be unidirectional, then we would expect to find a language in which the recessive vowels on one side of a dominant vowel are affected by harmony, while those on the other side remain unaffected. *Such a pattern is entirely unattested.*" (Bakovič 2000: 8; italics added)

The Karajá data, as we have seen, demonstrate that this is definitely not the case. Such a pattern is actually rather common in Karajá, not only in polymorphemic constructions such as (2), but in tautomorphemic words as well (15-18). Thus, Karajá provides a strong counterexample to such claims, showing that strict directionality can also be found in dominant-recessive vowel harmony systems, constituting in such cases an independent parameter of assimilation.

3. Final remarks

The discussion presented in this paper hopefully shows that Karajá, a Macro-Jê language from Brazil, presents a straightforward case of dominant-recessive vowel harmony with strict right-to-left directionality, *contra* the assumption that such systems would be always bidirectional (Bakovič 2000). Rather than describing a universal state of affairs, such an assumption probably reflects the fact that our understanding of vowel harmony systems is drawn mainly from a limited sample of languages, a number of which are genetically or geographically related. As van der Hulst and van de Weijer (1995) remind us, "our knowledge of the structure and classification of harmony systems is still extremely limited." As linguistic research expands to comprise lesser-known language areas, such as South America, a more complete picture starts to emerge, revealing otherwise 'unattested' patterns.

References

Archangeli, Diana & Douglas Pulleyblank. 1989. Yoruba vowel harmony. *Linguistic Inquiry* **20.2**: 173-217.

Bakovic, Eric. 2000. Harmony, dominance, and control. Doctoral dissertation. Rutgers University.

Beckman, Jill. 1995. Shona height harmony: markedness and positional identity. *UMOP 18*, 53-76.

Beckman, Jill. 1997. Positional faithfulness, positional neutralization, and Shona vowel harmony. *Phonology* **14**: 1-46.

Beckman, Jill. 1998. Positional faithfulness. Doctoral dissertation, University of Massachusetts, Amherst.

Cavalcante, Marita Pôrto. 1992. Fonologia do Karajá. *Revista do Museu Antropológico* **1**. Goiânia: Editora da UFG.

Fortune, David & Fortune, Gretchen. 1963. *The Phonemes of the Karajá Language* (manuscript). Rio de Janeiro: Arquivo Lingüístico do Museu Nacional.

Fulop, Sean A., E. Kari, & P. Ladefoged. 1998. An acoustic study of the tongue root contrast in Degema vowels. *Phonetica* **55**: 80-98.

Hulst, Harry van der & Jeroen van der Weijer. 1995. Vowel harmony. In Goldsmith, John (ed.) *Handbook of Phonological Theory*, pp. 495-534. Cambridge, MA: Blackwell.

Lombardi, Linda. 1996. Restrictions on direction of voicing assimilation: an OT account. *University of Maryland Working Papers in Linguistics* **4**: 84-102.

Lombardi, Linda. 1999. Positional faithfulness and voicing assimilation in Optimality Theory. *NLLT 17*, 267-302.

Noske, Manuela. 1995. The ternary use of distinctive features: In defense of the number three. PhD dissertation, Chicago: University of Chicago.

Ribeiro, Eduardo R. 2000. [ATR] vowel harmony and palatalization in Karajá. *Santa Barbara Papers in Linguistics* 10 (Proceedings from the 3rd Workshop on American Indigenous Languages), pp. 80-92. Santa Barbara: UCSB.

Ribeiro, Eduardo R. 2001. Female versus male speech in Karajá. Paper presented at the meeting of the Society for the Study of the Indigenous Languages of the Americas/4th Workshop on American Indigenous Languages. Santa Barbara, July 6, 2001.

Rigsby, Bruce & Michael Silverstein. 1969. Nez Perce vowels and Proto-Sahaptian vowel harmony. *Language* **45.1**: 45-59.

Eduardo Rivail Ribeiro
University of Chicago
Department of Linguistics
1010 E. 59th Street
Chicago, IL 60637

erribeir@midway.uchicago.edu

The Phonetics and Phonology of Unreleased Stops in Karitiana

LUCIANA STORTO[i] and DIDIER DEMOLIN[ii]
[i]*Museu de Arqueologia e Etnologia/Universidade de São Paulo and Phonology*
[ii]*Laboratory, Université Libre de Bruxelles*

0. Introduction

Karitiana, an endangered language from the Tupi stock, Arikem family, spoken in the state of Rondonia in Brazil, exhibits a number of interesting phenomena concerning stop consonants at the end of words. In particular, voiceless stops are always unreleased word-finally. This phenomenon has been observed as an occasional variant of stops in English (Laver 1994, Ladefoged & Maddieson 1996) and in a number of other languages such as Bamileke (Westermann & Ward 1952) and Efik (Cook 1969), and is systematic in Karitiana. Unreleased final stops are also common in Tupi languages and have been described by several researchers (e.g. Moore 1984; Galucio 1994, 1996; Gabas Jr. 1998, 1999; Picanço 1999; Storto 1999; Rose 2000). Such facts have rarely been described in a detailed manner. This paper describes some phonetic and phonological aspects of the nature of these sounds when they appear word finally. More specifically, three points will be examined: (i) voiceless stop consonants (there are no phonemic voiced consonants in Karitiana); (ii) nasal consonants, which have been described as unreleased word finally by Storto (1999), and (iii) the phonological behavior of these consonants in context.

1. The Phonetics of Word Final Stop Consonants

1.1 Materials

Data were collected in Brazil with 4 male subjects. Measurements were made to record acoustic, aerodynamic and articulatory data. All subjects contributed to all of the different modes of recording and measurement. Acoustic recordings were made with a Shure headset microphone, which enables the maintenance of a constant distance between the subject's lips and the microphone. This method allows data to be recorded in a way which is both comfortable for the speaker and prevents amplitude variations due to head movement during recording. Acoustic and aerodynamic data were recorded with a portable MacQuirer workstation and have been processed with the Signal Explorer and Formants software. The recordings were digitized at 20 kHz. Video data were recorded with a Sony DV camera which captures images and acoustic data simultaneously. Finally, palatograms and linguograms were made using stimuli containing intervocalic and word-final alveolar nasals. Data are based on Storto's 1999 description of Karitiana phonology. The stimuli used for the experiments are shown in (1) and (2) below.

(1) Karitiana words used for the acoustic and aerodynamic recordings

Stops		Nasals	
gep˺	'lice'	ŋãm˺	'rotten'
gop˺	'wasp'	nõm˺	'breasts'
gɨp˺	'termite'	mãm˺	'to tighten'
ʔot˺	'to fall'	mãn˺	'husband'
ʔit˺	'son' (father speaking)	põn˺	'to play'
ʔet˺	'son' (mother speaking)	pasẽn˺	'cricket'
ãndik˺	'buttocks'	sõŋ˺	'firewood'
tʃak˺	'to bite'	sɨ̃ŋ˺	'to grind'
apibmbik˺	'to push'	ãmãŋ˺	'to plant'
põm˺	'to play'	pat˺	'macaw'
kat˺	'to sleep'	õmp˺	'tadpole'

(2) Karitiana words used for the articulatory recordings

Palato/linguogram		Video	
pat˺	'macaw'	mẽm˺	'to enter'
ota	'friend'	sõm˺	'red'
mãn˺	'husband'	opo	'penis'
õnɨ̃	'that'	sop˺	'hair'

1.2 Methods

Acoustic data were recorded in two sessions: the first time separately from the aerodynamic measurements, the second time in tandem with those measurements. In each recording session, subjects were asked to repeat each of the words in (1) multiple times: once in isolation and three times in a short carrier sentence, *Karitiana haadna pip X nakaat Y* 'In Karitiana X is Y' (where Y is the gloss of X in Portuguese). Length measurements were made on the audio waveform using a spectrogram. The duration of inter-vocalic nasals was measured starting from the last cycle of the preceding vowel to the first cycle of the following vowel. When in initial position, this was done from the first observable cycle of voicing to the first cycle of the following vowel. The duration of final nasals was measured starting from the last cycle of the preceding vowel to the last observable cycle.

The aerodynamic data consist of pharyngeal pressure, nasal airflow, and oral airflow measurements. Oral airflow measurements were taken with a small flexible silicone mask placed against the mouth. Nasal airflow was measured with a nasal mask set around the nose. Pharyngeal pressure was recorded with a small flexible plastic tube (ID 2mm) inserted through the nasal cavity into the oro-pharynx for one subject. One session involved the recording of acoustic and aerodynamic parameters simultaneously and another session recorded only oral and nasal airflow. In both cases, the microphone was placed next to the mask used to record oral airflow.

As the three other subjects did not tolerate the tube used to make pharyngeal pressure measurements, pressure measurements were only made for labial consonants with those subjects. This was done by asking the subjects to hold a plastic tube (ID 5mm) between the lips.

Articulatory data were also collected in two different sessions. The first involved simultaneous recording of face and profile images using a digital video camera. Profile images were obtained by putting a mirror against the cheek of the speaker, at a 45° angle to the sagittal plane. These data were intended to study the relative timing of lip movements during the realization of the sounds [m˺, n˺, ŋ˺]. The camera was used in order to allow for comparisons to be made with non-final (released) nasals. The audio signal was recorded simultaneously by the camera. The second session involved the use of the "classical" method for making palatograms and linguograms (Dart 1991). Subjects were asked to pronounce a word containing the consonant under study after their tongues or palates had been painted with an equal mixture of charcoal and olive oil. In the case of palatograms, a mirror was then inserted into the mouth and the print on the hard palate was photographed with the video camera. For linguograms, subjects were asked to pronounce a word after their hard palates were painted with the charcoal/olive mixture. The print on the tongue was then photographed with a video camera.

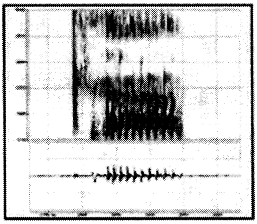
2. Results

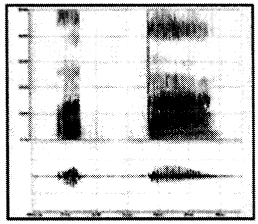
2.1 Acoustics

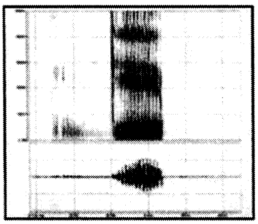
2.1.1 Stops

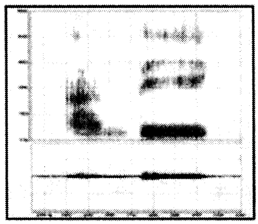
Results from the acoustic measurements, presented in (3a-d) show that formant transitions always identify the place of articulation of the consonant. Analysis of the corresponding audio waveform shows that these consonants are indeed stops with no bursts. There is no burst at the end of the spectrogram of [kat˺] 'to sleep' in (3a). Starting from the velar transition, there is a lowering movement of F2 towards the locus of the alveolar consonant and a rising movement of F3 towards the locus estimate (2500 – 2700 Hz). Figure (3b) shows that when no consonant follows the final vowel [a], there are no observable formant transitions. Formant transitions at the end of the spectrogram of [gep˺] 'lice' in (3c) show a lowering of F2 and F3 towards their loci estimates (1100 – 1500 Hz) and (2200 – 2400 Hz). Formant transitions at the end of the word [ãndiːk˺] 'cold' in (3d) show the convergence of F2 and F3, as expected for velar stops.

(3) Spectrograms and waveforms of [kat˺] 'to sleep', [ota] 'friend', [gep˺] 'lice', and [ãndi:k˺] 'cold'


(a) [k a t˺]


(b) [o t a]


(c) [g e p˺]


(d) [ã n d i: k˺]

2.1.2 Nasals

Acoustically, the main feature differentiating a non-final nasal, which is in fact a post-stopped nasal (see Storto & Demolin (2002) for details on post-stopped nasals), from an unreleased final nasal is duration. (Post-stopped nasals are transcribed as [m^b, n^d, ŋ^g].) Table (4) shows that the average duration of unreleased final nasals is greater when compared to the non-final nasals.

(4) Average duration in ms of non-final nasals and of final unreleased nasals (n=54 for each speaker).

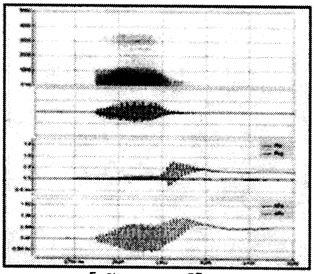
Non-final nasals	ms	Final unreleased	ms
m^b	191	m˺	221
n^d	172	n˺	255
ŋ^g	120	ŋ˺	176

2.2 Aerodynamics

2.2.1 Stops

Results of aerodynamic measures show that when the closure is maintained until well after the word, as for an unreleased stop, pharyngeal pressure dissipates gradually. Sometimes an increase in nasal flow can be observed at the beginning of the final consonant, as can be seen in (5).

(5) Spectrogram, waveform, pharyngeal pressure and nasal airflow for [õmp˺] 'tadpole'



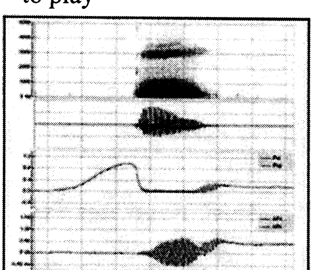
[õ m p˺]

When an unreleased stop is followed by a word starting with a voiceless stop consonant in a compound form or in a sentence, a burst is produced because pressure is released sharply. Sometimes, but much less frequently, the unreleased stop assimilates completely to the following stop without any observable drop in pressure (see Sections 3.1 and 3.2 for more details on this phenomenon). In context, a clear drop in pharyngeal pressure can sometimes be observed between two consecutive voiceless stops, the first being the final consonant of a word and the second the initial consonant of the following word. This drop in pressure accounts for the observed burst, but it should be noted that it marks a pause between two words and not a complex consonant cluster.

2.2.2 Nasals

It can be observed in (6) that an increase in nasal airflow starts when the oral airflow stops. Note that nasal airflow is maintained well after the nasal voicing ceases. The maintenance of nasal airflow and also of lip closure accounts for the unreleased nature of these nasals, which differ from post-stopped nasals that are realized with a clear burst when the lip closure is released (see Storto & Demolin (2002) for a detailed description of post-stopped nasals).

(6) Spectrogram, waveform, pharyngeal pressure, and nasal airflow for [põm˺] 'to play'



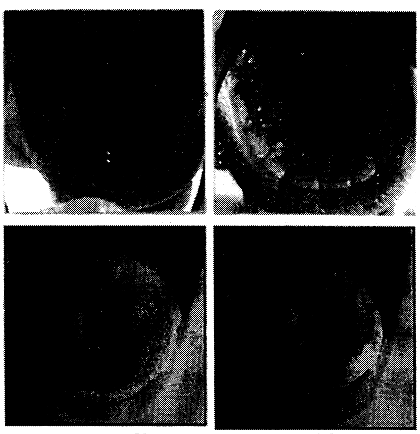
[p õ m˺]

2.3 Articulatory Data

2.3.1 Palatograms and Linguograms

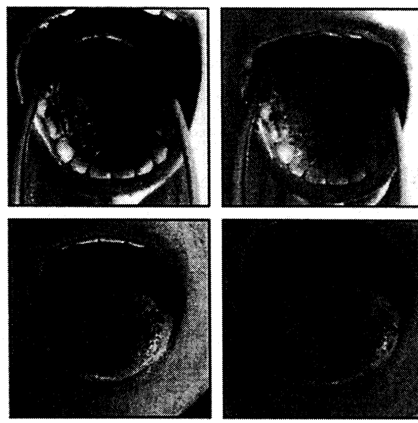
The main observations that can be made from the palatograms and linguograms realized with the speakers of this study is that when a stop is unreleased word finally it has a wider tongue contact compared to the same stop in intervocalic or initial positions. This can be seen in (7), which shows palatograms and linguograms of the words [pat˺] 'macaw' and [ota] 'friend' for one of the speakers. In addition to a wider contact, unreleased stops also have a more fronted articulation. Figure (8) shows a comparison between an intervocalic and a word-final nasal in the words [mãn˺] 'husband' and [õnɨ̃] 'that'. As for the stops, it can be seen that word final nasals have a wider and more fronted tongue contact. The same patterns of contact have been observed with the other speakers of the study.

(7) Palatograms and linguograms of the words [pat˺] and [ota]



[t˺] [t]

(8) Palatograms and linguograms of the words [mãn˺] and [õnɨ̃]

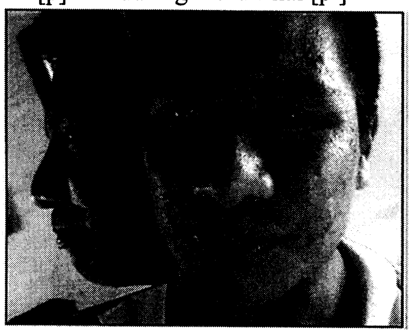


[n˺] [n]

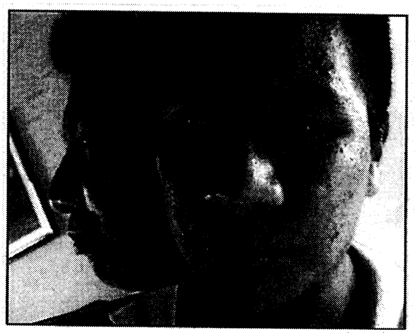
2.3.2 Video

Video data provide a good way to observe the behavior of labial consonants. Figure (9) compares lip closure during an intervocalic stop [p] in the word [opo] 'penis' and in a final [p˺] in the word [sop˺] 'hair'. The image for the intervocalic [p] is taken 20 ms before lip opening, whereas the image of the final unreleased [p˺] is taken 200 ms after lip closure. Comparison of the images suggests firmer closure with the final unreleased stop. This firmer closure accounts for the continuation of lip closure necessary to produce an unreleased stop. Figure (10) compares lip posture during an initial post-stopped nasal [mᵇ] and an unreleased final nasal [m˺] in the word [mᵇẽm˺] 'to enter'. The images are taken 20ms before lip opening for [mᵇ] and 200 ms after lip closure for [m˺]. Examination of the images suggests a firmer closure with the final unreleased nasal.

(9) Lip closure during intervocalic [p] and during word final [p˺]

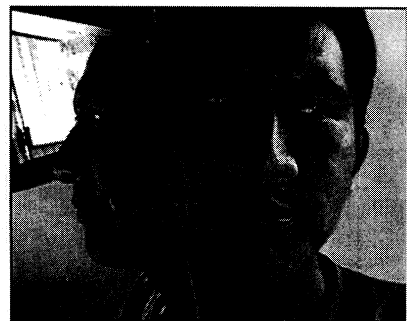


[p]



[p˺]

(10) Lip closure during intervocalic [mᵇ] and during word final [m˺]



[mᵇ]

[m˺]

3. Phonology

When they are not followed by a pause or at the end of a sentence, unreleased stops become voiced if followed by a voiced segment. If a voiceless stop consonant follows, the unreleased stop is exploded. In a few cases it has been observed that an unreleased stop assimilates completely to the following voiceless consonant, which geminates. These cases are described in the following sections.

3.1 Progressive Assimilation of a Nasal to Place Features of an Oral

The examples given below in (12), (13) and (14) show examples of the assimilation of the nasal place features to the preceding oral. For example in (12), the word [ʔednã] 'pregnant' is realized when the final unreleased stop [t˺] assimilates to the voicing features of the following nasal when a suffix [nã] (adjectivizer) is added to the word [ʔet˺] 'child'. This can be summarized by (11) which states that an unreleased voiceless stop becomes voiced when followed by a nasal.

(11) Voicing Rule
C˺ [-voice] > C [+voice] /__ N

Note that two possible ways have been observed to produce what is interpreted as voicing assimilation for [t˺]. In the first there is clear acoustic evidence of voicing whereas in the second voicing does not appear in the acoustic signal but voicing is perceived probably due to the short duration of the inter-segmental period.

(12) ʔet˺ + nã > ʔed nã 'pregnant'
child adjectiviser

(13) ʔop˺ + nã > ʔobmã 'pierced'
hole adjectiviser

(14) ãndik˺ + nã > andigŋã 'with buttocks'
buttocks adjectiviser

3.2 Vowel epenthesis

Vowel epenthesis occurs between two adjacent underlying consonants in the following environments: 1) in underived environments; 2) lexically, between a root and an affix; 3) post-lexically, between two words inside a compound or clitic-host unit. The quality of the epenthetic vowel is the same as that of the preceding underlying vowel.

(15) Epenthesis Rule
$Ø \rightarrow V_1$ / V_1C_C

(16) Examples of epenthesis in environment 1
/kɨrk/ → [kɨrik˺] 'mosquito'
/pɨkp/ → [pɨkɨp˺] 'tree bark'
/ept/ → [eβet˺] 'thin'

(17) Examples of epenthesis in environment 2
bik + pa → bi.ki.pa 'seat, bench'
to sit nominalizer

pi.hop + pa → pi.hopo.pa 'drier'
to dry nominalizer

ko:kot + pa → ko:kotopa 'bridge'
to pass nominalizer

In post-lexical epenthesis, we have one additional condition in the environment of the rule—the epenthetic vowel must separate a stress clash.

(18) Ø → V / 'V.C__CV

(19) Examples of post-lexical epenthesis in environment 3
ʔep sap → e.pe sap 'tree leaf'
tree flat.object

ʔep kɨ̃ⁿɨmbɨ̃ → ʔepkɨ̃mbɨ̃ 'tree root'
tree root

Note that epenthesis occurs lexically in ko:kotopa, but not post-lexically in ʔepkɨ̃mbɨ̃, although, in both cases, no stress clash takes place.

In a few cases it has been observed that an unreleased stop assimilates completely to the following consonant, which geminates. Such an example is given in (20).

(20) Karitiana haadna pip˺ tʃak˺...*morder* 'In Karitiana *tʃak˺* is *to bite*'
[...pitʃ:ak˺...#]

4. Syllable structure

One important thing to note is that when assimilation or epenthesis takes place, resyllabification take place too. This is because in this language a non-nasal in root-final position triggers epenthesis whereas a nasal does not, as shown in (21) and (22) (see Storto (1999) for more details on prohibited codas and obligatory onsets).

(21) bik + pa → bi.ki.pa 'seat, bench'
to sit nominalizer

(22) hõ.rõn + pa → hõ.rõn.pa 'basin'
to wash nominalizer

5. Conclusion

In this paper it has been shown that unreleased stops are identified by formant transitions and that these consonants are produced by sustaining the stop closure for a longer period compared to an exploded stop. This affects pharyngeal pressure, which decreases gradually. In the case of nasals, nasal airflow perseveres longer for final nasals than for initial or intervocalic nasals. The tongue contact is wider and more fronted for unreleased stops and nasals. In context, unreleased stops assimilate or are exploded and resyllabify as the onset of the following syllable.

Acknowledgments

We wish to thank Luiz Karitiana, Nelson Karitiana, Ignacio Karitiana, and João Karitiana for their patience and their collaboration during the experiments. Our colleagues of the Tupi comparative project contributed to fruitful discussions on various aspects of this work. Data were collected during several fieldwork trips in Brazil.

Part of this research has been supported by the Wenner-Gren Foundation, by the ARC convention 98-02 n°226 of the Belgian Ministry of Scientific Research, and by two Brazilian research funding agencies: FAPESP (Post-Doctoral grant 01/00172-9) and CNPq (PROFIX grant 540532/01-0 NV).

References

Cook, T.L. 1969. Efik. In E. Dunstan (ed.) *Twelve Nigerian Languages,* 35-46. London: Longmans Green.

Dart, S. 1999. Articulatory and acoustic properties of apical and laminal articulations. *UCLA Working Papers in Phonetics*, **79**. University of California Los Angeles.

Gabas Junior, N. 1998. *Estudo fonologico da lingua Karo*. Munich: Lincom Europa.

Gabas Junior, N. 1999. *A* Grammar of Karo, Tupi (Brazil). PhD Dissertation, University of California, Santa Barbara.

Galucio, A.V. 1994, Fonologica segmental da lingua Mekens. IX Encontro Nacional da ANPOLL. Caxambu, MG. *Anais do IX Encontro Nacional da ANPOLL*, Vol. 1. 988-998.

Galucio, A.V. 1996. *Mekens phonology*. Ms. University of Chicago.

Ladefoged, P. & I. Maddieson. 1996. *The Sounds of the World's Languages*. Oxford: Blackwell.

Laver, J. 1994. *Principles of Phonetics*. Cambridge: Cambridge University Press.

Moore, D.A. 1984. Syntax of the language of the Gavião indians of Rondonia, Brazil. PhD Dissertation, City University of New York.

Picanço, G. 1999. Estudo preliminar da fonologia da lingua Munduruku. Senior thesis, Universitade Federal do Para, Belém.

Rose, F. 2000. Elements de phonétique, phonologie et morphophonologie de l'Emerillon (Teko). *Mémoire de DEA de Sciences du Langage*. Université de Lyon II.

Storto, L. 1999. Aspects of Karitiana Grammar. PhD dissertation, MIT.

Storto, L. & D. Demolin 2002. Control and timing of articulatory gestures in pre- and post-oralized nasals in Karitiana. Paper presented at the 8th Conference of Laboratory Phonology, Haskins Laboratories, Yale University, June 27-30.

Westermann, D. & M.A. Bryan 1952. *Handbook of African languages, Part II, Languages of West Africa.* London: Oxford University Press.

Luciana Storto
MAE, Universidade de São Paulo
Av. Prof. Almeida Prado, 1466
CEP 05508-900 São Paulo, Brazil

luciana_storto@yahoo.com

Didier Demolin
ULB, Phonology lab., CP 175
50 av. F. Roosevelt
1050 Brussels, Belgium

ddemoli@ulb.ac.be

Before You Record Anything Else: Topics and Questions to Consider When You Interview a Speaker for the First Time

SUZANNE WASH
University of California, Davis

0. Introduction

In recent years there has been a renewed interest in field linguistics and language documentation. One reflection of this trend is the appearance of several relatively new books about doing linguistic fieldwork (e.g. Newman and Ratliff 2001, Vaux and Cooper 1999, Payne 1997). Yet, there remains one area of fieldwork methodology that is rarely dealt with in publications and linguistic field methods courses generally: how and why a linguist should gather information about the life history of the speaker, including, especially, information about the speaker's sociolinguistic background. Such information is critical for linguists investigating moribund languages, where it is not uncommon to find significant differences in the speech of speakers who belong to the same family, let alone the same speech community. The reasons behind such differences become clearer if the linguist first determines what sociolinguistic factors, in the speaker's personal history, have affected that speaker's knowledge and use of the language in question.

To deal with this area of inquiry, I developed the oral questionnaire given in the following pages of this article. With the exception of section 1.2.4., which is a relatively recent addition, I have used this questionnaire in my fieldwork with some of the last speakers of Northern Sierra Miwok, a moribund California Indian language. This language was formerly spoken in an area that now roughly includes the counties of Amador and Calaveras in California.

I first started doing fieldwork among speakers of Northern Sierra Miwok in 1992, when I was a graduate student in linguistics at the University of California, Santa Barbara. I did not become aware of the importance of recording sociolinguistic information about speakers until 1995, when I started working as a research assistant for the Corpus of Spoken American English, headed by John W. Du Bois. Under this project, researchers were sent out across the United States to record naturally spoken American English. The researcher who made the recording would collect information about each speaker's background on a brief, speaker information form. Among other areas of inquiry, this form asked about the speaker's native language(s), including the speaker's first and/or dominant language, age, occupation, education, religion, language used during school, residence history, the speaker's parents' place of origin, and parents' native language(s). I realized that if such background information were important for research on a world language like English, how much more important must it be for research on an endangered language like Northern Sierra Miwok? The resulting oral questionnaire is thus inspired from Du Bois' speaker information form, but it is modified and greatly expanded for working with speakers of Northern Sierra Miwok. The format is general and flexible enough, however, that it can be easily customized for working with speakers of any language.

This oral questionnaire will help you obtain critical information about the speaker's sociolinguistic background and history of language use *before* you begin detailed investigations of the language in question. For example, you will find out what languages the speaker knows or has otherwise been exposed to, the contexts in which the speaker learned and used these languages, and the extent to which the speaker has used these languages. You will gain an insight into the differences perceived by the speaker between his or her language and the languages of nearby communities which the speaker may know. You can obtain a better understanding of the speaker's attitudes about speaking particular languages, and what factors shaped those attitudes. For moribund languages, the questionnaire helps you obtain information about the number of other speakers there may be and their degree of fluency in those languages. Finally, in cases where members of the same family show unexpected differences in the way they speak a given language, this questionnaire will help you discover the reasons for such differences. In one particular family whose members speak Northern Sierra Miwok, I found that the pronominal paradigm for one of the speakers showed unexpected similarities with the pronominal paradigm of Central Sierra Miwok, a closely related language. The similarities stem in good part from the fact that this particular speaker spent a significant amount of his childhood living with his paternal grandmother, who spoke to him in Central Sierra Miwok. His parents did not speak to him in that language. Without this questionnaire, I might not have ever found out these particular but crucial facts about his history of language use.

1. The Oral Questionnaire

1.1. Preliminaries: What to say when you first start the recording

When you first start the recording, it is a good idea to state the date, the name of the speakers(s) you're interviewing, your name, and the place of the interview. Regarding the place of the interview, include particulars such as the town and state or province in which it is located. State also whether the interview is taking place in the speaker's home, or in a public place such as a park. Mention the name of anyone who is assisting you in the recording process, such as the person attending to the video camera. Are there listeners in attendance? If so, mention this fact, and identify them on the recording if circumstances allow it.

State the number of the tape or disk that you are using for that day's recording session. For example, if you're starting on the first tape of that day's session, say into the tape recorder *This is tape number one*. Likewise, say *This is tape number three* if you have already recorded two tapes during that same interviewing session and you are starting on the third tape.

Record the above preliminary information every time you begin a new tape and/or recording session. Doing so will save you from some hassles later on, especially if the tapes or disks are unlabeled and get stacked out of sequence in your equipment bag.

1.2. Questions about the speaker's age, residence- and work history, education, and religion

1.2.1. Age and residence history

(1) When were you born?

(2) Where were you born?

(3) Could you tell me the places you have lived throughout your life, about how long you lived in each place, and who you lived with during these times? Start with your place of birth. (Note: Make a list of the people the speaker has lived with, because you will later want to ask about the languages that the speaker has used with each of these people.)

1.2.2. Formal schooling and language use

As is well known, formal schooling had a great, negative impact on the lives and native languages of many Native Americans, especially those who were forced to attend boarding schools in earlier times. The questions in this section were aimed at discovering what the formal schooling experience was like for the Northern Sierra Miwok speakers I interviewed, with a special focus on language use.

(4) Did you attend school in your youth?

(5) At what age did you begin going to school?

(6) At what age did you stop going to school?

(7) The school(s) you attended, were they located near your home, or were they boarding schools located some distance from your home? (Note approximate locations.)

(8) Did you live at home while going to school?

(9) What was/were the language(s) of instruction in the classroom?

(10) What was the ethnic make-up of your classmates? Were they mostly Indian (if so, from what communities), white, or other (e.g. Chinese or Mexican)?

(11) Did your classmates speak languages other than English?

(12) Did they speak languages other than English inside the classroom?

(13) Did they speak languages other than English outside the classroom, e.g. on the playground?

(14) Did you ever talk in Indian while at school to your classmates or to your teacher? (Note: The Miwok elders I have worked with often used the term *Indian* to refer to the Northern Sierra Miwok language, so that is the term I use in question (14) and elsewhere in this questionnaire.)

(15) Did your teacher and classmates know you were Indian?

(16) Were you ever punished or otherwise discouraged from speaking in Indian?

(17) Did your teacher or classmates ever encourage you to speak in Indian, or teach them some of the language?

(18) After you left school, would you say you spoke in Indian with your family just as much as you did before you entered school as a child, or did you talk in the language much less often than before? Would you say the same was true for your friends or other members of your family who attended school?

1.2.3. Work history and language use

(19) What kinds of work have you done throughout your life, starting with the work or chores you would do with your family and relatives as a young man/woman?

(20) Did your bosses and co-workers know you were Indian?

(21) Did you ever speak with your bosses and co-workers in Indian? If yes, did you do so while working, or did you do so only after hours?

(22) Did any of your non-Indian bosses/co-workers understand the language?

(23) Were there only certain times or situations that you would speak in Indian at work? (Possible examples of such situations include greeting friends, joking around, arguing, or wanting to say something without other people understanding, etc.) Or, would you converse in Indian as often as you would in English? What kinds of things would you say or talk about in Indian?

(24) Did people at your work place ever encourage you to speak in Indian or teach them some of the language?

(25) Did anyone ever ask you *not* to speak in Indian?

(26) What other languages did you use at work?

(27) Where there some languages you heard but didn't speak yourself?

1.2.4. Religion and language use

This part of the questionnaire is a relatively recent addition. I added it because a growing number of Native Americans, whose language of heritage is no longer spoken, nonetheless desire to pray in their language of heritage. Therefore I believe it is an area of inquiry that a linguist should cover if at all possible. To my great regret I neglected to ask my most fluent Northern Sierra Miwok consultants, who are now deceased, about this very important aspect of language use.

(28) Did your family practice any particular faith or religion while you were growing up?

(29) Do you practice the same faith/religious traditions today?

(30) What language did you hear people use in ceremonies, prayers, or blessings?

(31) Is there any particular language you or your family members prefer to use when you pray?

(32) About how often do you pray in this language?

(33) Have you yourself ever said prayers in public or in group gatherings? In what language did you say these prayers?

(34) Would you be interested in recording some blessings and prayers, ones you say and ones you have heard others say?

If the speaker answers yes, then the following are some suggestions that might help the speaker remember specific blessings or prayers, for either now or in a future recording session:

- Prayers expressing thanks
- Prayers or blessings appropriate at Indian Big Times/Powwows
- Blessings or prayers during opening/closing ceremonies
- Blessing for a new baby or addition to the family
- Blessing a home and the land the home is on
- Blessings/prayers at coming-of-age ceremonies or other rites of passage, such as graduations
- Blessings/prayers at weddings
- Prayers said at funerals or memorial services
- Prayers for spirits whose burial place has been disturbed
- Prayers to restore harmony and peace
- Prayers for guidance and wisdom
- Prayers for healing
- Prayers used in addressing a plant or tree that you wish to harvest from or cut down
- Prayers used in addressing a river, or the ocean, before you take fish or other food from it, or before you traverse it
- Prayers used for more/less rain or snow
- Prayers used before any undertaking, such as a long journey, a hunt, building a house, etc.
- Translations of other blessings or prayers, such as the Lord's Prayer

1.3. Questions about the speaker's mother

(35) Speaker's mother's name:

(36) Her date of birth:

(37) Her place of birth:

(38) Where did your mother live throughout her life?

(39) What languages did/does she speak?

(40) About how often did you hear her speak these languages with other people?

(41) Were there any languages she could understand but did not speak herself?

(42) What language(s) did/does she speak with you?

(43) How often (i.e. daily/weekly/occasionally)?

(44) When (what kinds of contexts) would she most often talk in Northern Sierra Miwok (or in other indigenous languages) with you? (Possible examples: while telling you to do chores, telling you stories, talking with you during the preparation of meals, during mealtimes, scolding you, joking around, gathering basketry materials, making baskets, or around people who didn't know the language so that she could say something to you without others listening in, etc.)

a. During your childhood:
b. During your youth/teen years:
c. During your adult years:

(45) When (what kinds of contexts) would she most likely speak in English (or Dominant Language X) with you?

a. During your childhood:
b. During your youth/teen years:
c. During your adult years:

(46) Did she ever discourage you from speaking in Indian or in English?

(47) In your best estimation, what were her thoughts and attitudes about speaking in Indian versus in English?

1.4. Questions about the speaker's father

Ask the same questions given in section 1.3., but for the speaker's father.

1.5. Questions about the speaker's grandparents

Ask the same questions given in section 1.3., but for the speaker's grandparents.

1.6. Questions about the speaker's aunts, uncles, and acquaintances of his/her parents' generation

Ask the same questions given in section 1.3., but for the speaker's aunts, uncles, and others of the speaker's parents' generation whom the speaker has talked with or listened to.

1.7. Questions about the speaker's siblings, cousins and acquaintances of his/her generation

Ask the same questions given in section 1.3., but for the speaker's siblings, cousins, and friends or acquaintances of the speaker's generation.

1.8. Questions about the speaker's spouse(s) and in-laws

Ask the same questions given in section 1.3., but for the speaker's current spouse, former spouse(s) and in-laws, regarding the extent to which the speaker spoke with these people in the languages in question.

1.9. Questions regarding the speaker's perception of and exposure to languages of nearby communities

The following are examples of questions I have asked of elders who live in or come from Calaveras County, California, and who speak what they themselves call *Calaveras*, a dialect of the Northern Sierra Miwok language. The questions and subsequent answers help me understand more about a speaker's particular dialect and idiolect, and whether that speaker perceives this language to be more or less like another, nearby Miwok language or dialect that the speaker has spoken or heard. You can customize these questions in your own interviews to make them more appropriate to the homeland of the speaker you are interviewing.

(48) Is it easier for you to understand people who speak Tuolumne (Central Sierra Miwok), or is it easier for you to understand people from Amador County (who speak a different dialect of Northern Sierra Miwok), especially those from the Ione area?

(49) Does the Calaveras language strike you as being more like the Tuolumne language, more like the Amador language, or equally like both of them?

(50) In general, do you think you spent more time around Tuolumne-speaking people or Amador-speaking people, or not much time with either group, during your life?

(51) Could you do an imitation of someone speaking in Tuolumne and someone speaking in Amador?

(52) Are there other places where you spoke or heard a Miwok language spoken (e.g. at Big Times, in town, at the health clinic, the grocery store, in a bar, etc.)?

(53) Have you ever heard the Mariposa (Southern Sierra Miwok) language, and can you speak it?

(54) Have you ever spoken or heard any other American Indian languages?

1.10. Questions about the speaker's use of language with his or her children, grandchildren, nieces, nephews, and others of their generation

(55) Do you or have you spoken to children (your own or others) in Indian?

(56) When your children were young, did you talk to them solely in Indian, or in both Indian and English?

(57) Were there particular contexts in which you would speak only in Indian or only in English with them?

(58) Is there anything you remember about how people used to talk to babies and children in the Indian languages you know?

(59) Did the children you spoke with answer you in Indian, or did they answer in English or Spanish?

The speaker's answers to the questions in (60) – (62) will give you a clearer idea of how to make your future recordings as relevant as possible to language-revitalization programs that the speaker's community has or wishes to develop.

(60) Would you be interested in making recordings that could be used to help the younger generation learn your language?

(61) Would you like to teach your language to younger members of your community?

(62) What kinds of things would you like to see happen or not happen, in the future, with regard to your native language(s) and the traditional knowledge and history of your community? What role would you like to have in these efforts? For example, would you like to see these languages taught in the schools? Would you like to see bilingual signs (using your language) throughout town? Would you like to help the general public develop a greater awareness of your language, traditions, history, and community (e.g. through adult education classes for the general public or through language tutorials on the world wide web)?

Other suggestions for the list in (62) could of course include a user-friendly dictionary, illustrated books that have bilingual stories with accompanying exercises, or multimedia materials on CD-ROM. Even if the speakers you work with are not interested in having or creating language teaching materials, their future descendants might well be, especially if the language in question is endangered.

2. Final remarks on using this questionnaire

There are several things to keep in mind when using this oral questionnaire:

First, it can easily take more than one, two-hour recording session, depending upon how much detail the speaker wishes to give in answering the questions. In my own fieldwork on Northern Sierra Miwok, this questionnaire has taken several recording sessions. Since Northern Sierra Miwok is a dying language, I encouraged the elders to feel free to expand upon any of the topics as they saw fit. I also told them to pass on any topic or question that they did not care to talk about.

Second, you can and should customize the questionnaire to fit the languages and experiences of the speakers you are interviewing. For example, if the speaker you are interviewing never had any formal schooling, you can change most of the questions to focus on any other kind of specialized training or apprenticeship the speaker may have had, and his or her experiences with language use therein. Or, to give another example, if you want to ask the speaker some questions about religion and language use (section 1.2.4.), but there is no direct translation for the English words *blessing*, *pray* or *prayer* in that speaker's language, then simply use terms from languages that the speaker knows and that convey the closest meaning to these words, in the context of the speaker's own culture and experiences. A third example of how you could customize this questionnaire concerns the questions about the speaker's relatives and their language use with the speaker. In some cultures, it is inappropriate to say the names of deceased relatives. (Such was true in traditional Miwok societies, though most of the speakers I interviewed no longer followed that custom.) Simply delete the question that asks for the name of anyone who happens to be deceased, in that case, and use an appropriate kinship term, e.g. *maternal grandmother,* as a substitute for the name of the speaker's deceased maternal grandmother, when you need to refer to the deceased person. As for the wording of the questions generally, I encourage anyone who uses this questionnaire to put the questions into his or her own words, in whatever way is more comfortable to say.

Finally, as the title of this paper suggests, I recommend using the topics and questions in this oral questionnaire as a starting point in linguistic fieldwork. The questionnaire allows you to gather, in a systematic fashion, the kind of information that will contextualize the *rest* of the linguistic data that you obtain from the speaker. And, if you work on a language whose most competent speakers are elderly, as is often the case with endangered and moribund languages, you risk losing the opportunity to obtain this crucial information if you wait too long.

References

Newman, Paul & Martha Ratliff (eds.). 2001. *Linguistic fieldwork.* Cambridge University Press.

Payne, Thomas E. 1997. *Describing morphosyntax: A guide for field linguists.* Cambridge University Press.

Vaux, Bert & Justin Cooper. 1999. *Introduction to linguistic field methods.* Munich: Lincom Europa.

Native American Language Center
Department of Native American Studies
University of California, Davis
Davis, CA 95616

swash@ucdavis.edu

Rethinking the Observer's Paradox and Data "Purity"*

SUZANNE WERTHEIM
University of California, Berkeley

0. Introduction

In this paper, I will examine some of the difficulties faced by the linguistic fieldworker who is attempting to observe and record "natural" conversations, and I will reconsider the long-held sociolinguistic notion of the observer's paradox by recasting it within Bell's (1984) framework of audience design theory. Using data gathered during my own fieldwork, I will once again call into question the idea of a single, unmarked, unperformed vernacular, the access to which is supposedly blocked by the observer's paradox. Finally, I will demonstrate that "performed" or "self-conscious" speech produced for the fieldworker can be useful in systematic linguistic analysis, and in gaining insights into local language ideologies and linguistic norms.

The usual goal of sociolinguistic investigation is to gather recordings of "natural" speech, which is to say, somehow accessing and recording what is generally referred to as the vernacular, as "untainted" by interactions with the fieldworker as possible. Although specific definitions vary, among sociolinguists it is generally agreed that the vernacular is "the relatively homogenous, spontaneous speech reserved for intimate or casual situations...taken to reflect the most systematic form of the language acquired by the speaker, prior to any subsequent efforts at (hyper-) correction or style shifting..." (Poplack 1993: 252). Historically, sociolinguistic investigators have also used the term "vernacular" to refer to a low-prestige variety in contradistinction to a standard and high-prestige variant, associating the vernacular both with social groups (e.g., African Americans) or with localities (e.g., Belfast English) (Milroy 1987b: 58).

My fieldwork situation was bilingual rather than bidialectal, and without a ready binary distinction of standard vs. vernacular varieties. The language I am studying, Tatar, is a is a Turkic language currently spoken by one quarter of the

* The fieldwork upon which this paper is based was made possible by generous grants from the International Research and Exchanges Board and from the Academy for Educational Development.

four million residents of Tatarstan,[1] a semi-autonomous republic of the Russian Federation, and it has been under significant stress from Russian for hundreds of years. However, since the fall of the Soviet Union and the declaration of Tatarstan's autonomy, the republic's language policies have been promoting the use of Tatar and the expansion of its functional domains. Russian remains the dominant language and is high prestige, but in the milieu of the Tatar Social Club that was the locus of my fieldwork, as well as among the Tatar intelligentsia, the Tatar language, more precisely "pure" Tatar (*saf tatar tele*) and "literary" Tatar (*ädäbi tatar tele*), is also awarded high prestige. Additionally, as will be seen later, there is a continuum of linguistic performance for urban Tatar bilinguals, ranging from completely Russian to completely Tatar, further complicating the designation of a single "unmarked" vernacular.

1. Observation and audience design

While some linguistic fieldworkers (e.g., Baugh 1993, Besnier 1994) claim that informants can forget about recording and produce naturalistic speech, I found that in my own fieldwork, the presence of the language investigator, with or without recording materials, would often be enough in and of itself to precipitate "performed" speech rather than "unperformed." Performance speech, as defined by Schilling-Estes (1998), is "associated with speakers' attempting to display for others a certain language or language variety, whether their own or that of another speech community" (53). Only after my return home from the field, and extensive review of my recordings and fieldnotes, did I realize that the linguistic performance and style shifting of the Tatar speakers I had observed and recorded could be best interpreted by taking into account two major factors: (1) the speaker's assessment of my social role, particularly as in-group or out-group member, and (2) my participant role in the speech event in question.

The first sociolinguistic investigations conceptualized style shifting on a single continuum ranging from careful to casual speech. Labovian sociolinguistic interviews were designed to elicit more- and less-careful styles, and topics introduced by the interviewer were meant to create contexts for casual speech.

Bell, following Brown and Levinson (1979), dismissed Labov's attention-to-speech continuum as an "impoverished" view. He proposed an alternate explanation, that of *audience design*, which holds as a basic tenet that "at all levels of language variability, people are responding primarily to other people. Speakers are designing their style for their audience" (1984:197). Variables such as topic and setting are seen to have less effect upon stylistic variation than audience, which is the "responsive, critical forum before whom the utterances are performed" (161): for empirical testing of this claim cf. Rickford and McNair-Knox (1994) and Lewis (2002), *inter alia*. The audience design framework is generally held to be superior to the attention-to-speech continuum, which has fallen into disuse.

[1] And elsewhere in the territories of the former Soviet Union.

Refining upon Goffman's (1981) "participant framework," Bell offers five participant roles for any given speech situation, which are as follows:

(1) Bell's Partipant Roles (Bell 1984)

1. Speaker – uses the 1st person,
2. Addressee – addressed in 2nd person,
3. Auditor – referred to in 3rd person,
4. Overhearer – unratified to participate in the speech event, not addressed or referred to,
5. Eavesdropper – both unratified and unknown.

Each step down on the list increases the distance from the speaker, sometimes physically. Figure (2) below summarizes the attributes of audience roles in a speech event. A fieldworker can participate in any of these audience roles.

(2) Hierarchy of attributes and audience roles (adapted from Bell (1984))

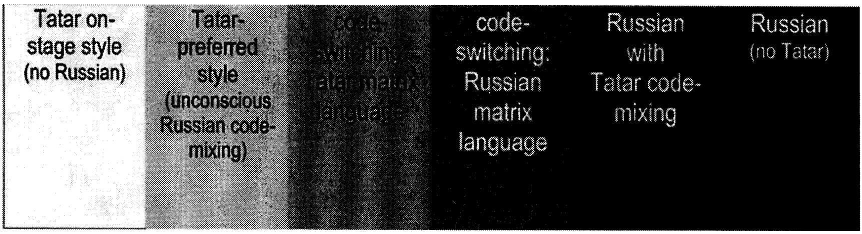
	Known	*Ratified*	*Addressed*	*Person*
Addressee	+	+	+	2nd
Auditor	+	+	-	3rd
Overhearer	+	-	-	n.a. (unratified)
Eavesdropper	-	-	-	n.a. (unratified)

Audience roles are assigned by the speaker, and will have different levels of salience for the speaker's style design—for example, auditor effect is usually lower than addressee effect. Style shifting can result in convergence to or divergence from the audience—convergence is seen as accommodation, and an expression of what Brown and Gilman (1960) call "solidarity," while divergence is interpreted as "referee design," which marks the speaker as a member of a social group not present in the speech situation—but referred to by his or her divergent style—and marks the audience (most commonly an addressee) as a non-member of the referred-to group.

Therefore, one must take into account both the fieldworker's participant role in a speech event and her or his position as a socially located being. As a member of neither the majority nor minority group, yet somehow aligned with minority group interests and culture, my own social position was unique. My attempts to speak Tatar would elicit commentary on my performance, on the linguistic performance of other Tatars, and on the Tatar language itself. It quickly became clear that Tatar speakers felt responsible for presenting me with the best possible Tatar, both so I could have appropriate models for learning, and so I could represent the language well in my research. I received invitation after invitation to be taken "back to the village," where I could hear "real" Tatar spoken, Tatar that was "purer" than the urban dialects, which had been "tainted" by Russian. As my Tatar competence improved, I found that this high level of language awareness, with its stated ideals of "pure" and "literary" Tatar, was not merely provoked by the presence of the language learner and investigator. This "discourse of purity" and constant awareness of the level of Russification of Tatar, combined with the idea of "saving" the Tatar language from both misuse and disuse, was not only presented to me, but also found in newspaper articles and opinion pieces, on television, on Internet bulletin boards, and on the radio.

Over time, as speakers' familiarity with me increased and my positioning within the community became increasingly in-group, I eventually gained access to a wide range of styles. What I found was that urban bilingual Tatars had a continuum of linguistic performance that was related to language mixing, as shown in Figure (3), below.

(3) Cline of language mixing for urban Tatar bilinguals [2]



Tatar on-stage style is often found in the public sphere in formal registers, particularly when aiming for a high literary standard, but can also found in informal register. In this style, speakers will de-Russify their Tatar to the best of their ability. Performances in Tatar on-stage style will generally take place at Tatar cultural events, in radio and television interviews, in political speeches, and in public presentations or comments at the Tatar Social Club. Private use of Tatar on-stage style seems to only occur in conversations with or for investigators of Tatar language and culture.

Tatar-preferred style is always in informal register, and is found in private conversations where Tatar has been in some way established as the preferred language of communication: for example, this is the main style of intergenerational family communication in Tatar-speaking homes, both urban and rural. For intragenerational family interactions, generally Tatar-preferred style will be used by parents, while siblings will interact in one of the code-switching

[2]I would like to acknowledge up front that these styles are arranged according to a single stylistic variable, the level of use of Russian, even though there are other markers that differentiate, for example, Tatar on-stage style from Tatar-preferred style. Additionally, linguistic performance within each style can vary, e.g., Tatar on-stage style can be both in formal register and in informal register. And within formal register of Tatar on-stage style there are further nuances, such that speeches given at a literary tribute evening are different from the on-stage performance of the master of ceremonies at a concert. I in no way wish to suggest that this single variable, the level of Russification, is the sole differentiating marker of Tatar style and performance.

styles, or even in Russian with Tatar code-mixing. In Tatar-preferred style, the level of language awareness and verbal hygiene (Cameron 1995) is lower than in Tatar on-stage style, and although speakers believe themselves to be speaking "pure" Tatar, they actually are unconsciously using what I call Russian "metalanguaging" words (following Maschler 1994). These code-mixed words—not standard borrowings, and mostly function words, adverbs of manner, and interjections—all have pragmatic functions and both structure and comment upon discourse (for a more complete examination of Tatar-preferred style and code-mixing, see Wertheim 2003).

2. The role of the language investigator in style shifting

Upon reviewing my data, I found that I was able to discern patterns in style shifting that showed that a speaker's choice of style was dependent in part upon the audience role assigned to the fieldworker participating in a speech event. This can be demonstrated by the style shifting of one Tatar bilingual, who I will call Galimä,[3] a 46-year old Tatar philologist with whom I had both a professional and personal relationship, and who I observed in a variety of situations, locations, and speech events. I believe that for Galimä my social role remained reasonably constant—I was non-peer (15 years younger), outsider, language learner and language investigator. Although our relations were quite warm, and we discussed personal matters frequently, she would only address me using the formal version of 'you' (*sez*) and never once used the informal 'you' (*sin*).

(1) In dyadic conversations (where I was both a speaker and an addressee) in either a professional or a personal context, Galimä was always in Tatar on-stage style. As with many other Tatars who felt responsible for helping me in my Tatar language learning and investigations, she would choose ideology over communication, and never used Russian with me, opting for a simplified Tatar—or even drawings—to explain words or phrases I hadn't understood, when a single Russian word would have sufficed.

(2) Galimä would use Tatar on-stage style with colleagues, friends, and acquaintances when ratifying them as participants in conversations in which I would be speaker, addressee, or auditor. She would do this even with speakers whom she knew to have limited Tatar competence. This was clearly a self-conscious "performance" of Tatar, and arrangement of performance by others, one that was staged for me. Very often it was only Galimä's participation in the conversation that kept it in Tatar on-stage style—if she would leave, speakers would frequently ask in Tatar if I knew Russian, and when I answered in the affirmative, would either switch to Russian with no Tatar in it, or would code-switch with Russian as the matrix language.

(3) If I was an auditor of a conversation with family and friends, Galimä would speak in Tatar on-stage style, with no Russian whatsoever. For example, if we were drinking tea in her kitchen and talking, and her son came in and asked a

[3] A name I have chosen for its meaning of "scholarly, knowledgeable."

question, she would use only Tatar with him. Additionally, if he used any Russian, she would upbraid him and tell him to speak in Tatar only. However, if I was not an auditor but rather an overhearer, and thus unratified as a conversation participant, Galimä would code-switch with Tatar as the matrix language. For example, one time I arrived early for a meeting with Galimä and entered the room while she was still consulting with her younger cousin, advising her on how to best teach Tatar periphrastic verbs. Galimä saw me enter the room—I sat at the opposite end of a long table, read a newspaper, and surreptitiously listened as the two spoke mostly in Tatar, but with code-switching into Russian of higher-level grammatical constituents and code-mixing of isolated words. As soon as their meeting was over, I became a ratified conversation participant, and when Galimä introduced me to her cousin, she switched into Tatar on-stage style, with no Russian, and her cousin followed suit.

(4) Galimä would speak Russian with those bilingual Tatars with whom Russian is the usual language of interaction, but if I was an auditor of one of these Russian-language interactions, my auditing would cause linguistic meta-commentary on Galimä's part. For example, one time we stopped off at the post office, where the transaction was conducted in Russian. As we turned away to leave, resuming our Tatar-only conversation, Galimä said to me, "I don't know why I speak Russian with that woman. I've been coming here for years. She knows that I speak Tatar, and I know that she speaks Tatar. So why do we speak in Russian?" I don't want to say that my presence brought this fact to Galimä's attention for the first time, but I believe that my auditing did trigger the meta-commentary. Perhaps this is because of the conflict that my auditing caused in Galimä's style-shifting. Recall that Galimä's usual behavior when I was a ratified conversation participant was to use Tatar whenever possible with anyone whom she knew to be a Tatar speaker, regardless of their competence. However, local linguistic norms required her to shift to Russian when transacting post office business, and this requirement seems to have superseded the style-shifting patterns that were based on my presence. Perhaps this conflict in linguistic presentation of identity—culturally competent citizen on the one hand, and speaker of pure Tatar on the other—is what caused her explicitly stated dissatisfaction with her linguistic performance.

(5) When conversing with Russian monolinguals, Galimä would speak Russian only, regardless of my participation role. This would sometimes lead to an interesting phenomenon—a sort of Russian hangover, where Galimä would be "out of phase" in her style shifting, such that after the Russian-language conversation had ended, she would return to our conversation and address me in Russian. However, this Russian performance would only last for one conversational turn, because regardless of the language of my response, Galimä would become immediately aware of her "inappropriate" style. I could respond in Russian, or I could respond in Tatar, but her response, always in Tatar, would be the same: "Why am I speaking Russian with you? We don't speak Russian

together. Let's speak Tatar." And we would continue our conversation with her in Tatar on-stage style.

In summary, we can see that my various participation roles seem to have influenced both Galimä's style shifting and her awareness of style shifting. Perhaps this can be seen more clearly in Figure (4), below.

(4) Effect of changing participant roles of fieldworker on Galimä's speech

Fieldworker role	*Speech event participants*	*Situation*	*Galimä's speech style*
Speaker/ addressee	friends; family; colleagues; acquaintances; no other participants	private conversation	Tatar on-stage style
Auditor	friends; family; colleagues; acquaintances	private conversation	Tatar on-stage style
Overhearer	friends; family; colleagues; acquaintances	private conversation	Code-switching: Tatar as matrix language
Auditor	service personnel	business transaction	Russian with post-transaction metacommentary
Auditor	Russian monolinguals	all	Russian with periodic "Russian hangover" followed by metacommentary

3. The role of the language investigator in recording speech

The various performance styles I have just described were produced in the presence of a fieldworker who was not explicitly in her role as fieldworker at the time of the speech events in question, and who was without any obvious note-taking or recording equipment. Now I would like to turn to the recorded speech event, and briefly examine this too from within the framework of audience design.

A recorded private-domain speech event, regardless of the physical presence or absence of the fieldworker, is an atypical, even extraordinary situation that is not classifiable within Bell's ordinary hierarchy of audience roles. And here is why: the recording equipment, previously analyzed by some sociolinguists as a participant itself in the speech event, actually represents an end-listener or listeners whose identity is not known at the time of the speech event. This means that the speech event participant represented by the recording equipment is simultaneously *ratified* (providing that permission to record has been requested and granted) and *unknown*—a participant role that is not analogous to any found in Bell's framework because it is unique to the experience of being investigated by a fieldworker of some sort, linguistic or otherwise. In (5) below I have added the end-listener of recorded speech to the audience role hierarchy so it can be easily compared with the other standard audience roles. The "strange" or "unnatural" behavior of recorded speech event participants can thus, in part, be interpreted as speakers trying to grapple with a participant role they have never dealt with before, that of the unknown eavesdropper who is nonetheless ratified. Perhaps what is unnatural is not so much the behavior and performance of speakers as the uniquely conflicting audience role attributes of the recording's end listener, for whom speech must be designed, but whose social role and identity remains cloaked in mystery (or at least unclear).

(5) Adjusted hierarchy of attributes and audience roles for a recorded speech event

	Known	*Ratified*	*Addressed*	*Person*
Addressee	+	+	+	2nd
Auditor	+	+	-	3rd
Overhearer	+	-	-	n.a. (unratified)
Eavesdropper	-	-	-	n.a. (unratified)
End-listener of recorded speech	-	+	-/+	2nd/3rd/n.a.

Recall that according to Bell, speakers design their speech with audience members in mind, and those participants whose audience roles are higher up on the hierarchy (as seen in Figure 1) will usually have a greater effect on the linguistic performance of the speaker. I submit that the unnatural audience attributes of the end-listener, highlighted by the act of recording and the presence of recording equipment, can cause this audience role to be of primary salience and effect. The fieldworker may be entirely absent and yet still the most salient participant. By understanding the identity and traditional audience role assigned to this end-listener by the speakers being recorded, it is possible to more systematically account for her or his effect on recorded speech. The end-listener can be seen as the fieldworker, and the fieldworker alone; can be seen as the fieldworker in combination with other language investigators; and can be seen as a person or persons completely unknown to the speaker. For example, some of the people I recorded pictured me alone as the end listener, such that the end-listener took on the attributes of addressee. In three separate recorded conversations, people performed for me in the most standard sense of the word, singing Tatar songs directly into the recording equipment. Several of them later volunteered the information that they had sung as a memento, so that later on when I was back home in America, I would have these nice songs to listen to, and think of them fondly. We can extrapolate from this performance and subsequent meta-commentary that these speech events were performed with consciousness of my eventual review of the recorded material and subsequent linguistic analysis. These people were speaking "good Tatar" both for me and for posterity, and their level of verbal hygiene remained quite high.

4. The value of performance speech

Bauman (1975, 1977) argued that in addition to "vernacular" speech, "performed" speech is available and meaningful for analysis, yet it remains understudied by sociolinguists and linguistic anthropologists. Ochs (1988) used performance speech to illuminate the linguistic counterpart to the Samoan system of spatial contrasts, and her observations prompted her to study the relationship between children's acquisition of language and their acquisition of culture. Schilling-Estes (1998) found performance speech in Ocracoke English to be both regularly patterned (contra Labov), and useful in gaining insight as to what elements of the dialect were most salient to its speakers.

Tatar on-stage style, particularly when combined with metacommentary and speaker evaluations by members of the speech community, was absolutely indispensible in my understanding of a variety of local norms and practices. Performance speech pointed the way to the "discourse of purity" that was the most significant of the language ideologies that I found, and exemplified speaker's attempts to produce "pure" Tatar. It demonstrated verbal hygiene in practice, and showed me who were the community's standard keepers. When speakers remained in Tatar on-stage style rather than accommodating their uncomprehending audience[4] by switching into Russian, it alerted me to the existence of Tatar speakers who choose ideology over communication, for whom out-group referee design supersedes accommodation and convergence. Tatar on-stage style, when produced by less-competent Tatar speakers who feel compelled to refrain from using any Russian at all, can provide excellent examples of morphosyntactic interference. For fully competent speakers, a comparison of their linguistic performance in Tatar on-stage style and in Tatar-preferred style gives evidence for what appears to be two different underlying competences and grammars: the one accessed by Tatar on-stage style has limited Russian influence, and the one accessed by Tatar-preferred style appears to be what is referred to as a "composite matrix language," a language with lexical, morphological, and syntactic elements of both languages in contact that is posited by some linguists to be an intermediary step in multi-generational language shift. Using only one of these styles for a grammatical description would lead to an incomplete and erroneous analysis. Finally, placing "performance" speech within a continuum of language-mixing styles gives insight into the identities and roles that are associated with certain kinds of styles, and how choosing styles when constructing identity may play a role in the stylistic and domain shrinkage of a language, and thus play a role in language shift.

5. Conclusions

In her early critique of the Labovian sociolinguistic interview as a method for collecting spontaneous speech, Wolfson called into question the idea of a "single, absolute entity answering to the notion of natural/casual speech" (1976: 202). And indeed, not one of the styles that I observed among urban Tatar bilinguals can be regarded as more natural or unmarked than any of the others: they are all designed to be appropriate for the audience, topic, and setting of the speech event in which they are produced.

[4] Sometimes, but not always, me.

Abandoning the search for a natural and unmarked vernacular has important ramifications for linguistic fieldwork and analysis. The first is that no speech should be dismissed as unnatural and thus unusable, but that instead speech events should be analyzed according to audience and participant roles in the context of local norms and ideologies. Second, performance speech can be utilized for linguistic analysis in a variety of ways, for example, to elucidate verbal hygiene patterns, or to demonstrate which elements of a language or dialect are most salient to its speakers. Third, the fieldworker should try to gain access to as many styles and registers as possible by using her or his unique social status. Varying one's recording methodology—recording when the fieldworker is both present and absent, recording in different settings, and having community members record for you—can increase access to various styles, particularly those that are in-group only, and reduce the salience of the "unnatural" audience role of the unknown yet ratified end-listener.

To conclude, I have briefly addressed just one of the difficulties of sociolinguistic fieldwork, the gathering and contextualizing of spontaneous speech. I have attempted to demonstrate that by framing the observer's paradox within audience design theory, it is possible to more systematically account for and analyze the effect of the fieldworker upon speech production. I have once again called into question the idea of a single, unmarked, unperformed vernacular, the access to which is supposedly blocked by the observer's paradox. Finally, I have shown that a "performance" or "on-stage" style can be just as useful for sociolinguistic analysis as other styles. By maintaining what Hyman (2001) calls the "fieldworker mentality," which keeps one open to unanticipated phenomena, the language investigator may find that performance speech will lead to both unexpected and unexpectedly fruitful avenues of linguistic inquiry.

References

Baugh, John. 1993. Adapting dialectology: The conduct of community language studies. In D. R. Preston (ed.) *American dialect research.* Amsterdam: John Benjamins.

Bauman, Richard. 1977. *Verbal art as performance.* Rowley, MA: Newbury House.

—. 1975. Verbal art as performance. *American Anthropologist* **77**: 290-311.

Bell, Alan. 1984. Language style as audience design. *Language in Society* **13**: 145-204.

Besnier, Nico. 1994. The truth and other irrelevant aspects of Nukulaelae gossip. *Pacific Studies* **17**:3.

Brown, Roger & Albert Gilman. 1960. The pronouns of power and solidarity. In T. A. Sebeok (ed.) *Style in language.* Cambridge, MA: MIT Press.

Brown, Penelope & Stephen Levinson. 1979. Social structure, groups, and interaction. In K. R. Schere and H. Giles (eds.) *Social markers in speech.* Cambridge: Cambridge University Press.

Cameron, Deborah. 1995. *Verbal hygiene.* London/New York: Routledge.

Goffman, Erving. 1981. *Forms of talk.* Philadelphia: University of Pennsylvania Press.

Geertz, Clifford. 1983. *Local knowledge: Further essays in interpretive anthropology.* New York: Basic Books.

Hyman, Larry. 2001. Fieldwork as a state of mind. In Paul Newman and Martha Ratliff (eds.) *Linguistic fieldwork.* Cambridge/New York: Cambridge University Press.

Labov, William. 1972. Some principles of linguistic methodology. *Language in Society* **1**: 97-120.

Lewis, Julie. 2002. *Social influences on female speakers' pitch.* PhD dissertation, University of California.

Maschler, Yael. 1994. Metalanguaging in bilingual conversation. *Language in Society* **23**: 325-366.

Milroy, Lesley. 1987a. *Language and social networks*, 2nd edition. Oxford/New York: B. Blackwell.

——. 1987b. *Observing and analysing natural language: A critical account of sociolinguistic method.* Oxford/New York: Blackwell.

Ochs, Elinor. 1988. *Culture and language development: Language acquisition and language socialization in a Samoan village.* Cambridge: Cambridge University Press.

Poplack, Shana. 1993. Variation theory and language contact. In D. R. Preston (ed.) *American dialect research.* Amsterdam: John Benjamins.

Rickford, John & Faye McNair-Knox. 1994. Addressee- and topic-influences style shift: A quantitative sociolinguistic study. In D. Biber and E. Finegan (eds.) *Sociolinguistic perspectives on register.* New York/Oxford: Oxford University Press.

Schilling-Estes, Natalie. 1998. Investigating 'self-conscious' speech: The performance register in Ocracoke English. *Language in Society* **27**: 58-83.

Wertheim, Suzanne. 2003. *Linguistic purism, language shift, and contact-induced change in Tatar.* PhD dissertation, University of California, Berkeley.

Wolfson, Nessa. 1976. Speech events and natural speech: Some implications for sociolinguistic methodology. *Language in Society* **5(2)**: 189-210.

1203 Dwinelle Hall
University of California, Berkeley
Berkeley, CA 94720-2650

wertheim@socrates.berkeley.edu

BERKELEY LINGUISTICS SOCIETY

Linguistics Department, 1203 Dwinelle Hall, Berkeley, CA 94720-2650, USA
Telephone: (510)642-5808 Fax: (510)642-5808 E-mail: bls@socrates.berkeley.edu
http://www.linguistics.berkeley.edu/BLS

Price List

Vol.	Year	Parasession/Special Session (S) Topic	Indiv.	Inst.
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 Ling.	$8.00	$12.00
INDEX		Author Index to BLS 1-16S	$2.00	$4.00
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 Lgs	$8.00	$12.00
BLS-20	(1994)	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 Sociolgx /Social Issues in Hist. Lgx	$16.00	$20.00
BLS-21S	(1995)	Discourse in Southeast Asian Linguistics	$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	$4.00	$8.00
BLS-25	(1999)	Loan Word Phenomena	$20.00	$25.00
BLS-25S	(1999)	Caucasian, Dravidian and Turkic Linguistics	$8.00	$12.00
BLS-26	(2000)	Aspect	$20.00	$25.00
BLS-26S*	(2000)	Syntax and Semantics of the Indigenous Lgs of Americas	TBA	TBA
BLS-27*	(2001)	Language and Gesture	TBA	TBA
BLS-27S*	(2001)	Afroasiatic Linguistics	TBA	TBA
BLS-28	(2002)	Field Linguistics	$20.00	$25.00
BLS-28S*	(2002)	Tibeto-Burman/Southeast Asian Linguistics	TBA	TBA

Shipping and Handling (USPS ground)**

	First Book	Each Addtl book
Within U.S.	$4.00	$2.00
Canada and Mexico	$7.00	$3.50
All other countries	$9.00	$4.00

Orders must be accompanied by pre-payment. We accept checks and money orders in US dollars drawn on a US bank. Sorry, no credit cards. California residents must include sales tax according to county of residence. *Still in publication at the time of this printing; check our website for availability. **Please note that our shipping rates increase as USPS rates increase; check website for current rates.