UCLA Working Papers in Phonetics, No. 103, 48-63

PHONETIC ENCODING OF PROSODIC STRUCTURE*

Patricia A. Keating

INTRODUCTION

Prosody is the organization of speech into a hierarchy of units or domains, some of which are more prominent than others. That is, prosody serves both a grouping function and a prominence-marking function in speech. As examples of the grouping function, some ways in which smaller units are combined to form larger ones (perhaps via intermediate groupings) include: segments combine to form syllables, syllables combine to form words, and words combine to form phrases. As examples of the prominence-marking function, there are at least two levels of prominence in English: lexical stress, or prominence at the word level, and pitch accent, or prominence at a phrasal level.

Some prosodic constituents:	**Some levels of prominence:**
• Utterance	• Nuclear accent
• Intonational Phrase	• Pitch accent
• Phonological Phrase/ Intermediate Phrase /Accentual Phrase	• Lexical primary stress
• Phonological Word	• Lexical secondary stress
• Foot	
• Syllable	
• Mora	

The grouping function and the prominence-marking function can be seen together in a prosodic tree of an utterance. Figure 1 is a partial prosodic tree of the phrase "that new propaganda", showing 4 levels of prosodic domains (leaving out several other levels): it shows a single Intonational Phrase (IP), containing 2 intermediate phrases (ip), the first of which contains 3 prosodic words (Wd). The third word has the phrasal accent or prominence, and it contains 4 syllables (sigma), of which the third has the lexical stress or prominence. This partial tree has syllables as its smallest prosodic units. Each syllable contains some segments, and the features of the segments are the terminal nodes of the prosodic tree. A few of these segmental features are included in the figure. As can be seen, each segment and thus each feature has a position in the tree relative to the domains and prominences.

For the purposes of this paper, we can consider any one interval of speech that is grouped into a single prosodic domain and ask, At what level of prosodic structure does this domain occur? What speech events occur at the beginning and end of this

---

* Based on a presentation at the 6th International Seminar on Speech Production in Manly, Sydney in December 2003, draft of paper to appear in a book edited by J. Harrington and M. Tabain for the Psychology Press series "Macquarie Monographs in Cognitive Science".

domain? How prominent is this domain relative to its neighbors? All of this information will be relevant phonetically.

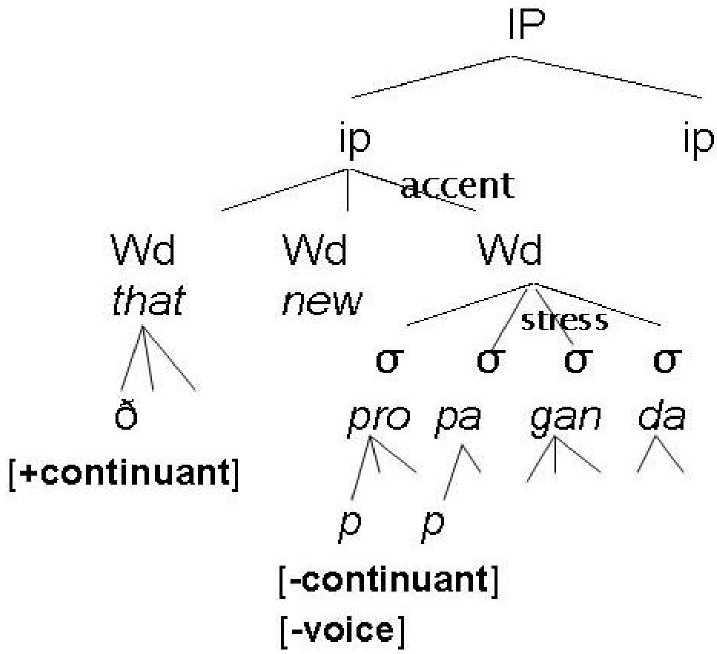

Figure 1. Sample partial prosodic tree for a phrase, « that new propaganda ». Phrasal accent and lexical stress are indicated ad-hoc on the appropriate branches in the tree. Only three segments and selected feature values are shown.

The phonetic dimensions that are most obviously involved in realizing prosody are pitch, duration, and loudness, which are generally thought of as the suprasegmental dimensions. But the phonetic dimensions that are typically thought of as more segmental than suprasegmental also serve to realize prosodic distinctions. For example, it is well-known that vowel quality varies not only with phonemic vowel identity, but also with such suprasegmental factors as stress and length (Lehiste, 1970). Put generally, the phonetic realization of an individual speech segment (vowel or consonant)'s phonological properties depends in part on that segment's position in the entire prosodic structure. The exact pronunciation of any one such feature will depend on the other features in that segment, features of neighboring segments, *and the position of the feature in the overall tree*. Thus segmental phonetic dimensions are as much about prosody as are the traditional suprasegmental dimensions.

In Levelt, Roelofs & Meyer (1999)'s important model of planning for speech production, a distinction is made between (1) *phonological encoding*, or generating a complete phonological representation, including prosody, from lexical entries and syntactic structure; and (2) *phonetic encoding*, which specifies the surface phonetic shape of the phonological representation. At each of these stages, Levelt et al., relying on the traditional distinction between segmental and suprasegmental phonological representations and speech parameters, envision segmental and prosodic planning as virtually independent. Phonetic encoding of segments is thought to operate at the level of the word, and it consists largely of retrieval of stored syllable plans. Segmental and prosodic planning interact in only a minor fashion, at the end of the encoding process, when the results of these two independent processes are brought together. As discussed at length in Keating & Shattuck-Hufnagel (2002), the missing ingredient from their model (and similarly, from the model in Levelt, 1989) is the close link between prosodic structure and segmental phonetic properties. We outlined instead an opposing view, according to which segmental and prosodic planning are not independent, since planning segmental articulation depends crucially on prosody. We stressed that even if phonetic encoding relies on stored syllable plans, the work of phonetic encoding has just begun with their retrieval, as adjustments to them are required on the basis of all kinds of prosodic information; and even if phonetic encoding relies on stored exemplars, then that retrieval operation itself must be highly sensitive to prosodic structure and the retrieved exemplar may still require further processing. The present chapter, like Keating & Shattuck-Hufnagel (2002), defends this claim, by reviewing a variety of ways in which phonetic encoding must be sensitive to prosodic structure, focusing especially on results from my own laboratory. In particular, I will review results of experiments suggesting a strength relation between prosodic positions and phonetic realizations.

The general idea of a strength relation is that prosodic positions are *stronger or weaker*; segment/feature phonetic realizations are also stronger or weaker; and *segment strength matches position strength*, with *stronger pronunciations in stronger positions*. Thus in turn segment strength serves as an indicator of positional strength. What are some strong positions? Not surprisingly, prominent syllables are prosodically strong. It turns out that other strong positions are the edges of the various sized prosodic domains. In particular, segments at beginnings of domains -- in domain-initial positions -- are generally strengthened, while segments within those domains are not.

What exactly does *strengthening* mean? Strengthening is articulatory, meaning that the articulations themselves are stronger, or more extreme. For consonants in strong positions, for example, the primary oral constriction is more extreme, meaning that the primary articulator moves farther from a neutral position into a more extreme position which reduces the size of any mouth opening. Such strengthening (often called *fortition* in the historical linguistics literature) is the opposite of *weakening* (or *lenition*), by which a more reduced primary consonant articulation results in a greater mouth opening. In historical sound changes (e.g. Hock, 1991), consonant strengthenings and weakenings are observed in different prosodic positions; for example, word-initial consonant strengthening may change an approximant consonant into an obstruent, or a continuant into a non-continuant. Word-initial segments are also preserved more often than other segments, since lenition in weak positions often leads to complete loss. Glottalization of word-initial vowels at the beginnings of prosodic domains (Dilley et al., 1996), which gives them a more consonantal quality, can also be seen as a strengthening, though of a very different sort (Pierrehumbert & Talkin, 1992).

Many previous studies, including several from Australia, have been concerned with the articulation of prominent vowels, including the direction of displacement for prominent high vowels, and differences between prominence and other strengthenings (e.g. Edwards et al., 1991; Beckman et al.,1992; Fletcher &

Vatikiotis-Bateson, 1994; de Jong, 1995; Harrington et al., 2000; Cho, 2002; Erickson, 2002). Other studies (reviewed in Epstein 2002, 2003) have been concerned with the phonation qualities associated with prominence. Strengthenings at the beginnings of prosodic domains and in prominent syllables are complemented by domain-final lengthenings, the well-known phenomenon where segments at the ends of domains have longer durations (e.g. Wightman et al., 1992). See Shattuck-Hufnagel & Turk (1996) and Fougeron (1999) for reviews of the articulation of prominence and final lengthening, which will not be discussed in any detail in this chapter.

The total effect, then, is that at the end of one domain there is a slowing down, then at the beginning of the next domain a strong attack, with another strong moment associated with any prominence. Such phonetic effects are also seen in synchronic phonological patterns. At the word level, languages may license or distribute inherently stronger segment types in stronger positions, e.g. a preference for initial obstruents, especially stops, as opposed to non-initial sonorants or continuants (e.g. Martinet, 1955; Bell & Hooper, 1978). Overall, then, there is a tendency for prosodic domains, such as words, in a given language to have a phonological and/or phonetic shape conditioned by the prosodic structure of the language.

BASIC DOMAIN-INITIAL EFFECTS

In my work (with various collaborators) I have been most interested in domain-initial articulatory strengthening, that is, strengthening associated with the beginnings of prosodic domains. I have put forward a specific claim about how this strengthening works through the whole of prosodic structure: that it is cumulative, in the sense that the higher in the prosodic tree an initial position is, the stronger that position and the segment in it. The empirical support for this claim is somewhat mixed, as I will make clear below, but there is an interesting range of data that seem to work this way.

The prosodic positions we compare are generally edges of domains. A domain-initial segment (or syllable) is at the beginning of some prosodic domain. A domain-final segment (or syllable) is at the end of some prosodic domain. Because prosodic domains are hierarchically organized, a given segment (or syllable) is usually initial or final in multiple (nested) domains. To see how this works for domain-initial syllables, look at the six syllables shown in the partial tree in Figure 1. The first is initial in all the domains shown. We generally refer to the *highest* domain in which some segment or syllable is initial, so in this case the syllable is IP-initial. The second and third syllables are word-initial, but IP- and ip-medial. The other syllables are not initial in any domains. The last syllable is final in all of the domains shown.

Experimentally, we measure something we take to be related to segment strength, for example linguo-palate contact as a measure of oral constriction. At UCLA we have primarily used electropalatography (EPG) to infer the strength of segment articulation. With the Kay Elemetrics EPG system, a speaker wears a custom-made false palate embedded with 96 contact electrodes. When the tongue touches any electrodes, a circuit is completed, current flows, and the contact is thereby registered. Figure 2 shows the electrode layout, which concentrates electrodes around the inner tooth surfaces and thus registers variation in tongue height. A computer samples the contact over the entire palate every 10 msec, and each frame of data shows which electrodes were contacted at that time. Figure 3 shows a sample frame of contact. Our general method is to construct speech materials that put a test consonant into different prosodic positions, and then take the simplest measure of strength, namely the maximum amount of contact between tongue and palate found during that consonant in each condition. This measure ignores where on the palate, and when during the consonant, this peak contact occurs, but those aspects can be measured as follow-ups. For stop consonants, we also generally measure the duration of the stop seal, that is, the amount of time (in number of data frames) that the vocal tract is completely sealed off by the stop occlusion.

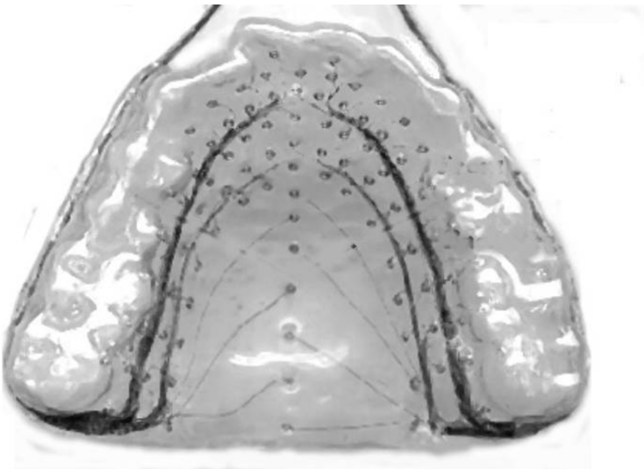

Figure 2. Sample pseudo-palate used with Kay Elemetrics Palatometer, showing arrangement of 96 contact electrodes. The front of the mouth is at the top of the picture.

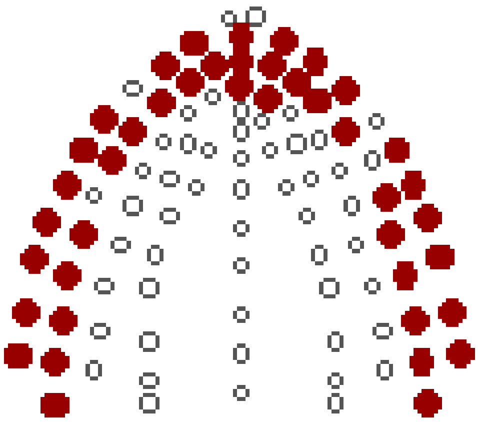

Figure 3. Sample frame showing peak linguopalatal contact for a Korean word-initial /n/. The orientation is the same as in Figure 2. Each circle is an electrode on the palate; filled circles have been contacted by the tongue (here, 42% contacted).

Our first study, Fougeron & Keating (1997), looked at English /n/ and /o/ in initial and final positions in several domains. Three American speakers read reiterant versions of sentences, using the reiterant syllable /no/ to replace every syllable of the sentences. The sentences were arithmetic expressions in which the use of parentheses was crucial to the meanings of the expressions. One speaker produced a larger set of sentences in which different numerals occurred in the sentences, but the other two speakers produced only sentences with the numeral “89”. This numeral was chosen because its lexical stress is generally on the last syllable, and thus would not be a factor in the articulation of the initial syllable. The prosodic phrasing of the utterances was coded post-hoc, and the prosodic position of each reiterant syllable was determined from these phrasings. The domains coded were the Utterance, the IP, the ip (or PP for Phonological Phrase, for typographical clarity), the Word, and the Syllable.

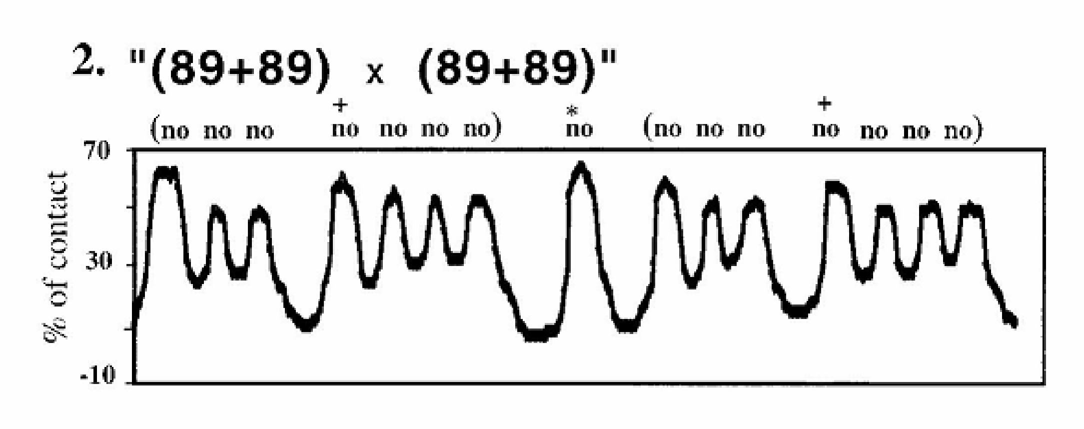

Figure 4. Sample trace of percent contact over time for a reiterant utterance. The reiterant utterance is modeled on the arithmetic expression shown at the top of the figure; the 15 reiterant syllables are indicated above the trace. High values of contact occur for /n/ while low values of contact occur for /o/.

The percent contact was calculated for each frame of data; Figure 4 shows such data for a sample utterance. This figure illustrates that not only the consonants but also the vowels vary in contact across the utterance. The measured contacts for each segment are the values of the consonant peaks and vowel valleys as seen here. The temporal pattern can also be seen (though our duration measure involves inspection of individual frames to verify that there is a complete occlusion; no duration measures are taken from these contact profiles). Articulatory seal duration was not measured or reported in Fougeron & Keating (1997) but was measured subsequently and reported in Keating, Cho, Fougeron, & Hsu (2003). Figure 5 shows the results for both measures (peak contact on the left, seal duration on the right). This figure, and the next one, shows the mean peak contact for all /n/s initial in each prosodic domain, laid out like a prosodic hierarchy, with the highest domains at the top of the figure. If the prosodic positions have a cumulative effect on a measure, then the dark bars at the top will be longest and the light bars at the bottom will be shortest. And overall, this is what the figure shows, though not completely consistently.

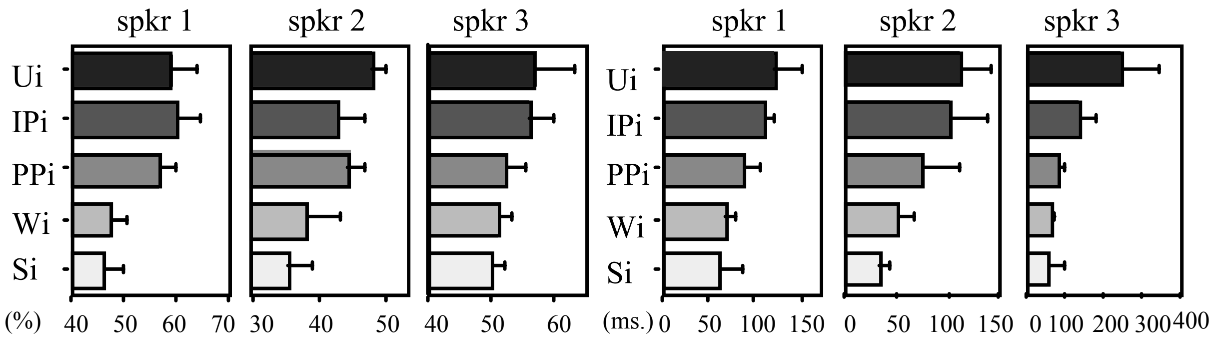

Figure 5. Mean results for three speakers of English: peak contact on the left, seal duration on the right, for /n/s initial in five prosodic domains (Utterance-initial Ui, Intonational Phrase-initial IPi, Phonological Phrase-initial PPi, Word-initial Wi, Syllable-initial Si). From Keating et al. (2003).

The original hypothesis motivating this study did not involve domain-initial strengthening; we expected to see some kind of final or progressive weakening (*declination*), as discussed in detail in Fougeron & Keating (1997). Yet many explicit tests for declination failed to provide any evidence for it. What we found instead in our data for /n/ was fairly cumulative domain-initial strengthening: each speaker made three or four pairwise cumulative distinctions between /n/s in domain-initial positions. However, no speaker made all possible distinctions, and there was no pairwise distinction that was made by all three speakers. There was also a trend towards strengthening (meaning less contact) of the vowels in initial /no/ syllables; and strengthening of /n/ in domain-final syllables was occasionally found. Domain-final vowel strengthening was observed, but it was not as strongly cumulative, in that phrase-final /o/s at different levels were not distinguished by amount of opening. Also, acoustic duration of /o/ did not vary much according to domain. In contrast, acoustic and articulatory durations of /n/ were more consistently cumulative than linguopalatal contact; nonetheless, correlations between /n/ durations and contact were minimal to modest.

We then followed up with a study of domain-initial strengthening in three other languages which differ in their prosodic properties: French, Korean, and Taiwanese (Keating et al., 2003). None have lexical stress – French and Korean have phrasal tone patterns, Taiwanese has lexical tones. In this study we used real-word utterances rather than reiterant speech, with test words beginning with the consonants /n/ or /t/, but similar prosodic domains as in the English study[1]. Overall, each language showed cumulative initial strengthening – as in English, contact and duration are generally greater in higher prosodic positions. In fact, the surprising result was how similar the results were for the three languages, despite their prosodic differences. The pattern was most consistent for Korean, which we had predicted could show the most articulatory strengthening, as its domain beginnings are generally thought to be prosodically strong. Figure 6 shows the results for Korean /n/; /t/ is similar. Here the smaller phrase is the Accentual Phrase (AP).

[1] Two points deserve mention in this connection. First, there is no reason that the prosodic domains should be exactly comparable across the languages, the small phrases in particular seem to vary across languages. Second, the speakers of the different languages differed in their use of pauses; the French speakers sometimes paused at IP boundaries, while the Korean and Taiwanese speakers were instructed not to do so.

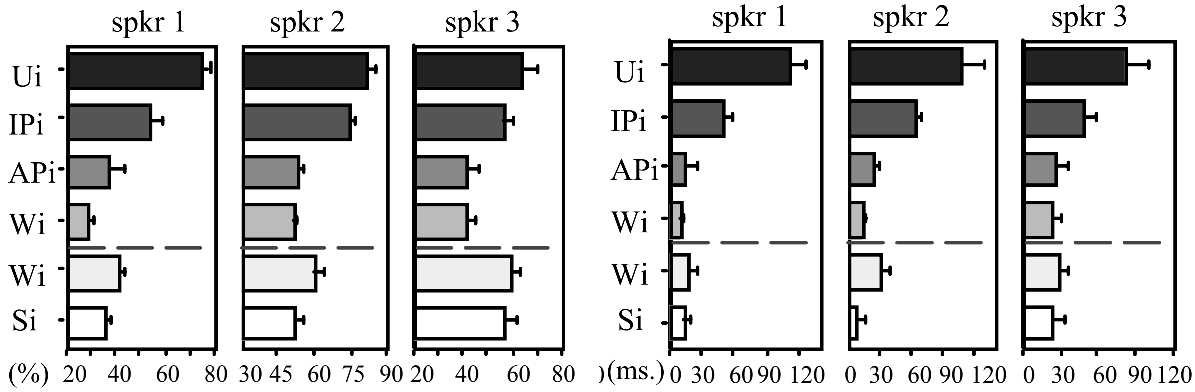

Figure 6. Mean results for three speakers of Korean: peak contact on the left, seal duration on the right, for /n/s initial in five prosodic domains (Utterance-initial Ui, Intonational Phrase-initial IPi, Accentual Phrase-initial APi, Word-initial Wi, Syllable-initial Si). The dashed horizontal line divides one set of comparisons, of levels above the Word, from another comparison, Word vs. Syllable. From Keating et al. (2003).

A further result concerned the relation of duration to linguopalatal contact in the different languages: while all the languages had fairly consistent cumulative initial lengthening, only in Korean was strengthening (contact) strongly related to that lengthening. In our English data, the correlations of contact with duration were low to modest ($r^2 < .3$); in French there was a stronger relation ($r^2 > .6$); but in Korean the correlations were very high ($r^2 < .9$). This strong relation suggests a sort of undershoot mechanism. Strengthening in Korean seems to be related to how much time is available for the articulation: in Cho & Keating (2001) we showed that up to about 80 msec, the amount of contact is a function of the time, with the peak contact coming at the end of the consonant and shorter consonants undershooting their target; but above about 80 msec, there is no additional contact. Thus it seems that in Korean, there is little if any independent effect of strengthening apart from lengthening. However, to the extent that the other languages are not like this, they in turn provide evidence that initial strengthening is a separate effect from lengthening – two effects, but both sensitive to prosodic position. Further support for the independence of strengthening and lengthening comes from Byrd et al. (2000), a study of Tamil that found lengthening of certain initial consonants without greater spatial displacement.

Other researchers have contributed to our knowledge about initial strengthening in a variety of languages, including Gordon (1999), Lavoie (2001), Tabain (2003), and Tabain et al. (2003). Most studies that have included several prosodic domains have found an overall tendency, but not a perfect pattern, of cumulative domain-inital strengthening. Exceptions are Byrd et al. (2000) on Tamil, mentioned above, which found no effect of prosodic position on the extent of articulation, and Tabain (2003), which likewise found no effect of prosodic position on the peak displacement of French domain-initial consonants / b d g f s ʃ /. And, Byrd & Saltzman (1998) found a very different result for English than we did: comparing lip movements at the boundaries of what were probably three different prosodic domains, they found that displacement to the postboundary consonant was highly correlated with duration[2]. It seems likely that displacement, the measure used in these EMA studies, does not depend on prosodic position in the way that peak linguopalatal contact does. Finally, in an EPG study of Japanese by Onaka (2003), only one of two speakers showed only a tendency to cumulative strengthening, a weaker result than in the languages we have studied.

FURTHER RESULTS

Further observations can be made about domain-initial strengthening on the basis of additional analyses in Fougeron & Keating, as well as later studies at UCLA: on English, Cho (2002, 2004), Keating, Wright & Zhang (1999); on Korean, Cho & Keating (2001), Cho & Jun (2000), and Kim (2001); and on French, Fougeron (1998, 2001).

First, some segments vary more than others; indeed some segment types show no initial strengthening. For example, Fougeron (1998) looked at EPG contact for French /n t k l s i/, and found that prosodic position had less effect on contact for /s/ and /i/ than for the other segments. For the sibilant fricative /s/, Fougeron measured several aspects of the fricative constriction, not just overall contact. By all measures, it varied very little across prosodic positions, presumably because the production of sibilance constrains the articulation. That fricative /s/ is highly constrained in its articulation is no surprise, and the same has been shown at the word level for English, e.g. Byrd (1996) and Keating et al. (1999). However, Kim (2001) did find differences in Korean /s/, specifically in the contact in the mid-palate and the fricative channel region, across three prosodic positions, showing that /s/ in at least one language is free to vary. As another example, Cho & Keating (2001) compared the four coronal stops of Korean, and while all showed an effect of prosodic position on contact, the range of variation differed across the consonants, such that the prosodic effect was larger for some than for others.

Second, domain-initial strengthening appears to be a very local effect largely limited to the first segment after a boundary, and is thus unlike final lengthening, which extends over a larger span. For example, in the French study (Fougeron 1998, 2001) domain-initial strengthening was limited to only the /k/ in a /kl/ cluster, and to a vowel only when there is no preceding initial consonant (/i/ in /#ip/ but not /a/ in /#na/). This very local effect is perhaps consistent with the fact that in French, final lengthening is more limited in extent than in English (Fletcher, 1991). Still, in English a vowel in #CV varies very little with prosodic domain (Fougeron & Keating, 1997; Cho, 2002).

Third, it is noteworthy that strengthening does not result in discrete phonetic categories, corresponding to the domains of the prosodic hierarchy, even though descriptions of prosody are couched in terms of these discrete domains. In Cho & Keating (2001) we compared pooled measures of acoustic domain-final lengthening

---

[2] This result might seem to suggest that English is in fact like Korean, with a relation between strengthening and lengthening, contrary to our own result. However, our correlations are with peak contact, not displacement; peak contact is more closely related to constriction degree, and in Fougeron & Keating's analysis, is not strongly related to displacement (as measured by EPG).

and EPG domain-initial contact, as seen here in Figure 7. In statistical comparisons, both sets of measures support a four-way prosodic distinction. However, only the data for final duration clearly fall into discrete categories, and only two of those; the data for initial contact belong to a single large unimodal distribution. That is, it cannot be claimed that there are four (or however many) categories of strength, each of which gets some additional increment of constriction; the effect appears instead to be continuous. Our strict prosodic categories are a marker, but not a trigger, of a scalar effect.

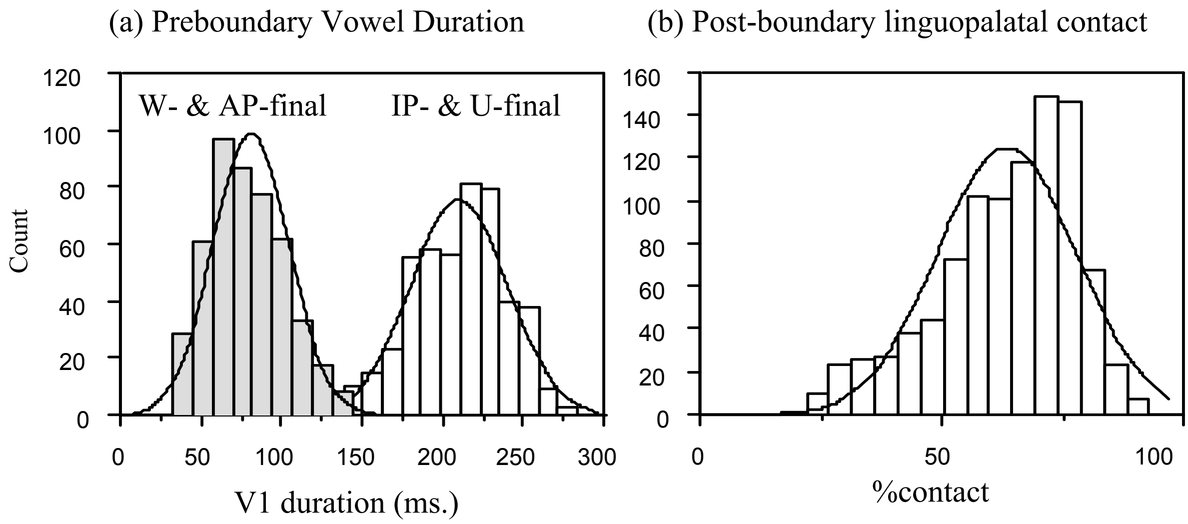

Figure 7. Histograms of all measurements of (a) domain-final vowel duration and (b) domain-initial consonant contact, pooled across consonants and speakers. From Cho & Keating (2001).

Finally, it is worth stressing that Fougeron & Keating (1997) were careful to distinguish *initial strengthening* from *final weakening* and *declination*. While these terms might sound as if they are different names for the same thing (e.g. some sort of downtrend), they in fact describe different outcomes when a large enough span of speech is considered. Fougeron & Keating called their effect *initial strengthening* because of explicit tests that favored that interpretation. Simple comparisons of two positions cannot decide this point. For example, a comparison of initial vs. final positions by itself cannot distinguish these three possibilities, and much of the literature compares only two positions in this way. In such cases the terms *initial strengthening*, *final weakening*, and *declination* all come to the same thing, and no importance can be placed on the choice of descriptive term.

Fougeron & Keating did not look at domain-final consonants, since all their test syllables were CVs. Given the historical linguistics literature, is plausible that domain-final consonants should show some weakening. In Keating et al. (1999) we made a very limited comparison of domain-initial and domain-final consonants / t d n l /; in the test corpus, word-initial consonants occurred utterance-initially vs. utterance-medially, while word-final consonants occurred utterance-medially vs. utterance-finally. The maximum EPG contact depended on position in both the word and the utterance. Overall, as expected, word-initial consonants had more contact than word-final, and also as expected, word-initial consonants had more contact when they were also at the beginning of an utterance. However, utterance-final consonants had more contact than other word-final consonants. That is, there appears to be no cumulative domain-final weakening of consonants; instead we see some strengthening at the end of the largest domain. The role of domain-final lengthening in this apparent strengthening deserves further study.

ACOUSTIC AND OPTICAL CORRELATES

The kinds of articulatory variation discussed above give rise to two kinds of potentially perceivable variation: acoustic and optical. Some production studies have included acoustic measures, e.g. Cho & Keating (2001) on Korean. An interesting example of prosodically-conditioned acoustic variation, involving the glottis, is seen in Hsu & Jun's (1998) study of Taiwanese VOT. Taiwanese has voiced (often prenasalized), voiceless unaspirated, and voiceless aspirated stops. Hsu & Jun found that when /k^h/ and /b/ are initial in higher domains, the /k^h/ is more aspirated and the /b/ is more voiced. Thus the phonetic voicing categories are acoustically more distinct in stronger positions.

The optical correlates of prosody have been studied very little to date, mainly the correlates of prominence, but since many articulations involve the face, at least some aspects of phrasing should also be visible. Certainly durational differences should be apparent, even if subtler articulatory differences are difficult to see. A study with colleagues at the House Ear Institute in Los Angeles (Auer et al., 2004) is testing the visual perceptibility of prosodic boundaries. Talkers read minimal pairs of sentences differing in the presence/absence of boundaries, while movements of reflective markers on their faces were tracked by the Qualysis infrared system. Some sentences contrasted in presence/absence of a word boundary (e.g. *He was noted for getting the right number* vs. *He just hated forgetting the right number*), while others contrasted in presence/absence of a phrase boundary (e.g. *When you sing, his songs are better* vs. *When you sing his songs, they're better*). The duration of the movement of the marker on the chin was measured, and were found to be longer around a boundary. Acoustic durations were also longer across a boundary. Perceivers then saw sentence fragments and had to decide which sentence (with or without boundary) the fragment had come from. Perceivers were generally good at doing this.

IMPLICATIONS OF STRENGTHENING

In this section I consider two functional implications of domain-initial strengthening. First, because it depends on phrasing, segment strength may convey information about the local coherence vs. disjuncture in connected speech. A strengthened segment indicates a break and the start of a new domain, while domain-internal spans of segments are not interrupted by strengthening. On this view, speakers manipulate strength to indicate the degree of break/cohesion between words in connected speech. While stronger segments may consume more calories of a speaker's articulatory energy than do weaker ones, weaker segments do not reflect laziness or inattention on the part of the speaker. Rather, we should think of the speaker's energy as constantly directed to control of the modulation of articulation, because all levels of strength carry information.

Second, strengthening may increase information about segment identity, and in just those positions where such information is most important. Psycholinguists (e.g.

Levelt et al. 1999) have noted the special status of word-initial position; word-initial segments

●are vulnerable to speech errors

●are exchanged in speech errors

●show strong similarity effects in speech errors

●are easy to detect with speech errors

●are vulnerable to mis-hearings

●are preserved in TOT (tip of tongue) state

●are accurate in phoneme monitoring

And these word-initial segments are produced with articulatory strengthening. Is this a paradox? Phonetically, word-initial position is said to be "strong", yet much of the psycholinguistic evidence concerns errors in production. If these segments are strong, why do they seem so vulnerable? A possible resolution to this quandary can be found in the contributions of Frisch (2000) and Dell (2000) to the *Laboratory Phonology V* volume, in which they suggest that speech errors arise in word- (or in most such experiments, utterance-) initial position due to lack of a constraining prior context. Because other words are possible, other word candidates are activated and compete with the correct word. Furthermore, word-initial position is a position of competition between many competitors, in the sense that words generally can begin with more segments than they end with, such that there are more (different) segments in strong competition word-initially than elsewhere. That is, word-initial segments are more vulnerable to errors because there are more possibilities when context does not provide strong constraints.

This explanation seems comparable to a suggestion by Fougeron & Keating (1997) concerning initial strengthening. They noted that, from the perspective of the listener, initial segments are less determined by prior context, and that therefore the acoustic signal must bear a greater load in the recovery of the message in those positions. Initial strengthening could thus help the listener by enhancing segmental properties in positions of uncertainty[3]. Thus the resolution of the paradox would be that initial segments are contextually weak, that is, relatively unconstrained by their prior context. Because of this contextual weakness they are more vulnerable to competition from other lexical entries in the process of speech production, and for the same reason they are more vulnerable to mis-hearing in speech perception.

Finally, as shown earlier, phonetic strengthening occurs in higher prosodic domains. The larger the phrasal domain, the more likely is the initial position to be

[3] It must be noted that the hypothesis that strengthening serves an enhancing function for listeners is controversial. Various views on the nature and possible function of segmental enhancement in domain-initial positions can be found in, for example, Pierrehumbert & Talkin (1992), Fougeron & Keating (1997), Hsu & Jun (1998), Fougeron (1999, 2001), and Cho & McQueen (2004).

unconstrained by context (for example, the first segment of a sentence is far less predictable than the first segment of most words within a sentence). If contextual uncertainty is indeed the connection between the psycholinguistic phenomena and phonetic strengthening, then we would also predict that speech errors and mis-hearings should be more frequent in initial positions of higher phrasal domains. As Keating & Shattuck-Hufnagel (2002) review, errors across phonological words are far more common than errors within words, so this prediction is plausible.

CONCLUSION

This review of findings about domain-initial articulatory strengthening counters the view, espoused by Levelt et al. (1999), that segmental and prosodic planning for speech production can proceed separately. Since phonetic encoding of segments is highly sensitive to prosodic structure, prosody needs to be computed first, not last as in Levelt's model. Returning to the example utterance in Figure 1 ("that new propaganda"), consider the effects of prosodic position on the feature values indicated. The word "that" is in the strongest position, initial in the highest domain, so the initial continuant consonant is likely to be strengthened to a stop articulation. The word "propaganda" is in a strong position because of its prominence, which will affect the stressed syllable "gan". The /p/ at the beginning of the word should be somewhat stronger than the /p/ in the second syllable, and thus should have a closer oral consonant constriction, while its glottal abduction should be larger. Phonetic plans for those, and all other, aspects of the utterance must refer to prosody.

How might such prosody-dependent phonetic encoding of features be modeled? One possibility would be Byrd et al. (2000)'s π-gestures, by which the prosody could modulate precompiled syllable scores. Another possibility is a window-style model, in which prosody could modulate articulatory targets. Window models (Keating 1990, 1996) posit ranges, rather than fixed points, as the targets of articulatory movements. Guenther (1995) first suggested that such target ranges could be sensitive to prosody, expanding or contracting over the course of an utterance. At an edge or a prominence, target ranges would shift towards extreme values. While this proposal remains to be worked out, it has the potential advantage of extending readily to other kinds of variability, on other timescales, that are not accounted for by fixed windows, or by gestures associated with prosodic boundaries. In addition to shifts that are local to a prosodic position, target ranges could shift at the word level, as a function of lexical difficulty due to competion, and more globally, as a function of discourse factors.

In sum, when a speaker plans for the phonetic aspects of speech production, prosodic structure organizes the treatment of possibly every feature in every segment, and the interactions of segments. One aspect of this dependence is the relation between the strength of a prosodic position, and the phonetic strength of a segment in that position. A theory of phonetic encoding that incorporates this basic fact is a major challenge, but an important one.

ACKNOWLEDGEMENTS

This research was supported by an NSF Linguistics grant to P. Keating (#95-11118), an NSF dissertation improvement grant to T. Cho and P. Keating (#BCS-0001716), and an NSF KDI grant to L. Bernstein et al. (#99-96088). Collaborations with all my co-authors on the various papers cited here, and the contributions of the former students cited, are gratefully acknowleged.

REFERENCES

Auer, E.T. Jr., Kim, S., Keating, P., Scarborough, R.A., Alwan, A., & Bernstein, L.E. (2004, November). Optical Phonetics and Visual Perception of Lexical and Phrasal Boundaries in English. Paper to be presented at the meeting of the Acoustical Society of America, San Deigo.

Beckman, M.E., Edwards, J. & Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. In G. J. Docherty & D. R. Ladd (Eds.), *Papers in Laboratory Phonology II: Gesture, segment, prosody* (pp. 68-86). Cambridge, UK: Cambridge University Press.

Bell, A. & Bybee Hooper, J. (1978). Issues and evidence in syllabic phonology. In A. Bell & J. Hooper (Eds.), *Syllables and Segments* (pp. 3-24). Amsterdam: North-Holland Publishing.

Byrd, D. (1996). Influences on articulatory timing in consonant sequences. *Journal of Phonetics, 24,* 209-244.

Byrd, D. & Saltzman, E. (1998). Intragestural dynamics of multiple prosodic boundaries. *Journal of Phonetics, 26,* 173-200.

Byrd, D., Kaun, A., Narayanan, S. & Saltzman, E. (2000). Phrasal signatures in articulation. In M. B. Broe & J. B. Pierrehumbert (Eds.), *Papers in Laboratory Phonology V: Acquisition and the Lexicon* (pp. 70-87). Cambridge, UK: Cambridge University Press.

Cho, T. (2002). *The Effects of Prosody on Articulation in English*. New York: Routledge.

Cho, T. (2004). Prosodically-conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics, 32,* 141-176.

Cho, T. & Jun, S.-A. (2000). Domain-initial strengthening as enhancement of laryngeal features: Aerodynamic evidence from Korean. In J. Boyle, J-H. Lee, & A. Okrent (Eds.), *Chicago Linguistic Society, 36,* 31-44. [Also in *UCLA Working Papers in Phonetics, 99,* 57-69.]

Cho, T. & Keating, P. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics, 29*, 155-190.

Cho, T. & McQueen, J. (2004) Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics*, to appear.

de Jong, K. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America, 97,* 491-504.

Dell, G. (2000). Commentary: Counting, connectionism, and lexical representation. In M. B. Broe & J. B. Pierrehumbert (Eds.), *Papers in Laboratory Phonology V: Acquisition and the Lexicon* (pp. 335-348). Cambridge, UK: Cambridge University Press.

Dilley, L., Shattuck-Hufnagel, S. & Ostendorf, M. (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics, 24*, 423-444.

Edwards, J., Beckman, M.E. & Fletcher, J. (1991). The articulatory kinematics of final lengthening. *Journal of the Acoustical Society of America, 89,* 369-82.

Epstein, M. (2002). *Voice Quality and Prosody in English*. Unpublished UCLA dissertation.
Epstein, M. (2003). Voice quality and prosody in English. *Proceedings of the 15th International Congress of Phonetic Sciences,* 2405-2408.
Erickson, D. (2002). Articulation of Extreme Formant Patterns for Emphasized Vowels. *Phonetica, 59*, 134-149.
Fletcher, J. (1991). Rhythm and final lengthening in French. *Journal of Phonetics, 19,* 193-212.
Fletcher, J. & Vatikiotis-Bateson, E. (1994). Prosody and intrasyllabic timing in French. *Ohio State University Working Papers, 43,* 41-46.
Fougeron, C. (1999). Prosodically conditioned articulatory variations: a review. *UCLA Working Papers in Phonetics, 97,* 1-74.
Fougeron,C. (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics, 26,* 45-69.
Fougeron, C. & Keating, P. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America, 101*, 3728-3740.
Frisch, S. (2000). Temporally organized lexical representations as phonological units. In M. B. Broe & J. B. Pierrehumbert (Eds.), *Papers in Laboratory Phonology V: Acquisition and the Lexicon* (pp. 283-298). Cambridge, UK: Cambridge University Press.
Gordon, M. (1999). The effect of stress and prosodic phrasing on duration, acoustic amplitude and air flow of nasals in Estonian. *UCLA Working Papers in Phonetics*, *92,* 151-159.
Guenther, F. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review, 102*, 594-621.
Harrington, J., Fletcher, J. & Beckman, M. (2000). Manner and place conflicts in the articulation of Autralian English consonants. In M. B. Broe & J. B. Pierrehumbert (Eds.), *Papers in Laboratory Phonology V: Acquisition and the Lexicon* (pp. 40-51). Cambridge, UK: Cambridge University Press.
Hock, H. H. (1991). *Principles of Historical Linguistics* (2nd ed.). Berlin: Mouton de Gruyter.
Keating, P. (1990). The window model of coarticulation: articulatory evidence. In J. Kingston & M. Beckman (Eds.), *Papers in Laboratory Phonology I* (pp. 451-470). Cambridge, UK: Cambridge University Press.
Keating, P. (1996). The Phonology-Phonetics Interface. In U. Kleinhenz (Ed.), *Interfaces in Phonology (*pp. 262-278). *Studia grammatica, 41*. Berlin: Akademie Verlag.
Keating, P., Cho, T., Fougeron, C. & Hsu, C-S. (2003). Domain-initial strengthening in four languages. In J. Local, R. Ogden & R. Temple (Eds.), *Phonetic Interpretation: Papers in Laboratory Phonology VI* (pp.143-161). Cambridge, UK: Cambridge University Press.
Keating, P. & Shattuck-Hufnagel, S. (2002). A Prosodic View of Word Form Encoding for Speech Production. *UCLA Working Papers in Phonetics, 101,* 112-156.
Keating, P., Wright, R., & Zhang, J. (1999). Word-level asymmetries in consonant articulation *UCLA Working Papers in Phonetics, 97***,** 157-173.
Kim, S. (2001). Domain initial strengthening of Korean fricatives /s/ and /s*/. *Harvard Studies in Korean Linguistics, IX,* 164-173.

Lavoie, L. (2001). *Consonant strength: Phonological patterns and phonetic manifestations*. New York: Garland.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.

Levelt, W. J. M. (1989). *Speaking*: *From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W. J. M., Roelofs, A. & Meyer, A. S. (1999). A theory of lexical access in speech production. *Brain and Behavioral Sciences, 22*,1-38.

Martinet, A. (1955). *Economie des changements phonétiques: traité de phonologie diachronique.* Berne: Francke SA.

Onaka, A. (2003). Domain-initial strengthening in Japanese: An acoustic and articulatory study. *Proceedings of the 15th International Congress of Phonetic Sciences,* 2091-2094.

Pierrehumbert, J. & Talkin, D. (1992). Lenition of /h/ and glottal stop. In G. J. Docherty & D. R. Ladd (Eds.), *Papers in laboratory phonology II: Gesture, segment, prosody* (pp. 90-117). Cambridge, UK: Cambridge University Press.

Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research, 25*, 193-247.

Tabain, M. (2003). Effects of prosodic boundary on /aC/ sequences: articulatory results. *Journal of the Acoustical Society of America, 113***,** 2834-2849.

Tabain, M., Perrier, P. & Savariaux, C. (2003). A kinematic study of prosodic boundary effects on /i/ articulation in French. *Proceedings of the 15th International Congress of Phonetic Sciences,* 2617-2620.

Wightman, C., Shattuck-Hufnagel, S., Ostendorf, M. & Price, P. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America, 91*, 1707-17.