THE PHONETICS AND PHONOLOGY OF NON-MODAL VOWELS: A CROSS-LINGUISTIC PERSPECTIVE

MATTHEW GORDON
University of California, Los Angeles

This paper provides phonetic explanations for a number of asymmetries in the distribution of non-modal voiced and modal voiced vowels found in a typology of over 50 languages with non-modal vowels. The distribution of non-modal vowels cross-linguistically is argued to result from the conflicting demands of perceptual salience and articulatory ease; this conflict can be formally modeled in an constraint-based grammar.*

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

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

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

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

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

First, non-modal vowels are characterized by less overall acoustic intensity than modal vowels as shown for different non-modal phonation types, e.g. breathy vowels in Kui and Chong (Thongkum 1987), creaky vowels in Chong (Thongkum 1987) and voiceless and creaky vowels in Hupa (Gordon 1998). Decreasing the acoustic intensity results in a decrease in loudness, the auditory correlate of intensity; it thus follows that non-modal vowels are less salient than modal vowels.

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

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

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

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

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

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

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

3.1. THE STATUS OF NON-MODAL VOICING: PHONETIC OR PHONOLOGICAL. Before preceding with the typology, it is appropriate to address the question of whether non-modal voicing is a phonological or phonetic phenomenon or perhaps both, depending on the language. Clearly in languages with an underlying contrast between non-modal voiced vowels and modal voiced vowels (e.g. Sedang, Gujarati, Jalapa Mazatec), non-modal vowels are a synchronic phonological feature[1]. However, as pointed out earlier, the number of languages with underlying or even surface contrastive phonation type for vowels is quite small. In the majority of languages in which they occur, non-modal vowels are a surface non-contrastive property, and thus less clearly belong to the phonology.

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

On the other hand, there are many languages in which vowel devoicing behaves like a phonological phenomenon. In some languages, voiceless vowels contrast on the surface with voiced vowels. For example, the word-internal contrast between short and long vowels is realized as a contrast between voiceless and voiced vowels in word-final position in Oromo and in Woleaian. Similarly, in Hupa as discussed in section 2, vowel devoicing is contrastive before many syllable-final consonants. If we adopt the standard assumption that contrastive properties are phonological, vowel devoicing would clearly fall under the purview of phonology in Oromo, Woleaian and Hupa. Furthermore, in other languages, vowel devoicing interacts with other phenomena which are typically assumed to be phonological. For example, vowel devoicing influences stress assignment in Awadhi, pitch accent placement in Tunica, and debuccalization and tone shift in Comanche. Furthermore, in Tongan, one of the prerequisites for vowel devoicing is that vowels be in morpheme-final position; such morphological conditioning would suggest that vowel devoicing is not merely a low level phonetic phenomenon. In summary, vowel devoicing thus appears to operate at a relatively deep level of the grammar in a fair number of languages[3].

Perhaps not surprisingly, languages in which vowel devoicing clearly appears to be phonological display vowel devoicing in the same environments (e.g. domain finally, adjacent to voiceless consonants) in which vowel devoicing is most likely to occur in languages where it has been demonstrated to be gradient. Even if one assumes a sharp distinction between vowel devoicing as a phonetic process vs. devoicing as a phonological one, the striking similarity between the distributions of phonetic and phonological vowel devoicing suggests that examination of phonetic devoicing may also provide insight into phonological devoicing. For this reason, the typology in this paper includes cases of vowel devoicing which are clearly phonological as well as others which may not be This paper will not address the issue of where to draw the line between phonological and phonetic processes. Crucially, because substantially the same phonetic factors condition devoicing in all languages, many aspects of the analysis of devoicing are more likely than not to be quite similar for all languages with voiceless vowels.

3.2. THE LENGTH ASYMMETRY. Of the 32 languages in the survey with contrastive vowel length in environments targeted by devoicing, devoice short but not long vowels occurring in the same environment (in many languages, only high vowels devoice; see section 3.4): e.g. Awadhi, Big Smokey Valley Shoshoni, Bulu, Mbay, Cheyenne, Cocama, Gadsup, Galla, Ik, Inuit, Oneida, Goajiro, Tarascan, Zuni, Japanese, Kawaiisu, Mokilese, Sámi, Sara, Shina, Bagirmi, Tongan, Tubu, Tunica, Turkish, Woleaian). Four languages (Boraana Oromo, Papago, Southern Paiute, Ket[4]) possess voiceless short vowels and also devoice a portion of long vowels in certain environments[5], one (Acoma) possesses voiceless short vowels and long vowels which are completely voiceless, and one language (Hupa) has partially voiceless long vowels but lacks voiceless short vowels. In one, Cheyenne, long vowels partially devoice in final position but do not devoice at all in word-medial environments in which short vowels devoice.

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

Up to this point, I have been intentionally vague in defining “final position”. The reason for this is that the domain of devoicing varies from language to language; however, these domains follow an implicational hierarchy. Devoicing in final position of a smaller domain (e.g. word) implies devoicing in larger domains (e.g. phrase, utterance); the reverse of this statement is not necessarily true. In 20 languages in the survey (Ik, Dafla, Cocama Galla, Bagirmi, Turkana, Sara, Tubu, Mbay, Malagasy, Campa, Tarascan[6], Ticuna, Ket, Ainu[7], Island Carib, Zuni, Washkuk, Goajiro, Woleaian[8]), voiceless vowels occur word-finally (and of course, by implication, finally in larger domains as well). In 14 languages, devoicing is characteristic only of final position of larger domains, e.g. phrase or utterance (Alabama, Papago, Greek, Tarascan, Totonac, Chontal, Gadsup, Oneida, Apinaye[9], Mixtec, Nyangumata[10], Boraana Oromo, Cheyenne, Kawaiisu[11]). Note that, from most descriptions, it is impossible to make distinctions among larger domains such as the phonological phrase, intonational phrase or utterance.[12] Interestingly, only one language, Cocama, regularly devoices initial vowels in addition to final vowels; in both environments, devoicing only affects vowels adjacent to a voiceless consonant.

After final position, the next most common position in which vowels devoice is adjacent to voiceless consonants. Word-medial devoicing is found in 19 languages in the survey (Mandarin, Brazilian Portugese, Malagasy, Mixtec, Quechua, Goajiro, Azerbaijani, Inuit, Chontal, Montreal French, Cheyenne, Mokilese, Big Valley Shoshoni, Japanese, Turkish, Korean, Tongan, Cocama, Papago). It is interesting to note that devoicing of final vowels in most languages (29 of 36), occurs not only after voiceless but also after voiced consonants[13]. In only 6 languages with final devoicing (Japanese, Korean, Tongan, Turkana, Cocama, Mixtec) must the vowel both be final *and* next to a voiceless consonant for devoicing to occur[14]. In 8 languages with word-medial devoicing it is sufficient to have a voiceless consonant on only one side of a vowel to trigger devoicing. In 5 of these languages, the triggering consonant is on the right side of the vowel (Big Valley Shoshoni, Comanche, Southern Paiute, Goajiro, Quechua), in 3, it is on the left side (Acoma, Mandarin, Chontal). In 9 languages with word-medial devoicing (Cheyenne, Mokilese, Japanese, Turkish, Korean, Montreal French, Tongan, Papago, Malagasy), devoicing is described as affecting vowels (almost) exclusively between two voiceless consonants. Further asymmetries between different voiceless consonants will be discussed in section 4.

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

3.5. THE STRESS AND TONE/INTONATION ASYMMETRY. Two other asymmetries in vowel devoicing relate to the closely related properties of tone and accent. In all languages in the survey for which data on accent location is reported and in which the other necessary preconditions for devoicing are present, accented vowels resist devoicing (Montreal French, Turkish, Tongan, Comanche[17],Cheyenne, Brazilian Portugese, Azerbaijani, Quechua). In keeping with this pattern, in Papago, the set of environments in which stressed vowels devoice is a subset of those in which unstressed vowels devoice. I know of no language with devoicing of stressed but not unstressed vowels.

A final asymmetry is that many tone and pitch accent languages fail to devoice high-toned vowels (Japanese[18], Cheyenne, Acoma[19]) Furthermore, in some languages with stress, intonational pitch accents (Greek, Boraana Oromo, Tunica) and high boundary tones (Inuit) can inhibit devoicing. The stress and tone/intonational asymmetries are presumably closely related since accented syllables often carry high pitch accents cross-linguistically, as in Japanese.

4. AN ARTICULATORY ACCOUNT OF VOWEL DEVOICING. The asymmetries in vowel devoicing discussed can, in large part, be explained in terms of a combination of articulatory overlap between neighboring glottal gestures and aerodynamic considerations. First, let us consider the patterns which are compatible with a gestural overlap account of devoicing. The reasoning given here basically follows

that of Dauer (1980), Jun and Beckman (1993, 1994), Jun et al. (1997, 1998). The types of vowels which devoice cross-linguistically are those which are likeliest to be produced with voicing gestures which are durationally shortest. Trivially, phonemic short vowels are phonetically shorter than phonemic long vowels. It is also well known that high vowels are shorter than non-high vowels (Lehiste 1970) and that unaccented vowels are shorter than accented ones.

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

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

Interestingly, in Korean, the asymmetry between stops and fricatives depends on whether the stop or fricative appears on the right or left side of the potential target of vowel devoicing (Jun and Beckman 1994, Jun et al. 1997, 1998). Fricatives and aspirated stops are more likely to trigger devoicing than fortis and lenis stops when they precede a vowel, an expected pattern given the greater glottal apertures of fricatives and aspirated stops. However, following a vowel, all stops, including fortis and lenis stops trigger devoicing more often than fricatives, even though fricatives have greater glottal openings than fortis and lenis stops. Jun and Beckman suggest that the closing gesture into a stop might be faster than the oral constriction gesture made for a fricative; this greater velocity of the oral closing gesture could lead to a more abrupt increase in oral pressure which could inhibit voicing in the preceding vowel. One might hypothesize that the asymmetry between Korean and those languages in which devoicing is triggered by a following fricative and not a following stop (Comanche, Goajiro, and Southern Paiute--singletons) is due to language specific differences in the relative magnitude and timing of glottal opening gestures in the two classes of consonants.

In summary, although the glottal overlap story does not account for all cases of devoicing[20], it nevertheless offer a coherent explanation for many of the devoicing asymmetries. There are two robust asymmetries, however, which do not fall out directly from a gestural overlap account without recourse to other factors. First, there is the tendency for low-toned vowels to preferentially devoice over high-toned vowels pointed out in section 3.5. The existence of this asymmetry is presumably linked to the inherent inability of voiceless segments to carry tone phonetically. Thus, devoicing both high and low toned vowels would lead to neutralization of a tonal contrast, either a lexical contrast in the case of languages with lexical tone, or a semantic contrast, in the case of sentence (or phrase-level) intonation, e.g. questions vs. statements. In order to maintain the contrast, a language could thus devoice either high toned or low toned vowels but not both. Perhaps low toned vowels devoice, because the articulatory gestures involved in producing low tone are more compatible with the glottal abduction gestures associated with devoicing. An explanation for the tone asymmetry must await instrumental research.

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

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

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

(1) **Ranking**	**Devoicing Targets**
*NON-MOD V, *NON-MOD VV >> DEVOICING	No vowels
*NON-MOD V >> DEVOICING >> *NON-MOD VV	Long vowels

*DEVOICING >> *NON-MOD V, *NON-MOD VV — Short + long vowels
*NON-MOD VV >> DEVOICING >> *NON-MOD V — Short vowels

6. SUMMARY AND CONCLUSIONS. I have argued that the distribution of non-modal vowels is governed by the conflict between perceptual demands, on the one hand, and articulatory and aerodynamic considerations, on the other hand. Although non-modal vowels are perceptually dispreferred to modal vowels, non-modal vowels may be articulatorily and aerodynamically preferred under certain conditions. The conflict between ease of perception and ease of articulation can be modeled in a constraint based grammar through different rankings of constraints.

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

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

NOTES

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

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

[2] Note that /h/ is not considered here as a voiceless vowel strictly speaking, since it lacks place features unlike "true" voiceless vowels.

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

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

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

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

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

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

[9] In Apinaye, phrase-final vowels optionally either lengthen, devoice or become creaky.

[10] Vowels may devoice before silence which I interpret to mean utterance finally.

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

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

[13] In Totonac, the only utterance final short vowels occur after voiceless consonants.

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

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

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

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

[18] High toned syllables could alternatively be considered accented syllables in Japanese.

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

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

REFERENCES

ALEXANDRE, P. 1962. Sur la voyelle suffixielle du bulu. JAL 1.243-252.

BECKMAN, M. 1982. Segment duration and the "Mora" in Japanese. Phonetica 39.113-135.

BECKMAN, M; EDWARDS, J; and FLETCHER, J. 1992. Prosodic structure and tempo in a sonority model of articulatory prominence. Papers in Laboratory Phonology II: Gesture, Segment, Prosody, ed. by G. Docherty and D. Ladd. Cambridge: Cambridge University Press.

BLANKENSHIP, B. 1997. The time course of breathiness and laryngealization in vowels. UCLA Ph.D. dissertation.

CEDERGREN. H.J. and SIMONEAU, L. 1985. La chute des voyelles hautes en français de Montreal: 'As-tu entendu la belle syncope?' Les tendencies

dynamiques du français parle a Montreal, ed. by Lemieux, M. and Cedergren, H.J. 57-144.
CHARNEY, J. 1993. A grammar of Comanche. Lincoln: U. of Nebraska Press.
CHO, Y. 1993. The phonology and phonetics of 'voiceless' vowels. Proceedings of the 19th Annual Meeting of the Berkeley Linguistics Society.
CRAPO, R. 1976. Big Smokey Valley Shoshoni. Desert Research Institute Publications in the Social Sciences 10.
CROTHERS, J.; LORENTZ, J.; SHERMAN, D.; and VIHMAN, M. 1979. Handbook of phonological data from a sample of the world's languages. Stanford.
DAUER, R. 1980. The reduction of unstressed high vowels in Modern Greek. Journal of the International Phonetics Association. 10.17-27.
DAVIS, I. 1962. Phonological function in Cheyenne. IJAL 28.36-42.
DIMMENDAAL, G. 1983. The Turkana language. Dordrecht: Foris.
FAUST, N. and PIKE, E. 1959. The Cocama sound system. Publicações do Museo Nacional, Série Linguistica Especial 1, 10-55.
FELDMAN, H. 1978. Some notes on Tongan phonology. Oceanic Linguistics 17(2).133-139.
FISCHER-JØRGENSEN, E. 1967. Phonetic analysis of breathy (murmured) vowels in Gujarati. Indian Linguistics 28.71-139.
GADEN, H. 1909. Essai de grammaire de la langue baguirmienne. Paris: Leroux.
GENDRON, J-D. 1966. Tendances phonétiques du français parlé au Canada. Québec: Les Presses de L'Université Laval.
GOLLA, V. 1970. Hupa Grammar. Ph.D. dissertation. U. of California, Berkeley.
GORDON, M. 1998. Laryngeal timing and correspondence in Hupa. UCLA Working Papers in Phonology 2.
HAAS, M. 1946. A grammatical sketch of Tunica. Linguistic Structures of Native America. New York: The Viking Fund. 337-366.
HAN, M. 1961. Japanese phonology: an analysis based upon sound spectrograms. Ph.D.diss. U. of Texas, Austin.
HARRISON, S.; and ALBERT, S. 1976. Mokilese reference grammar. Honolulu: University of Hawaii Press.
JANNEDY, S. 1995. Gestural phasing as an explanation for vowel devoicing in Turkish. OSU Working Papers in Linguistics 45.56-84.
JUN, S-A.; and BECKMAN, M. 1993. A gestural-overlap analysis of vowel devoicing in Japanese and Korean. Paper presented at the 67th Annual Meeting of the Linguistic Society of America. Los Angeles, CA.
JUN, S-A; and BECKMAN, M. 1994. Distribution of devoiced high vowels in Korean. Proceedings of the 1994 International Conference on Spoken Language Processing, vol. 1.479-482.
JUN, S-A; BECKMAN, M.; NIIMI, S.; TIEDE, M. 1997. Electromyographic evidence for a gestural-overlap analysis of vowel devoicing in Korean. Journal of Speech Sciences 1.153-200.
JUN, S-A; BECKMAN, M.; LEE, H-J. 1998. Fiberscopic evidence for the influence on vowel devoicing of the glottal configurations for Korean obstruents. UCLA Working Papers in Phonetics 96.
KAGAYA, R. 1974. A fiberscopic and acoustic study of the Korean stops, affricates and fricatives. Journal of Phonetics 2.161-180.
KIRK, P.; LADEFOGED, P.; and LADEFOGED. J. 1993. Quantifying acoustic properties of modal, breathy, and creaky vowels in Jalapa Mazatec. American Indian Linguistics and Ethnography in Honor of Laurence C. Thompson. UM Occasional Papers in Linguistics 10.435-50.
LEHISTE, I. Suprasegmentals. Cambridge, MA: MIT Press.
LÖFQVIST, A; and YOSHIOKA, H. 1980. Laryngeal activity in Icelandic obstruent production. Haskins Laboraties Status Report on Speech Research 63.275-92
LUKAS, J. 1953. Die Sprache der Tubu in der Zentralen Sahara. Berlin: Akademie-Verlag.
MADDIESON, I. 1984. Patterns of sounds. New York Cambridge University Press.
MASICA, C. 1991. The Indo-Aryan languages. New York: Cambridge U. Press.

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