

Semantically-Oriented Vowel Reduction in an Amazonian Language

CALEB EVERETT

University of Miami

Introduction

Numerous phonetic studies, both recent (Clopper and Pierrehumbert 2008) and less so (Lieberman 1963, Lindblom 1963), have demonstrated clearly that English speakers reduce vowels, both temporally and spectrally, in certain contexts. For example, findings in studies such as Fourakis (1991) have shown that vowels in words from portions of relatively fast speech are spectrally reduced when contrasted to their counterparts in the same words when uttered in segments of relatively slow speech. More interestingly perhaps, several studies (e.g. Munson and Solomon 2004, Scarborough 2006) have provided clear evidence that vowels are temporally and spectrally reduced when occurring in high frequency words.

Increased speaking rate and word frequency are not the only factors that correlate positively with vowel reduction, however. At least two other less straightforward but significant influences on vowel duration and vowel spectral composition have been evinced in the phonetic literature. The first of these is lexical density, i.e. the number of close phonological neighbors a given word has in the lexicon. Vowels in words with low lexical density, that is, words that have few corresponding words that differ according to only one phoneme, have been shown to exhibit a tendency towards spectral and temporal reduction (Munson and Solomon 2004). Put differently, vowels in words such as *choice*, which has a limited number of so-called phonological neighbors (e.g. *voice* and *chase*), tend to be reduced when contrasted to vowels in words such as *cat*, which has a relatively high number of phonological neighbors (*sat*, *scat*, *at*, *cot*, *kit*...). Munson and Solomon (2004) demonstrate that lexical density and word frequency interact, such that vowels in frequent words with few lexical neighbors are particularly reduced.

The intuitive and plausible motivation for the relationship between phonological density and vowel reduction (or expansion, depending on the perspective chosen) is that speakers have greater need to increase clarity when producing words that could more easily be confounded with a similar lexical item.

Conversely, speakers are afforded the convenience of reducing vowels when producing words that are unlikely to be confused with other lexical items. In other words, ease of intelligibility seems to motivate vowel reduction in many cases.

In a similar vein, phoneticians have demonstrated that vowels are often reduced in another context in which intelligibility is naturally facilitated. Specifically, vowels are often reduced when they occur in words that are predictable semantically. This finding was first made in Lieberman (1963), and has been replicated most recently in Clopper and Pierrehumbert (2008). In the latter study it was found that vowels in monosyllabic clause-final words that are fairly unpredictable from the semantic context denoted in the clause are less reduced than vowels in the same monosyllabic words produced clause-finally in clauses whose semantics are more predictive of the final word in question. Interestingly, Clopper and Pierrehumbert (2008) found that the effect of semantic predictability on vowel reduction was not uniform across the three American English dialects considered in their study. For example, they found that the way in which speakers from Chicago and northern Indiana reduced their vowels differed from the manner in which southern speakers reduced theirs. Specifically, the northern speakers tended to reduce their vowels, along the F1-F2 plane, in the direction of the Northern Cities Chain Shift that characterizes their speech more generally. For more details on the vowel reduction patterns observed, I refer the reader to the original study. What is relevant for our purposes is that the findings in that study and in Lieberman (1963) suggest that, generally speaking, American English speakers reduce the vowels of words when those words are somehow semantically predictable from clausal context.

In order to better illustrate what is meant by semantic predictability in such cases, it is worth providing some of the stimuli utilized by Clopper and Pierrehumbert (2008). In that study, the researchers had speakers produce clauses such as those in clause-pairs (1) and (2).

- (1a) The chicken pecked corn with its beak.
- (1b) She is glad Bill called about the beak.

- (2a) The shepherds guarded their flock.
- (2b) Paul should have discussed the flock.

Note that in (1a) and (2a), the words occurring at the end of the clause are relatively predictable given the semantic context established by the preceding portion of the clause. In (1b) and (2b), however, the words are less predictable. For instance, one would not be surprised to find that the instrument used for corn-pecking in (1a) is a beak, but one might very well be surprised that Bill called about a beak as in (1b). An analogous analysis holds for all of the clause pair stimuli utilized by Clopper and Pierrehumbert (2008).

Put simply, the fact that the vowels in clause-final words such as those in (1a) and (2a) tend to be reduced when contrasted with the clause-final words in (1b)

and (2b) suggests that vowels in words with greater semantic predictability are generally reduced. One could argue that, as in the case of the increased reduction of vowels in words with low phonological density, this phenomenon is driven by the increased ease of intelligibility associated with such contexts. That is, it appears that heightened intelligibility motivates semantically-oriented vowel reduction.

While it is an interesting phenomenon, semantically-oriented vowel reduction has only been demonstrated in studies with English-speaking participants. The primary objective of the current endeavor is to establish whether or not there is cross-linguistic evidence for semantically-oriented vowel reduction. In order to do so, we will consider some recently-collected data from Kariti'ána (K henceforth). K is a Tupí language spoken in southwestern Amazonia. It is a natural subject for the current study since the author regularly conducts phonetic research on the language, and since there are a number of recent phonetic studies available in the literature on the language (Everett in press a, 2008a, *inter alia*). An ancillary objective for this study is to contribute to the growing acoustic-phonetic literature on the language and on Amazonian languages more generally (see discussion in Everett in press b).

1 Relevant background on the language

Given that much has been published on the sound system of K, I choose not to represent the basics of the sound system here. The K examples below are transcribed using broad phonetic (IPA) transcription. The most typologically-remarkable facet of the sound system is an interesting phenomenon referred to by Everett (in press a) as ‘variable velic movement’, according to which the velum is lowered for a completely unpredictable duration in certain contexts. For example, the word for ‘thing’, /kina/, may surface as [kidnda], [kinda], or [kida] due to random perturbations in velic aperture duration. This variability of velic aperture duration contravenes the literature on nasality in ways described in Everett (in press a).

Not all of the literature on the K sound system focuses on such nasal patterns. Everett (2008a) presents a locus-equation based analysis of the three voiceless stops in the language and, significantly for our purposes, Everett (2008b) presents an acoustic analysis of the vowels of the language. As noted in that study and in numerous preceding studies beginning with the work of David and Rachel Landin (SIL members who worked on the K language in the 1970's), there are eleven phonemic consonants and five vowel types in the language. Vowel length and nasality are contrastive in the language, however, creating a system of fifteen potential vowels. Our focus here is on short oral vowels. These oral vowels, namely /i/, /e/, /ɨ/, /o/, and /a/, are also the focus of Everett (2008b). In that study the mean normalized vowel spaces for eight adult speakers, four males and four females, are presented. The normalized vowel spaces suggest that the only truly high vowel in the language is /i/. The /ɨ/ vowel is located between /i/ and the /e/ and /o/

vowels along the F1 (height) dimension. The /a/ vowel is the only low vowel in the language, and is located centrally along the F2 dimension, much like /ɪ/. Among the other findings presented in Everett (2008b) is one that is particularly relevant to the current study. Specifically, in that study it was found that unstressed vowels are characterized by significant phonetic reduction. In general, such vowels tend to occur in a more central location of the F1-F2 plane, nearer to tokens of other vowel types. Vowels in stressed contexts are generally more dispersed within the vowel space and occur in more restricted/specific regions of that space. The patterns of phonetic reduction characterizing unstressed vowels in K are relevant to the results of the current study in straightforward ways discussed briefly in the following section.

2 Methods and participants

The methods employed in this study are generally modeled after those in Clopper and Pierrehumbert (2008). The methods do deviate intentionally in some ways from those in that study, however.

As was noted in the introduction, Clopper and Pierrehumbert (2008) recorded speakers producing clauses such as (1a)–(2b) in which the analyzed vowels occurred in the final monosyllabic word of a given clause. For this study, K speakers were recorded producing clauses in which the word with the analyzed vowel occurred clause-finally or penultimately. While four English monophthongs were examined in Clopper and Pierrehumbert (2008), all five K monophthongs were examined for this study. The ten clauses utilized are listed in (3a)–(7b). The highlighted vowels in five different words were examined. In half the clauses, the relevant word occurred in a context that was fairly predictable semantically, and in the other half the same word occurred in a context that was less predictable semantically. In (3a)–(7b) the clauses are listed as pairs, with the (a) variants containing the relatively more predictable contexts.

Eight female K speakers were recorded producing three tokens of each of the ten clauses. The order of the recorded clauses was random in each case. Speakers were recorded onto a Mac powerbook at a sampling rate of 44.1 kHz. The duration and formant analyses described below were calculated via wide-band FFT-based spectrograms.

- | | | | |
|------|---|---------------------------------|--------------|
| (3a) | i | na-miŋgidna-t | kinda oti?ap |
| | 3 | NSAP ¹ -swallow-NFUT | medicine |
| | | ‘S/he swallowed the medicine.’ | |

¹ Glosses: NSAP: Verbal prefix used when the absolutive referent is not a speech act participant. CAUS: Causative. NFUT: Nonfuture. FUT: Future. INT: Semantic intransitivity marker. COP: Copula. S: Singular.

Semantically-Oriented Vowel Reduction in an Amazonian Language

- (3b) i naka-m-ʔa-t kinda otiʔap
3 NSAP-CAUS-make-NFUT medicine
'S/he made the medicine.'
- (4a) taso pikop na-okɪ-t ɔmbakɪ
man hunter NSAP-kill-NFUT jaguar
'The hunter killed the jaguar.'
- (4b) ɲõnso et na-atik-Ø ɔmbakɪ
woman daughter NSAP-throw-NFUT jaguar
'The woman's daughter threw the jaguar.'
- (5a) ɲõnso oʔit na-aka-t i-ambo-t ambi ohɪnt
woman adolescent NSAP-COP-NFUT INT-climb-NFUT house on/over
'The girl climbed on the house.' (Any built structure.)
- (5b) ɲõnso oʔit naka-i-t kinda o ambi ohɪnt
woman adolescent NSAP-eat-NFUT thing round house on/over
'The girl ate the fruit on top of the house.' (Any built structure.)
- (6a) ʃn naka-m-ʔa-t otʔep
1S NSAP-CAUS-be-NFUT bow
'I made the bow.'
- (6b) ʃn na-okot-Ø otʔep
1S NSAP-bite-NFUT bow
'I bit the bow.'
- (7a) ʃn na-okɪ-j sojja
1S NSAP-kill-FUT wild.pig
'I'll kill the wild pig.'
- (7b) ʃn na-mkɪwakɪwa-t sojja
1S NSAP-scrub-NFUT wild.pig
'I scrubbed the wild pig.'

It should be stressed that the vowels examined for this study (highlighted in the examples) were excised from polysyllabic words. Focusing on vowels from polysyllabic words allows us to test for semantically-oriented vowel reduction in a new phonological context, since previous studies have focused on monosyllabic words. This methodological difference may have influenced the results of this study in a specific non-significant manner discussed briefly in section 3.

Another way in which the methods used for this study differ from those in the literature is that the disparity in semantic predictability between the relevant

words in the (a) and (b) clauses is not as extreme as that evident in stimuli material in previous studies, e.g. (1a)–(2b). The goal here was to test whether semantic predictability affects vowel production even when the relevant vowels are excised from words that do not differ in radical ways according to this variable. Given that semantic predictability is necessarily a scalar variable, it seemed natural to test vowels excised from words that were found in contexts that were characterized as more or less predictable semantically, but were located at non-extreme points on the semantic predictability continuum.

3 Results

The FFT-based wide-band spectrograms for all of the words with the highlighted vowels in (3a)–(7b) were examined, for each of the tokens produced by the eight female adults. Three of these speakers were recorded outdoors, and as the data were being examined it became clear that the recordings for these three speakers would have to be discarded for the purposes of the present study. Unfortunately, the background noise levels were not controlled for properly in the outdoor setting, and it was quickly observed that the formant readings for the recordings in question were not sufficiently reliable to be included in the analysis. Nevertheless, as we will see momentarily, the five higher-quality indoor recording sessions resulted in a robust series of findings. Since each speaker was asked to produce all clauses three times, there were a total of 150 quality tokens on which to base the analysis (5 speakers x 10 clauses x 3 tokens-per-clause).

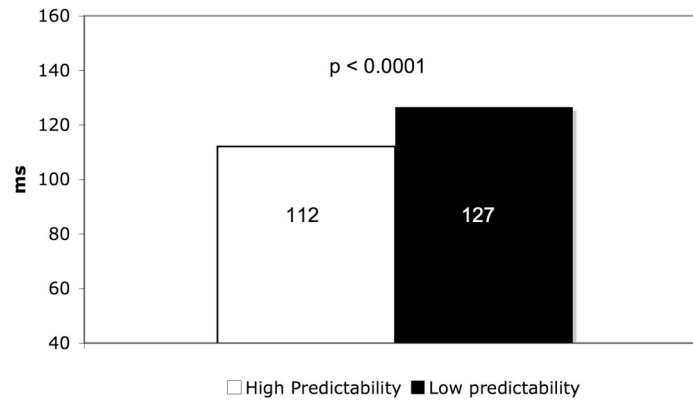
The focus of the data analysis was on two factors, as in the previous studies on vowel reduction: duration of the tested vowels, and placement in the F1-F2 plane. In the case of vowel position, vowels were tested for F1 and F2 values in Hz. The vowels were not normalized. For a more detailed analysis of normalized vowel spaces in the language, I refer the reader to Everett (2008b). The absence of normalization was not considered problematic in this case since eight females of similar size participated in the study.

Duration values for all of the tokens of the vowels highlighted in (3a)–(7b) were established by visually inspecting the FFT-based spectrograms, while also listening to the audio files on which they were based. The individual token durations gathered via this straightforward procedure were entered into a database and categorized according to three factors: semantic predictability, speaker, and vowel type. Analysis of all the tokens according to the first factor suggested that there was a significant disparity between vowels in semantically predictable contexts and those in less predictable contexts. Consistent with expectations, vowels in higher predictability contexts were shorter (mean duration=112 ms) than in low predictability contexts (127 ms). This disparity is represented graphically in (8). Note that the disparity is significant at the $p < 0.0001$ level, according to a paired two-tailed t-test. The reduced duration of vowels in contexts of higher semantic predictability was observed across all speakers. While there were inter-speaker differences in mean vowel durations, for all speakers vowels in contexts of greater

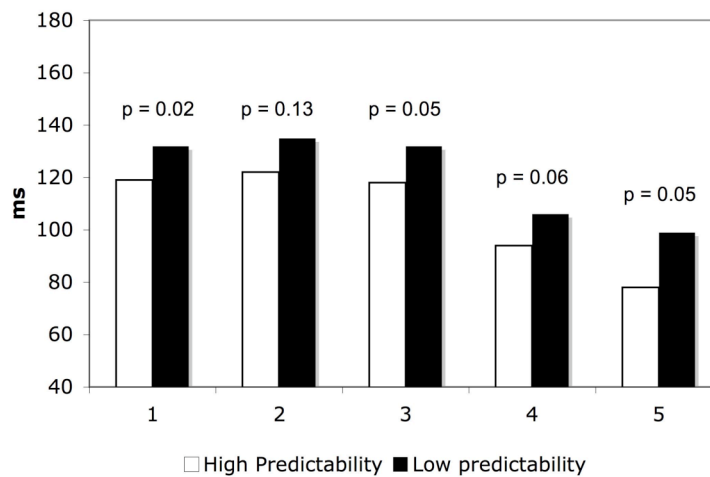
Semantically-Oriented Vowel Reduction in an Amazonian Language

predictability were temporally reduced when contrasted to those in less predictable environments. T-tests conducted for each speaker yielded significant or near significant disparities according to this variable in all cases, as we see in (9).

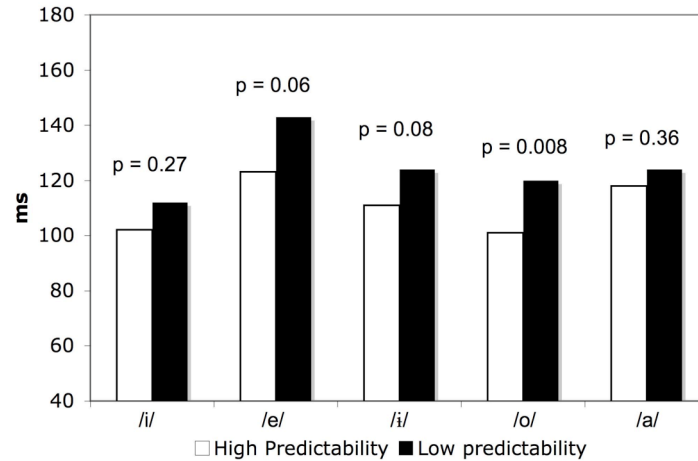
(8) Mean vowel durations according to context.



(9) Mean vowel durations according to context and speaker.



(10) Mean vowel durations according to context and vowel type.



Much as we can be confident that the inter-context duration disparities were not due to inter-speaker variation, we can also be confident that they were not conflated by variation across vowel types. In (10) we see that, for each of the five vowel types, vowel tokens in more predictable words tended to be shorter than those in less predictable words. This difference approached significance (at the $p < 0.05$ level) in the case of three vowels.

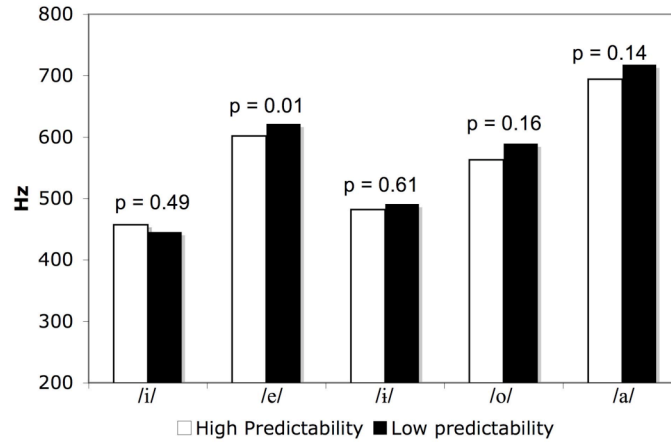
In general, then, the duration data suggest quite unequivocally that vowels in words that are more predictable semantically tend to be shorter. Interestingly, this difference is not easily perceptible in the audio stream, at least for a non-native speaker such as myself. The mean magnitude of the difference (15 ms) is generally in line with the range of differences evident in Clopper and Pierrehumbert (2008:1684). It should be noted that the vowels analyzed in this study tended to be shorter than the English vowels considered by those authors. This difference is likely the result of methodological factors however, since in the study of English vowels the relevant tokens were extracted from monosyllabic words and were stressed. In this study the vowels were generally unstressed and all occurred in polysyllabic words. The only stressed vowel analyzed here was /e/, which, as can be seen in (10), was longer than the four remaining vowels. This is not surprising given that vowel length is a correlate of word-level stress in K (cf. Everett 2008b).

With respect to vowel formant structure, the prediction made prior to vowel analysis was that vowels in semantically-predictable contexts would be reduced spectrally, i.e. occur in a more centralized portion of the vowel space. Recall that Everett (2008b) demonstrated that unstressed vowels in K are produced in a more centralized portion of the vowel space. More specifically, in that study it was observed that such vowels occur in less precisely-demarcated regions of the vowel space, when contrasted to their stressed counterparts. Based on the more specific findings on phonetic undershoot in that study (which cannot be fully delineated

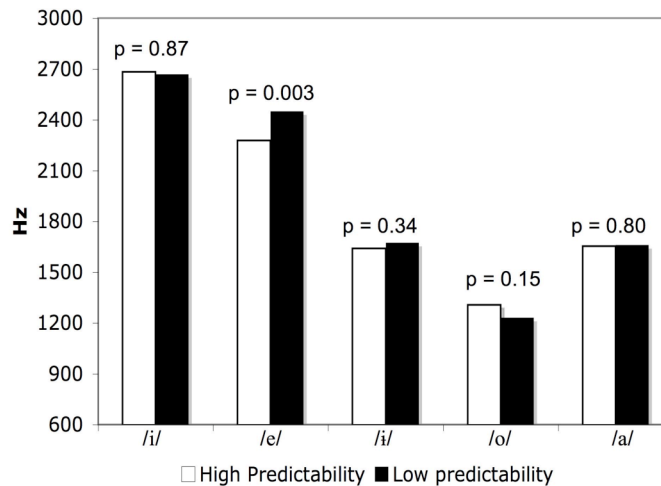
here for the sake of space), we can make three more fine-grained predictions. Reduced /a/ vowels in that study were found to be higher in the vowel space, i.e. characterized by lower F1 values. We might expect then that /a/ vowels in more predictable words would also be characterized by relatively low F1 values. Reduced /e/ vowels in that study were found to be more retracted in the vowel space, i.e. characterized by lower F2 values. We might also expect then that /e/ vowels in more predictable words would have lower F2 values. Finally, reduced /o/ vowels in Everett (2008b) were found to often occur further forward in the vowel space (likely due to decreased labialization), i.e. characterized by *higher* F2. We might naturally expect that /o/ vowels in more predictable words would have higher F2 values, then. The natural question is whether these three predictions are met by the data collected here. The pattern of reduction for /i/ and /ɨ/ vowels was less obvious in Everett (2008b), and so no clear prediction vis-à-vis the placement of these vowels can be made.

Let us begin by considering the F1 values of all the five vowel types, across both contexts. The relevant data are depicted graphically in (11). As we see in the figure, there is in fact a difference between the mean F1 values of /a/ vowels in the two contexts. While not significant, the difference trends in the predicted direction. In other words, vowels in more predictable contexts tended to have lower F1 values, i.e. were located higher in the vowel space. Interestingly, a similar unpredicted difference surfaced for the /e/ and /o/ vowels, significant in the case of the former vowel. For both vowel types, vowels in more predictable contexts had generally lower F1 values.

(11) F1 levels according to context and vowel type.



(12) F2 levels according to context and vowel type.



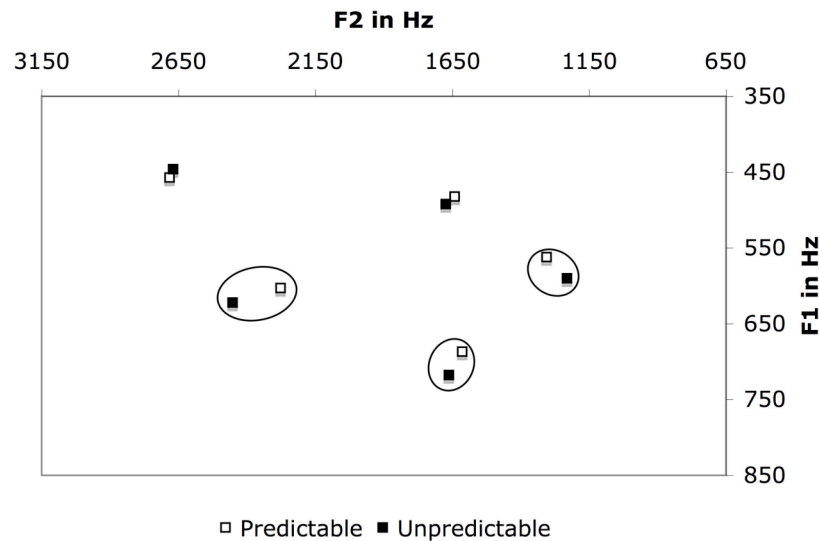
With respect to F2, the two major predictions made for vowels in more predictable contexts were that /e/ vowels would have lower F2 values and /o/ vowels would have higher F2 values in such contexts. As we see in (12), both of these predictions are clearly met by the data. In the case of the /e/ vowel tokens, those in more predictable contexts had significantly reduced F2 values, according to a two-tailed paired t-test. Conversely, in the case of /o/ vowel tokens, those in more predictable contexts had generally *greater* F2 values, though this difference only approached significance.

In general, the /a/, /e/, and /o/ vowels conformed to our expectations. In short, for all three vowels, those tokens in more predictable contexts were typically located in more centralized portions of the vowel space. This is particularly evident in (13), in which the vowels are plotted according to mean F1 and F2 values for

Semantically-Oriented Vowel Reduction in an Amazonian Language

all five speakers, according to context. Note that the /i/ and /i/ vowels do not exhibit any clear disparities in position across contexts. Significantly, no clear predictions were made regarding the position of these vowels in more predictable contexts. It should be stressed that the data in (11)–(13) do not demonstrate that these two vowel types are not reduced at all in more predictable contexts. They simply suggest that, based on the limited data considered here, there is no clear spectral reduction of these vowels in such contexts. Temporal reduction of these vowels in such contexts was already demonstrated, after all.

(13) Mean vowel locations according to semantic context.



4. Discussion and conclusion

The results of this study suggest quite strongly that vowels in K are more likely to be reduced when they occur in words that are relatively predictable based on the semantic context established by the clause they occur in. As was noted above, semantic predictability is a scalar variable and the clause pairs utilized in this study could easily have differed more according to this variable. The lack of an enormous disparity in semantic predictability may have influenced the results, that is, the extent of vowel reduction associated with semantic predictability may actually be less for these data than other possible data sets. From one perspective this may be taken as a strength of this study, since it demonstrates the correlation of vowel reduction and semantic predictability even when the degree of semantic predictability in question is relatively modest.

In a similar (somewhat speculative) vein, the results of this study may have been more robust had the vowels analyzed been excised from monosyllabic words and had they been stressed vowels in all cases. Recall from (10) that the only vowel type that was stressed in the clauses produced, /e/, exhibited the greatest

overall duration and a more readily apparent disparity between mean durations across contexts, when contrasted with the other vowel types.

In sum, the data presented suggest a significant disparity between the mean duration of vowels in less predictable words and more predictable words. For each speaker and for each vowel type, vowels in the latter category of words were typically temporally reduced. With respect to position within the vowel space, vowels in more predictable contexts tended to occur in a slightly more centralized portion of the vowel space. This spectral reduction was minor in most cases, and was restricted to the /e/, /a/, and /o/ vowels. Nevertheless, the vowels in more predictable contexts patterned in accordance with predictions for reduced vowels based on previous findings on vowel reduction in the language.

While I have suggested throughout that the data are consistent with the phenomenon of vowel reduction, it is of course possible to construe them as representing a case of vowel *expansion*. In other words, the data are also consistent with the claim that vowels in K are expanded temporally and spectrally in less predictable contexts. From this perspective, speakers may be more likely to hyperarticulate vowels in semantically unpredictable contexts. Such hyperarticulation would be plausibly motivated by the desire for speakers to heighten intelligibility, as in the case of vowel expansion in less frequent words and words with greater lexical density (Munson and Solomon 2004).

Whatever perspective on the data one takes, what seems clear is that semantic contexts affect vowel production in K. Tangentially, one might note that such semantic effects present difficulties for various phonological models in which phonetic production is seen as modular – influenced primarily by factors such as phonological context but not directly influenced by factors such as semantic context. The phenomenon described here is potentially more amenable to treatment with an exemplar-based approach to phonology (e.g. Bybee 2001), under which phonetic production is affected by a host of context-specific factors.

In conclusion, vowels in K are reduced in words produced in contexts of greater semantic predictability. To my knowledge, this study represents the first case in which the phenomenon in question has been demonstrated for a language besides English. Semantically-oriented vowel reduction is apparently crosslinguistic in nature, surfacing in dialects of English much as it does in the speech of the members of an Amazonian tribe.

References

- Bybee, Joan. 2001. *Phonology and Language Use*. Cambridge: Cambridge University Press.
- Clopper, Cynthia and Janet Pierrehumbert. 2008. Effects of Semantic Predictability and Regional Dialect on Vowel Space Reduction. *Journal of the Acoustical Society of America* 124:1682–1688.

Semantically-Oriented Vowel Reduction in an Amazonian Language

- Everett, Caleb. In press a. Variable Velic Movement in Karitiâna. *International Journal of American Linguistics*.
- Everett, Caleb. In press b. A Survey of Contemporary Research on Amazonian Languages. *Language and Linguistics Compass*.
- Everett, Caleb. 2008a. Aspects of Karitiâna Vowels. *Anthropological Linguistics* 50:266–291.
- Everett, Caleb. 2008b. Locus Equation Analysis as a Tool for Linguistic Field Work. *Language Documentation and Conservation* 2(2):185–211.
- Fourakis, M. 1991. Tempo, Stress, and Vowel Reduction in American English. *Journal of the Acoustical Society of America* 90:1816–1827.
- Johnson, K., E. Flemming, and R. Wright. 1993. The Hyperspace Effect: Phonetic Targets are Hyperarticulated. *Language* 83:505–528.
- Jurafsky, D., A. Bell, M. Gregory, and W.D. Raymond. 2001. Probabilistic Relations Between Words: Evidence from Reduction in Lexical Production. In J. Bybee and P. Hopper, eds., *Frequency and the Emergence of Linguistic Structure*, 229–254, Amsterdam: John Benjamins.
- Lieberman, P. 1963. Some Effects of Semantic and Grammatical Context on the Production and Perception of Speech. *Language and Speech* 6:172–187.
- Lindblom, B.E.F. 1963. Spectrographic Study of Vowel Reduction. *Journal of the Acoustical Society of America* 35:1773–1781.
- Munson, B. and Solomon, N.P. 2004. The Effect of Phonological Neighborhood Density on Vowel Articulation. *Journal of Speech, Language, and Hearing Research* 47:1048–1058.
- Scarborough, R. 2006. Lexical and Contextual Predictability: Confluent Effects on the Production of Vowels, paper presented at the tenth *Laboratory Phonology Conference*. Paris, France.

Caleb Everett
University of Miami
Department of Anthropology
PO Box 248106
Coral Gables, Florida 33124-2005

caleb@miami.edu