January 2005

A Phonetically-Based Model of Phonological Vowel Reduction

Edward Flemming
flemming@mit.edu

Abstract

This paper proposes an analysis of phonological vowel reduction according to which vowel contrasts are subject to neutralization in unstressed syllables because it is more difficult to keep vowels distinct where vowel duration is shorter. Where the durational difference between stressed and unstressed vowels is large enough, it may not be possible to realize the same number of vowel contrasts in stressed and unstressed syllables while keeping the unstressed vowels adequately distinct, in which case it is preferable to reduce the size of the unstressed vowel inventory. The analysis is formulated as a numerical model, extending Liljencrants and Lindblom's (1972) model of vowel inventories by incorporating aspects of the prosodic and segmental contexts of vowels, and allowing for different vowel inventories in different contexts. The model contributes to our understanding of vowel reduction, providing an explanation for the observation that vowel reduction primarily neutralizes height contrasts, and demonstrates how Liljencrants and Lindblom's modeling approach can be developed to analyze contextual restrictions on the distribution of speech sounds.

1. Introduction

In a seminal paper, Liljencrants and Lindblom (1972) (henceforth L&L) attempted to account for generalizations about the typology of vowel systems in terms of a principle of maximal perceptual contrast between vowels. This work made at least two fundamental contributions to linguistic theory. First, it was one of the first examples of what Lindblom has since referred to as the deductive approach to the analysis of language (Lindblom 1986:16f., 1990a:139). That is, L&L sought to derive the form of vowel inventories from considerations of efficient communication. They reasoned that communication is impaired if contrasting vowels are confused, so an optimal language would maximize the distinctiveness of vowel contrasts. The second contribution was methodological: they tested their hypothesis through construction of a model, providing an explicit implementation of the notion of maximal perceptual contrast and the relationship between this principle and the form of vowel systems. This paper takes up L&L's project, both in the broad sense of adopting a deductive approach to the analysis of linguistic phenomena, and specifically, developing their model of vowel systems.

Subsequent research has developed L&L's model in various ways, modifying the perceptual distance metric (Lindblom 1986; Diehl, Lindblom and Creeger 2003; Schwartz et al 1997), adding considerations of articulatory effort (Bosch, Bonder and Pols 1987), and explicitly modeling the process of vowel inventory optimization (de Boer 2001). However all of these models analyze vowel inventories in isolation, and thus cannot address contextual variation in vowel systems. For example, vowels may be subject to substantial variation depending on consonantal context. Arabic languages contrast three vowel qualities, often presented in broad transcription as [i, æ, u], but the realization of the vowels is altered by a preceding pharyngealized consonant. In particular, the low vowel is realized as back [ɑ] in this context (Card 1983). The models of vowel inventories just cited make the simplifying assumption that a language has a single vowel inventory and that each vowel has a single phonetic realization, so they cannot offer any account of this type of contextual variation.

The topic of this paper is another pattern of contextual variation in vowel systems, phonological vowel reduction. This is a pattern in which contrasts are neutralized in unstressed syllables, so a larger number of vowel qualities are distinguished in stressed syllables than in unstressed syllables. For example, in standard Italian, seven vowels are distinguished in primary stressed syllables (1a), but only five vowels appear in other environments (1b). These contrasts are exemplified in (2). Vowel reduction is particularly interesting because it involves contextual variation in the number of contrasting vowels, as well as in the details of their realization.

(1) (a) Primary stressed: (b) Elsewhere:

(a) Primary stressed:			(b) Elsewhere:		
i		u	i		u
e		o	e		o
ɛ		ɔ		a	
	a				

(2)

[i]	v**í**no	'wine'	v**i**nífero	'wine-producing'
[e]	p**é**ska	'fishing'	p**e**skáre	'to fish'
[ɛ]	b**ɛ́**lːo	'beautiful'	-	
[a]	m**á**no	'hand'	m**a**nuále	'manual'
[ɔ]	m**ɔ́**lːe	'soft'	-	
[o]	n**ó**me	'name'	n**o**mináre	'to name, call'
[u]	k**ú**ra	'care'	k**u**ráre	'to treat'

We will propose an analysis of certain cross-linguistic generalizations about the nature of phonological vowel reduction, based on constraints that have been argued to shape vowel inventories in general. The analysis is tested through implementation in an explicit quantitative model, building on that proposed by L&L. In order to analyze vowel reduction, the model departs from most previous work on vowel inventories by modeling aspects of the prosodic and segmental context, and in allowing for different vowel inventories in different contexts. While these additions to models of vowel systems are directed at the analysis of vowel reduction in particular, many of the same issues arise in modeling other contextual restrictions on the distribution and realization of contrasts, and are thus necessary if the deductive modeling approach exemplified by L&L is to be extended to a broader range of phonological phenomena.

We will first lay out some basic facts about the typology of vowel reduction, then motivate a particular explanation for these patterns in terms of basic constraints on linguistic contrasts, and some observations about the nature of speech production. This explanation is then formalized in terms of a model of vowel systems.

2. Phonological vowel reduction

The first observation about vowel reduction is that it is relatively common across languages (see Crosswhite 1999 for a survey), so we would like to explain the generalization in (3).

(3) Unstressed syllables are environments that can condition neutralization of vowel contrasts.

In fact we will argue below that this generalization needs qualification. It is typical correlates of lack of stress that condition neutralization, not stress per se, the relevant correlates being short vowel duration and perhaps reduction in articulatory effort.

Examination of the typology of vowel reduction reveals more specific generalizations about the nature of the contrasts that are eliminated in unstressed syllables. In particular, vowel height contrasts are generally eliminated before backness or rounding contrasts (cf. Barnes 2002). Indeed, backness and rounding contrasts are generally only lost where all vowel quality contrasts are neutralized to a single vowel, as in English reduction to 'schwa'.

The Italian case presented above exemplifies the typical pattern for more moderate reduction: the contrasts between higher and lower mid vowels [e-ɛ, o-ɔ] are eliminated in unstressed syllables. Similar patterns of vowel reduction are observed in Brazilian Portuguese (Mattoso Camara 1972) and Slovene (Lencek 1982), for example. Another common pattern of vowel reduction involves reduction from a five vowel system (4a) in stressed syllables to a three vowel system (4b) in unstressed syllables, as in Standard Russian (Halle 1959), Southern Italian dialects (Mazzola 1976), and Catalan dialects (Recasens 1991, Herrick 2003), eliminating the contrast between high and mid vowels. Two transcriptions are provided for the lowest vowel in (3c) because both are found in impressionistic transcriptions of languages of this type, but where phonetic evidence is available, it indicates that the lowest vowel in reduced inventories of this kind is generally closer to mid [ə]. Similar qualifications are in order regarding the transcriptions of unstressed vowels in (1b) and (4b): We will see evidence that the lowest unstressed vowel in a reduced inventory is not as low as its stressed counterpart and might more accurately be transcribed as [ɐ] or [ɜ] (section 6).

(4) (b) i   u   (c) i   u
e   o   a/ə
a

These are the most common patterns of vowel reduction in Crosswhite's (1999) comprehensive survey. Other attested patterns include reduction from a seven vowel system, as in standard Italian, to a three vowel system (3b), e.g. Eastern Catalan (Herrick 2003). We do not find cases in which only the contrasts between front and back vowels are neutralized, for example. This is true even in languages with more extensive backness or rounding contrasts: there are no cases in which front rounded vowels or non-low central vowels are allowed in stressed syllables, but excluded from unstressed syllables, unless all vowel contrasts are neutralized (as in English reduction to 'schwa')[1]. Where a language has front

[1] Crosswhite (1999) describes the Upper Carniolan dialect of Slovene as displaying such a pattern. According to her description, the short high vowels [i] and [u] neutralized with central [ə]. This change did occur in the

rounded or central vowels and vowel reduction, the reduction eliminates height contrasts but preserves central and front rounded vowels. For example, Scots Gaelic, as described by Borgstrøm (1940), has the stressed vowel inventory shown in (5), with high and mid central vowels. Some height contrasts are neutralized in unstressed syllables, but the contrast between front, central and back vowels is retained. The transcription implies that it is the mid central vowel that is retained, but no instrumental data on the quality of the vowels is available. Acehnese appears to have a similar pattern of reduction, except it eliminates all mid vowels (Durie 1985:21).

(5)	Stressed:	i	ɨ	u	Unstressed:	i		u
		e	ə	o		ɛ	ə	ɔ
		ɛ		ɔ			a	
			a					

Mantuan Italian (Miglio 1996) contrasts front rounded and unrounded vowels (6). The lower mid front vowel [ɛ] and the mid rounded vowels [ø, o, ɔ] are excluded from unstressed syllables, but the contrast between front unrounded [i] and front rounded [y] is preserved. Note also that Mantuan Italian preserves more height contrasts among front vowels than among back vowels. The same asymmetry between front and back vowels is observed in Bulgarian dialects where a stressed vowel system [i, e, a, o, u, ə] is reduced to [i, e, ə, u] in unstressed syllables (Wood and Pettersson 1988).

(6)	Stressed:	i	y	u	Unstressed:	i	y	u
		e	ø	o		e		
		ɛ		ɔ			a	
			a					

There are languages that restrict some backness or rounding contrasts to the first syllable of words. This can give the appearance of vowel reduction where stress is also initial, as in Estonian, which has nine vowels [i, e, æ, y, ø, a, o, u, ɤ] in initial stressed syllables, but does not allow mid back unrounded [ɤ] in non-initial syllables. In other words, a rounding contrast is neutralized without neutralization of any height contrasts, apparently contradicting the generalization just formulated. However there is good reason to believe that these cases actually involve a separate phenomenon of neutralization in non-initial syllables, which is independent of stress. This interpretation is supported by evidence that initial syllables can allow a greater variety of vowel contrasts than non-initial syllables even in the absence of stress, as in Turkish and Shona, for example (Beckman 1998, Steriade 1993). Beckman

historical development of Upper Carniolan Slovene according to Greenberg (2000) and Lencek (1982), and is referred to as vowel reduction, but it is not a vowel reduction process in the sense of that term used here – i.e. a reduction in the number of vowel contrasts in unstressed syllables. First, the sound change was not conditioned by stress – it applied to all short high vowels, stressed and unstressed (Greenberg 2000:174, Lencek 1982:174). Second, the change did not result in the loss of unstressed [i, u] - these vowels were ‘reintroduced’ by other developments (e.g. [Vj]>[i] and pretonic [o]>[u]) so the unstressed vowel system of Upper Carniolan is [i, e, a, o, u, ə] (Greenberg 2000:174).

(1998), MacEachern (1997) and Smith (2002) argue that the special status of initial position is related to the importance of initial syllables in lexical access. We will return to this phenomenon briefly in section 10, discussing how this analysis can be formalized in the framework developed here.

So the second generalization that we wish to explain is formulated in (7):

(7) Vowel reduction primarily eliminates height contrasts, and only eliminates backness or rounding contrasts under restricted conditions.

It will be argued that vowel contrasts are neutralized in short, unstressed syllables, because it is more difficult to keep vowels distinct in this environment. As a result it can be preferable to reduce the number of contrasts rather than tolerate less distinct contrasts. Height contrasts are neutralized first because it is low vowels in particular that become more difficult to produce in short unstressed syllables, so it is particularly difficult to keep height contrasts distinct. This analysis of vowel reduction is developed in more detail in the next section, and formalized in terms of a quantitative model in section four.

3. Outline of an analysis of vowel reduction

We will argue that vowel reduction results from basic constraints on sound systems of a kind adopted by much previous work on deductive models of phonological inventories, once we take into account the effects of stress-based variation in vowel duration and the role of duration in determining the articulatory effort involved in producing a vowel.

Two core constraints posited by most deductive models of sound systems are the preferences to maximize the distinctiveness of contrasts, and to minimize articulatory effort (e.g. L&L, Lindblom 1998, 1990b, Bosch et al 1987). The analysis of vowel reduction proposed here posits a third basic constraint, a preference to maximize the number of contrasting vowels (Flemming 2001, 2004). This constraint also derives from considerations of communicative efficiency because increasing the number of contrasting vowels increases the information that can be conveyed by each vowel. Given three contrasting vowels, uttering a single vowel can potentially differentiate between three words, whereas if there are five contrasting vowels, a single vowel could differentiate between five words.

Maximizing the number of contrasting vowels conflicts with the need to maximize the distinctiveness of contrasts because the space of articulatorily possible vowels is finite, so distinguishing more vowel contrasts implies packing the vowels closer together in that space. So the more vowel contrasts a language differentiates, the less distinct those contrasts can be. The size of a vowel inventory can then be analyzed as the result of balancing these conflicting constraints.

It is probable that other constraints shape segment inventories, such as featural economy (Clements 2003, Ohala 1980) or a preference to reuse articulatory gestures (Lindblom 1998), but only the three constraints outlined above will be incorporated into the formal model since these are the ones that are required to derive the basic properties of vowel reduction.

In essence, vowel reduction arises because unstressed vowels are usually shorter than stressed vowels and it is more difficult to achieve distinct vowel qualities where vowel duration is shorter. Where the durational difference between stressed and unstressed vowels is large enough, it may not be possible to realize the same number of vowel contrasts in stressed and unstressed syllables while keeping the unstressed vowels adequately distinct, in which case it is preferable to reduce the size of the unstressed vowel inventory.

The conclusion that shorter vowels are generally more effortful to produce than longer vowels follows from the assumption that faster articulatory movements are more effortful, other things being equal (e.g. Nelson 1983, Perkell et al 2002). The production of a vowel involves moving from the preceding segment to the position for the vowel, and then moving the articulators on to the position for the following segment. The shorter the vowel, the less time the articulators have to make these movements, and the faster they must move.

The most direct evidence for the significance of this increase in effort comes from Lindblom's (1963) study of vowel production in Swedish CVC sequences of various durations. He found that vowel formants varied systematically depending on the consonant context and the duration of the vowel. At long durations, vowel formants came close to consistent values that could be regarded as the vowel targets. But as vowel duration decreased, the vowel formants were displaced towards values characteristic of the neighboring consonants, and the magnitude of this displacement increased as vowel duration decreased. So as vowels shortened they progressively assimilated to the consonant context. This pattern of target undershoot can be understood as a consequence of minimizing effort by avoiding fast articulator movements at the cost of falling short of the vowel targets. This interpretation is further supported by the fact that the amount of undershoot increased as the distance between the vowel and the consonants increased – the greater the transition between vowel and consonant, the faster the articulator must move for any given vowel duration. For example, much more F1 undershoot is observed in low vowels than in high vowels (Fig. 1) because producing the low tongue and jaw position for a low vowel requires greater articulator movement from the constricted positions of adjacent consonants than the production of a high vowel. So the effects on undershoot of both duration and distance to the vowel target can both be understood in terms of a dispreference for rapid articulator movements[2].

The effects of vowel undershoot are illustrated in Fig. 1. This figure shows the formant frequencies of five vowels in the context [g_g] for three vowel durations: 200ms, 125 ms, and 100 ms. The values shown are not the measured formant frequencies, which are not reported in the paper, they are values derived from the model that Lindblom fit to his data. It is apparent that the separation between vowels in formant space is reduced as vowel duration decreases. Vowel reduction mitigates this decrease in distinctiveness by reducing the number of vowel contrasts, maintaining adequate separation between the remaining vowels.

Note also that the vowel space is primarily compressed in the F1 dimension (vowel height). It will be argued below that this is a systematic property of undershoot in most consonantal contexts, and is the basis for the typological generalization that vowel reduction primarily targets height contrasts (section 2).

The analysis just outlined does not relate vowel reduction to stress per se, but rather to typical correlates of stress, in particular decreased duration in unstressed syllables, and possibly also decreased effort. As discussed further in section 6, this implies that we should not find vowel reduction in languages where stress does not affect vowel duration, and that vowel reduction could arise in other circumstances that give rise to significantly shortened

[2] Barnes (2002: chapter 2) also argues that duration-dependent undershoot plays a key role in explaining patterns of vowel reduction.

vowel duration. For example, segmental context affects vowel duration, e.g. vowels are typically shorter before voiceless obstruents (Chen 1970) and before geminates (Maddieson 1985). If these effects can result in sufficiently low vowel duration, then the analysis proposed here predicts that they should be able to result in vowel reduction. This paper focuses on the case of unstressed vowel reduction because the typology of this phenomenon is well studied, whereas other duration-based effects are not. However, we will see evidence that extra short vowel duration results in reduction independently of stress in Northwest Caucasian languages (section 9).

In the next sections, we formalize the analysis of vowel reduction in the context of a model of vowel systems, and show how it derives the generalizations about the typology of vowel reduction discussed in section 2. The aim of the model is to derive these generalizations rather than the details of a complete typology of vowel reduction. Deriving the details of individual patterns of vowel reduction would depend on a better model of vowel inventories than is currently available, but this does not prevent us from showing that the proposed constraints give rise to qualitatively correct patterns of reduction. In particular we seek to derive the basic phenomenon of reduction – selection of a smaller inventory of vowels in short, unstressed syllables – and the observed restrictions on neutralization of backness and rounding contrasts.

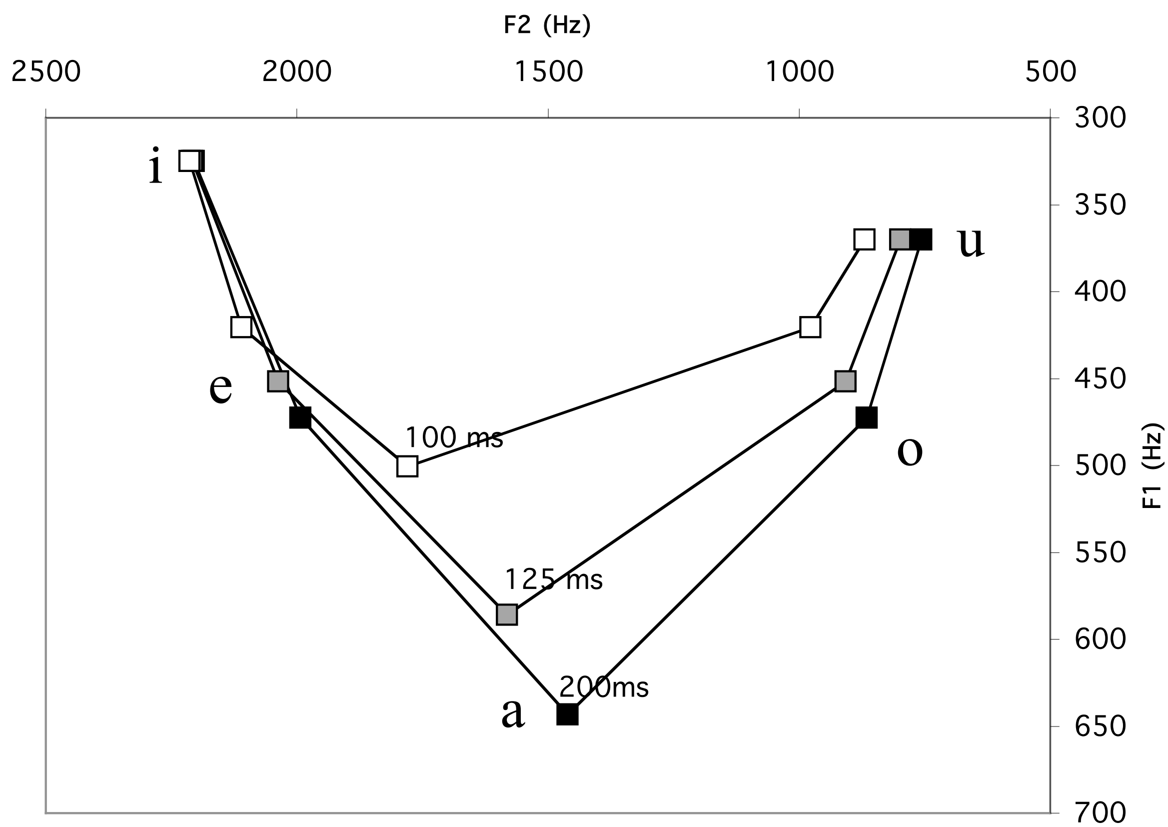


Fig.1. Lindblom's model of the effects of vowel duration on the formant frequencies of five Swedish vowels, produced in the context [g_g].

4. A model of vowel reduction.

Formalizing this analysis of vowel reduction requires explicit formulation of the three basic constraints: maximize the distinctiveness of contrasts, maximize the number of contrasts, and minimize effort. The basic structure of the model builds on L&L. As in that approach, inventories of vowels are selected from a space of physiologically possible vowels so as to minimize a cost function that implements the constraints on contrasts. The possible vowels are defined by a space of possible values of first and second formant frequencies, modeled on the space described in L&L (Fig. 2).

The distinctiveness of a contrast between two vowels is quantified, following L&L, as the euclidean distance between the two vowels in the formant space. That is, the distance d_{ij} between two vowels, V_i and V_j, is given by the formula in (8), where x_n is F2 of V_n in Bark and y_n is F1 of V_n in Bark. This formulation deviates from L&L in incorporating a weighting factor a, with a value less than 1, that reduces the distinctiveness of F2 differences relative to F1 differences. Without such a weighting factor, the model of L&L predicts that all vowel inventories with more than five vowels should contain high central vowels as a result of the substantial width of the top of the vowel space in Fig. 2. Schwartz et al (1997) adopt a similar approach to this problem, while recent work by Diehl, Lindblom, and Creeger (2003) suggests that the lower weighting of F2 differences might be derived from the lower intensity of F2 compared to F1. The proper analysis of this phenomenon is peripheral to our primary concern here, which is modeling the effects of stress on vowel inventories. Any model of the distinctiveness of contrasts could be incorporated into the present model. For now, we adopt the weighting factor of Schwartz et al because it is much simpler to implement than Diehl et al's whole-spectrum model of perceptual distinctiveness. Varying the value of a also allows us to derive a wider range of vowel systems (Vallée, Schwartz and Escudier 1999).

In the current implementation the role of F3 is neglected. L&L calculated distinctiveness based on F1 and F2', where F2' is the 'effective second formant', calculated from the frequencies of F2 and F3. It was not possible to incorporate F3 into the present model because no data is available on the effects of undershoot on F3. Keeping the vowel space two dimensional also helps to make the search for optimal vowel inventories more tractable. The main effect of substituting F2 for F2' is a slight narrowing of the vowel space in the F2 dimension, while preserving the same overall shape of the vowel space. Since the details of the precise shape of the space of possible vowels is still in question (see Atal et al 1978 and Ladefoged 2001:159 for two further analyses), this is a relatively unimportant difference between the models.

(8) $$d_{ij} = \sqrt{(a(x_i - x_j))^2 + (y_i - y_j)^2}$$

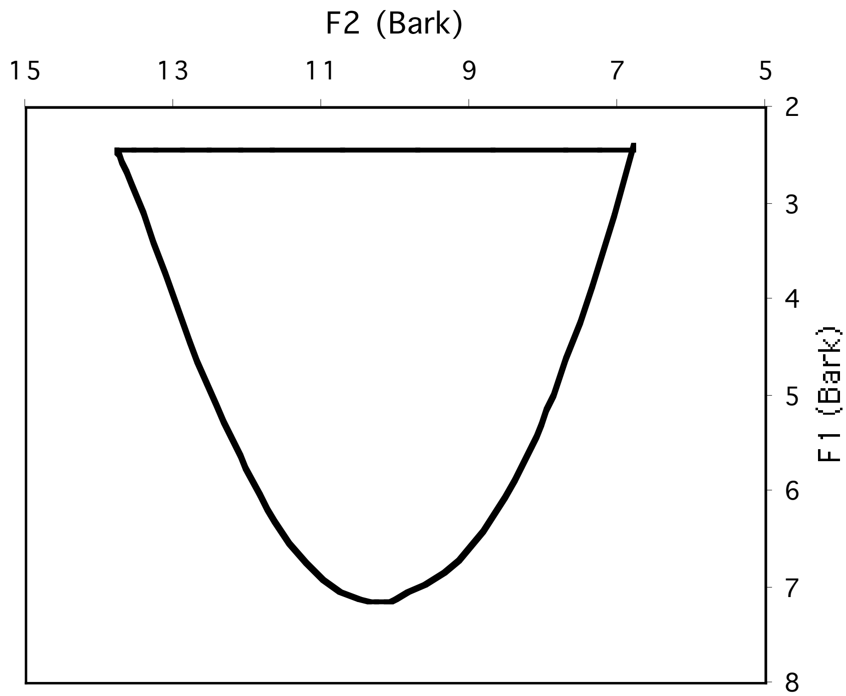


Fig.2. The boundaries of the space of possible vowels in the F1×F2 plane. Modeled after Liljencrants and Lindblom 1972.

We diverge from the model of L&L in formulating the constraint favoring maximization of distinctiveness as a requirement that the minimum distance between any pair of vowels should be maximized. That is, a system of vowels is only as good as its worst contrast. Maximizing the minimum distance between signals is a standard criterion for selecting optimal signal sets for communication systems (e.g. Anderson 1999). In that context it is motivated by the observation that the probability of confusing two signals in white noise falls rapidly as the difference between them is increase, so confusions between the least distinct signal pairs are a much more significant source of error than confusions between more distinct pairs. Consequently the rate of confusion can be reduced most effectively by increasing the minimum distance between signals. L&L employed a more global measure of the distinctiveness of a vowel system, based on the distances between all pairs of vowels in the inventory. We will see that adopting ‘maximize the minimum distance’ as a distinctiveness constraint has certain advantages in accounting for the properties of vowel reduction (section 8).

Maximizing the minimum distance between vowels favors distributing vowels evenly over the entire vowel space. This is because a given configuration of vowels can only be improved by increasing the separation of the closest pair of vowels. So it preferable to increase the separation of the closest pair even if it means reducing the distance between another pair of vowels, but only up to the point where the distances between the two pairs is equal. Further separation of the first pair beyond that point will mean that it is no longer the closest pair. So the optimal configuration of vowels will tend to be one in which the distances between any pair of vowels is the same, so no pair can be separated further without reducing the minimum distance of the system.

The cost function that is minimized in the selection of vowel inventories is the reciprocal of the square of the smallest distance between any pair of vowels in the inventory (9). Minimizing this function implies maximizing d_{min}. For undershoot in unstressed syllables to give rise to neutralization of contrasts, stressed and unstressed vowels must be subject to the same distinctiveness requirements, so d_{min} is the smallest distance found in either inventory. That is, both stressed and unstressed vowel inventories are evaluated together as a vowel system.

(9) $$\frac{1}{d_{\min}^2} \qquad \text{where } d_{\min} = \min_{i \neq j} d_{ij}$$

The preference to maximize the number of contrasts is implemented by adding a term to the cost function that increases as the number of vowels decreases (10). It is desirable to maximize the size of both the stressed and the unstressed vowel inventories, so this term is based on the average size of the two inventories. n_s is the number of vowels permitted in stressed syllables, and n_u is the number of vowels permitted in unstressed syllables. One could also imagine using a weighted average, perhaps reflecting the relative frequencies of stressed and unstressed syllables.

(10) $$\frac{1}{n_{ave}^2} \qquad \text{where } n_{ave} = \frac{n_{stressed} + n_{unstressed}}{2}$$

The specific form of these cost functions is somewhat arbitrary but designed to derive compromises between distinctiveness and number of vowels. To achieve this result, the cost function for distinctiveness should increase exponentially as minimum distance decreases, and the cost function for number of vowels should increase exponentially as the number of vowels decreases so neither constraint can completely dominate the other even if there is a substantial difference in their relative weights. That is, exponential cost functions ensure that very low minimum distances and very small vowel inventories always incur high costs, so a compromise that lies between these two extremes will generally minimize cost. Accordingly, we take $1/d_{min}^2$ as the measure of dispersion, and $1/n_{ave}^2$ as the cost contributed by the size of the inventory. The relative importance of these two factors is determined by a positive weighting factor w_n. So selection of a system of stressed and unstressed vowels is formulated as an optimization problem: The vowel system V of stressed and unstressed vowels is selected so as to minimize the cost function composed of the terms reflecting the distinctiveness and inventory size constraints (11).

(11) $$\underset{V}{\text{minimize}} \frac{1}{d_{min}^2} + \frac{w_n^2}{n_{ave}^2}$$

Effort constraints serve to restrict the space of accessible vowel qualities, depending on vowel duration. This limitation is modeled by contracting the space of possible vowels according to the undershoot functions proposed by Lindblom (1963). Lindblom found that undershoot increased exponentially as duration decreased. That is formant frequencies at the mid-point of a vowel are characterized by a function of the form in (12), where T represents vowel duration, F_{2t} is the second formant target value for the vowel, and F_{2i} is the frequency of the second formant at the release of the consonant preceding the vowel. So (12) states that a vowel undershoots its target by a proportion of the difference between the F2 target and F2 adjacent to the consonant release, and the size of that proportion increases exponentially as vowel duration T decreases. The constants k_2 and β_2 depend on the consonant context. For example, for the context [g_g], $k_2 = 1.5$ and $\beta_2 = 0.01$. This pattern of undershoot is illustrated in Fig. 3[3].

(12) $$F_{2V} = k_2(F_{2l} - F_{2t})e^{-\beta_2 T} + F_{2t}$$

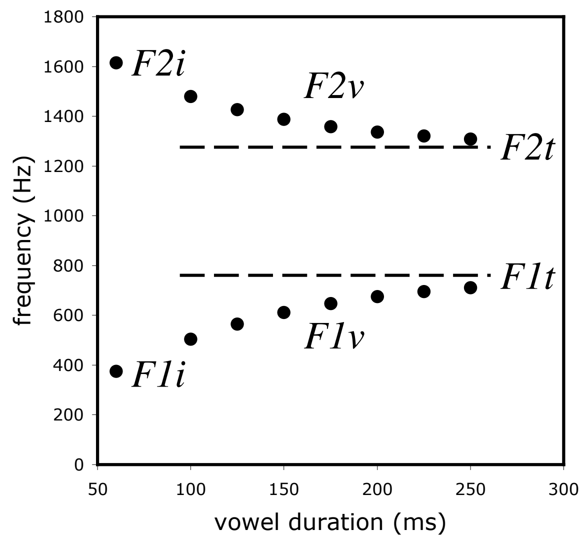


Fig.3. Lindblom's (1963) model of formant undershoot. Vowel formant frequencies, $F1_v$ and $F2_v$, are plotted as a function of vowel duration. The target values of the vowel formants, $F1_t$ and $F2_t$, are indicated by dashed lines, and the formant frequencies at the release of the consonant, $F1_i$ and $F2_i$, are plotted at the left of the figure.

It is important to note that the second formant frequency at the release of the consonant, F_{2i}, varies as a function of the vowel. Lindblom (1963) used the measured values for each vowel context in his models, but here we incorporate this variation into the model. We assign a fixed F2 target, or locus, F_{2l}, for a consonant context and then assume that F_{2i} assimilates to F_{2t} by a proportion $c<1$ of the distance between them, so $k_2(F_{2i}-F_{2t})$ in Lindblom's equation (12) can be replaced by $ck_2(F_{2l}-F_{2t})$ (13). This model of assimilation of consonants to adjacent vowels is similar to Klatt's (1987) characterization of the way in which F2 at consonant release varies as a function of F2 in the middle of the following vowel.

---

[3] Note that k_2 can be greater than 1, which implies that the vowel F2 does not converge on F2 at the consonant release, $F2_i$, as vowel duration approaches 0. Instead vowel F2 will exceed $F2_i$, which is inconsistent with the idea that undershoot represents assimilation of the vowel to the consonantal context. I take this to indicate that Lindblom's model is not valid at very low durations (which were not elicited in the experiment), and the function should level out as $F2_v$ approaches $F2_i$. That is, the true relationship between undershoot and vowel duration probably follows a sigmoid function. We will confine ourselves to the duration range where Lindblom's model is well motivated.

(13) $F_{2V} = ck_2(F_{2l} - F_{2t})e^{-\beta_2 T} + F_{2t}$

Undershoot for F1 follows a similar exponential duration-dependent function (14), except that F1 at the release of the consonant is fixed at 375 Hz for all consonants. There is no undershoot of F1 for vowels with target F1, $F1_t$, of less than 375 Hz. Again, k_1 and β_1 depend on the consonant context. Note that while $F2_i$ assimilates to the F2 of the following vowel, $F1_i$ does not show any assimilation. This proves to be significant in accounting for the observation that vowel reduction primarily neutralizes height contrasts.

(14) $F_{1V} = F_{1t}$ for $F_{1t} \leq 375$ Hz
$F_{1V} = k_1(375 - F_{1t})e^{-\beta_1 T} + F_{1t}$ for $F_{1t} > 375$ Hz

Effort limitations are enforced by reducing the boundaries of the space of possible vowels towards the locus for the consonant context. The size of the boundary shift in the F1 and F2 dimensions is given by equations of the form shown in (13) and (14), and is thus dependent on vowel duration (Fig. 4). Alternatively, one could regard the space of possible vowel targets as being unaltered by duration, but actual vowel realizations undershoot those targets, following the equations in (13) and (14) – the final result is the same.

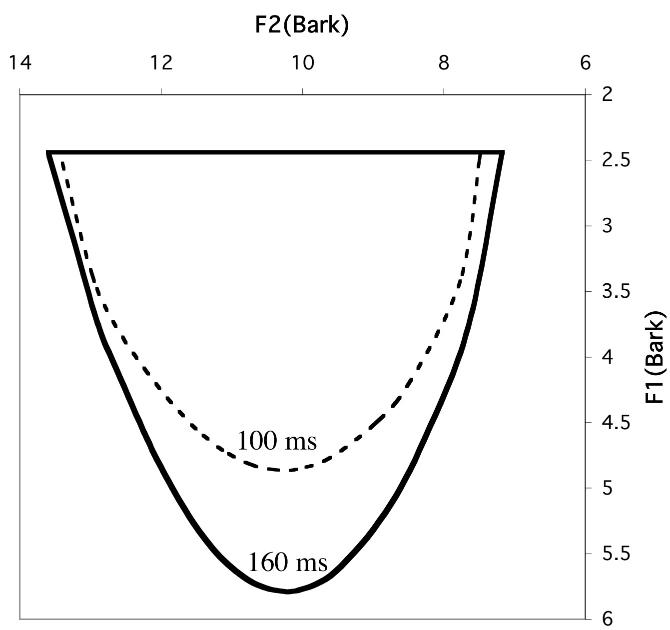


Fig. 4. Contraction of space of possible vowels as a function of vowel duration. The solid line marks the boundary of the space for vowels of duration 160 ms, the dashed line marks the boundary of the space for vowels of duration 100 ms.

5. Identifying optimal vowel systems

It is difficult to minimize the cost function (9) directly, since n_s and n_u are constrained to adopt integer values, so in practice the distinctiveness cost was minimized for a range of inventory sizes (e.g. n = 2-10) for both stressed and unstressed vowel durations and these costs are then used to identify the optimal vowel system, according to the procedure described in the following. For each inventory size, vowel formant frequencies that maximize d_{min} were found using the Matlab optimization routine, 'fminimax'[4]. Sometimes the algorithm gets stuck in a local minimum, but it seems to be possible to find the global minimum by running the optimization routine several times with random starting positions for the vowels (twenty runs were used in most of the simulations reported here). Fig. 5 illustrates the results of this procedure. The parameters were set as follows: the F2 weight $a = 0.14$, $k_1 = 1.5$, $\beta_1 = 0.008$, $k_2 = 1.5$, $c = 0.27$, $\beta_2 = 0.01$, $F2_l = 1400$ Hz, $w_n = 6$. Stressed vowels were 160 ms in duration, while unstressed vowels were 100 ms.

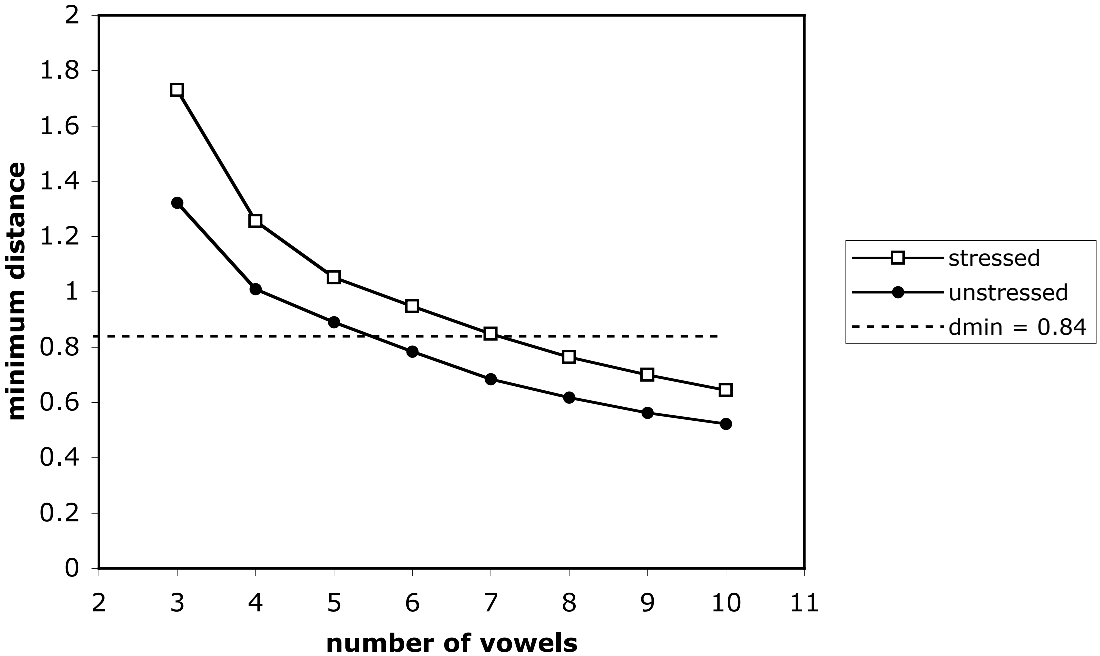


Fig. 5. Minimum distance in optimal inventories of different sizes.

In Fig. 5 we observe the expected trade-off between minimum distance and the number of vowels in the inventory – that is, as more contrasting vowels are packed into the vowel space, the distance between them decreases. The optimal vowel system can then be identified by observing how the number of stressed and unstressed vowels increases (and thus the number cost decreases) as the minimum distance decreases (and so the distinctiveness cost increases), and finding the optimal trade-off according to the relative weighting of these two costs. Fig. 6 illustrates this procedure. It shows plots of distinctiveness cost ($1/d_{min}^2$, plotted with triangles), number cost (w_n^2/n_{ave}^2, plotted with circles), and the total cost (the sum of these two costs, filled squares) as a function of minimum distance for the critical values where inventory size increases. With the parameter values specified above, the lowest total cost is achieved with a minimum distance of 0.84. This minimum distance is plotted with a dashed line in Fig. 5, where it can be seen that it allows seven stressed vowels and five unstressed vowels.

---

[4] *Fminimax* is an algorithm for solving 'minimax' optimization problems, i.e. problems in which the maximum value of a set of functions must be minimized. We have presented vowel system selection as involving the inverse problem, i.e. maximizing the minimum distance between any pair of vowels. This problem was converted into a minimax problem by negating the distances between vowels, so a larger distance is represented by a lower (negative) number.

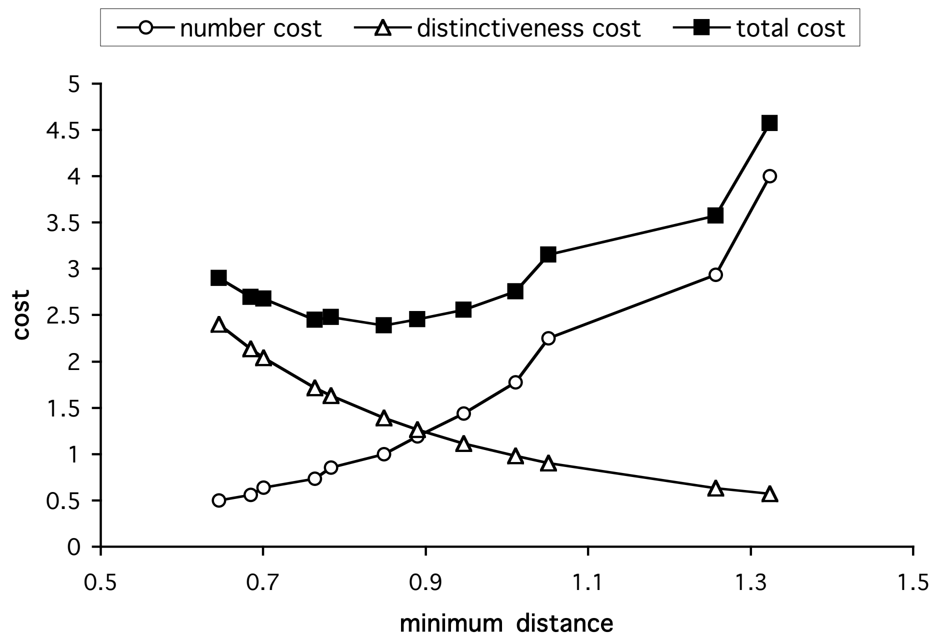


Fig. 6. Cost of stressed and unstressed vowel inventory pairs plotted against the minimum distance found in either inventory (see text for details).

The optimal stressed and unstressed vowel inventories, given these parameter settings, are shown in Fig. 7. This pattern is similar to Standard Italian (presented in (1) above) – seven vowels in stressed syllables, and five in unstressed syllables (see Fig. 10, below, for formant data for the Italian vowels). So the model outlined can derive the basic phenomenon of vowel reduction. That is, patterns where a reduction in vowel duration results in a reduction in the optimal number of vowel contrasts. This pattern is derived from relatively direct formulations of three basic constraints: maximization of distinctiveness, minimization of effort, and maximization of the number of contrasts.

The effort constraint is given the least general formulation, since it is derived from empirically observed consequences of effort minimization rather than an attempt to quantify the effort involved in producing vowels. However, the current formulation has the advantage of demonstrating that neutralizing phonological vowel reduction can be derived from patterns of undershoot observed in non-neutralizing 'phonetic' vowel reduction, so the same effort constraints can motivate both phonetic and phonological reduction. Indeed, reducing the weight of the constraint favoring large vowel inventories, w_n, to 1 while keeping other parameter values as above results in a pattern in which there are three vowels in both stressed and unstressed syllables – i.e. no phonological reduction – but the unstressed vowels are phonetically reduced (Fig. 8).

Note that the back vowel in the three vowel inventory is significantly lower than the front vowel, i.e. the inventory could be transcribed [i, a, o] rather than the more familiar [i, a, u]. This pattern is attested in languages such as Piraha (Everett and Everett 1984) and Axininca Campa (Payne 1981). This arrangement serves to increase the distance between the front and back vowels by differentiating them in F1 as well as F2. An inventory closer to [i, a, u] can be derived with a higher value of the weight that reduces the distinctiveness of F2 differences (a in (8)), although the present model predicts that the back vowel should always be somewhat lower than the front vowel. However, this prediction is peripheral to the analysis of vowel reduction which is the focus here, so we will not pursue it further.

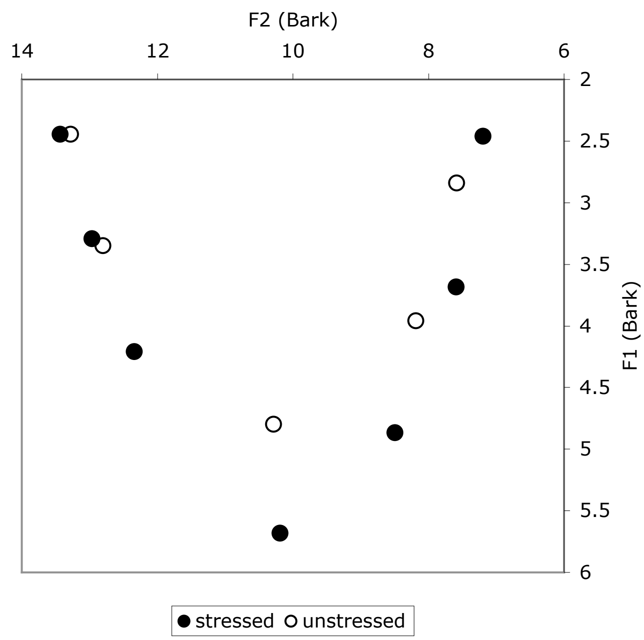


Fig. 7. Optimal stressed and unstressed vowel inventories, illustrating neutralizing reduction.

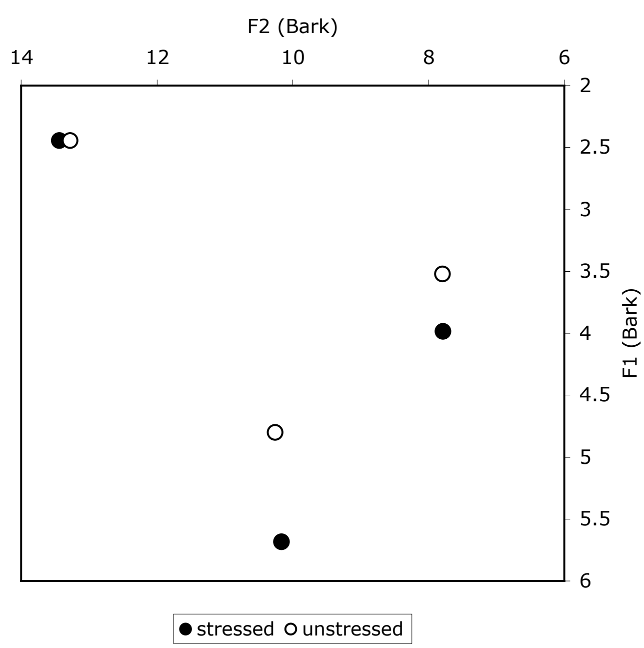


Fig. 8. Optimal stressed and unstressed vowel inventories illustrating phonetic reduction without neutralization.

By varying just the weight w_n from 1 to 6, the model derives the vowel systems shown in Fig. 9 in addition to those shown in Figs. 7 and 8. As explained at the outset, the goal of the model is to derive typological generalizations about vowel reduction. From this point of view, the most important observation about these vowel systems is that they represent two attested qualitative patterns of neutralization: the systems in Fig. 7 and Fig. 9(b, c) involve reduction of height contrasts, while Fig. 9(a, d) involve neutralization of the contrast between front and back low vowels. As discussed in section 2, neutralization of height contrasts is the most common pattern of vowel reduction, while neutralization of a contrast between front and back low vowels is attested in Chamorro (Topping and Dungca 1973, Chung 1983). Chamorro has six vowels [i, e, æ, ɑ, o, u] in stressed syllables, reduced to five in unstressed syllables [i, e, a, o, u], which corresponds to the pattern in (d). However, the mid and high vowels are in complementary distribution in the native vocabulary so, prior to the influx of Spanish loans, the vowel inventory in stressed syllables contained only four contrasting vowels reduced to three in unstressed syllables, which is closer to (a). Neutralization of backness contrasts involving low vowels is one of the limited cases in which F2 backness contrasts are eliminated in vowel reduction, as discussed further in section 7, below.

Although the focus of this paper is on modeling generalizations about vowel reduction rather than vowel inventories per se, it is necessary for the model to derive reasonable stressed vowel inventories in order to be able to assess its plausibility as a model of vowel reduction. All of the stressed vowel inventories in Figs. 7-9 correspond to attested inventories. As already noted, the stressed inventory in Fig. 7 is similar to Standard Italian, while that in Fig. 8 is similar to Piraha and Axininca Campa. The stressed inventory in Fig. 9(a) approximates [i, e, a, o], which is found in Navajo (McDonough 2003) and Klamath (Barker 1964), for example. The stressed vowels in (b) form the basic five vowel system [i, e, a, o, u], with the back vowels lower than their front counterparts, a pattern that is certainly attested if not in quite this extreme form (the Swedish vowels in Fig. 1 and the Italian vowels in Fig. 10 display tendencies in this direction). In (c) and (d) we find six stressed vowels [i, e, æ, a, o, u], which corresponds to the inventory of Farsi (Majidi 1991), and the surface inventory of Chamorro.

As for the details of the systems of reduction derived by the model, the pattern in Fig. 7 is comparable to standard Italian, while those in Fig. 9(a, d) are comparable to Chamorro. The patterns in (b) and (c) do not correspond directly to attested systems as far as I know, although they do instantiate the typical pattern in which only height contrasts are eliminated. In addition, the asymmetry between front and back vowels observed in (b), with elimination of mid back vowels but not mid front vowels, is attested in dialects of Bulgarian and in Mantuan Italian (6), as noted in section 2. This front-back asymmetry derives ultimately from the asymmetrical shape of the vowel space: there is more space along the front edge of the inventory than along the back (Fig. 2), so there is a tendency to allow more contrasts along the front edge of the space.

In the following sections we will examine the explanations offered by the model for the observations that vowel reduction applies in unstressed syllables, and that vowel reduction primarily neutralizes height contrasts.

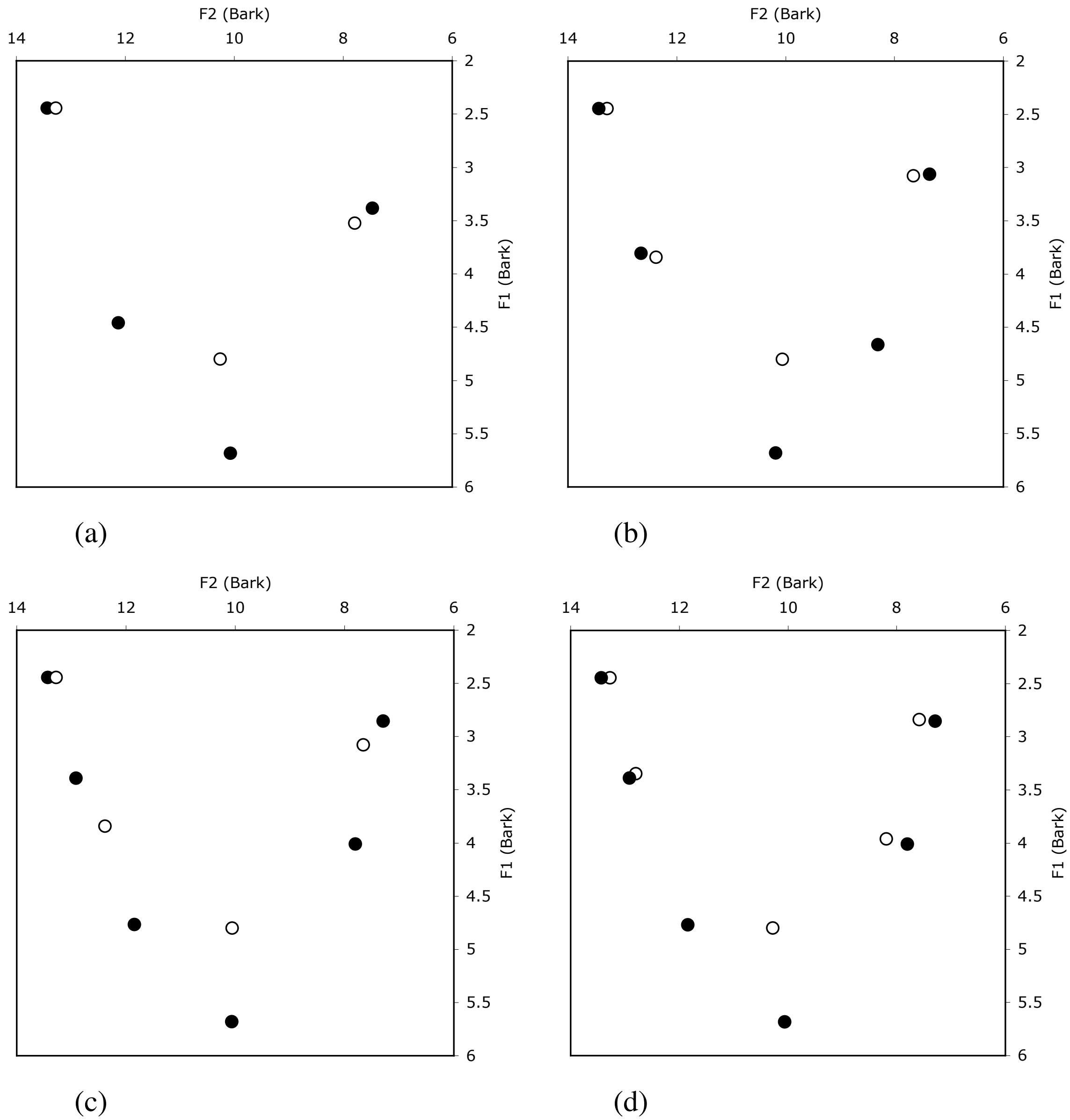


Fig. 9. Vowel systems derived by varying the value of w_n (see text for details).

6. The relations between reduction, undershoot, vowel duration, and stress

As formulated so far, the model implies that reduction occurs in unstressed syllables because of the shorter duration of unstressed vowels. So the analysis predicts that vowel reduction should not be found in languages like K'ekchi (Berinstein 1979), Toba Batak (Podesva and Adisasmito-Smith 1999), and Czech (Palkova 1994), where stress is not marked by duration[5]. That is, we have analyzed phonological vowel reduction as being

[5] Crosswhite (2001) also suggests that at least certain types of vowel reduction are restricted to languages with a 'strong duratio-based stress'.

driven by undershoot in unstressed syllables and, following Lindblom (1963), we have identified reduced duration as the primary cause of increased undershoot. But the extent of undershoot also depends on the rate of articulator movement, so undershoot can be offset by increasing articulatory effort (Moon and Lindblom 1994), or, conversely, it can be increased by a reduction in articulatory effort. So if the level of effort differs between stressed and unstressed syllables, that would also contribute to the occurrence of vowel reduction. The facts of the matter are unclear. Wouters and Macon (2002) provide some evidence of increased spectral rate of change in lexically stressed syllables which they attribute to faster articulator movements, but the effect was only observed for a limited set of onset-vowel transitions consisting almost entirely of [ɹV] sequences, so this might simply reflect an effect of stress on the realization of [ɹ] (cf. Hagiwara 1995:79-83)[6].

If effort can differ between stressed and unstressed syllables, then the relationship between duration and undershoot is less direct. The increase in undershoot associated with a decrease in duration would depend on whether it was accompanied by a decrease in effort. In principle, a very large difference in effort could give rise to vowel reduction without a difference in duration between stressed and unstressed vowels, although differences of this magnitude do not seem likely.

Whether undershoot is primarily a function of vowel duration or stress-conditioned variation in effort also plays a role, the model embodies the hypothesis that it is undershoot that motivates phonological vowel reduction. So the prediction remains that certain correlates of lack of stress should condition vowel reduction, not the lack of stress itself. Specifically it is those correlates that contribute to vowel undershoot that condition phonological reduction, i.e. short duration, and perhaps reduced effort. If stress is marked only by pitch movements, for example, no vowel reduction is expected. It is also predicted that phonological reduction should be accompanied by phonetic reduction in the sense of contraction of the vowel space, at least at normal speech rates. In particular, we should find that the lowest vowel of the unstressed inventory is raised compared to the lowest vowel in the stressed inventory. This is what we observe where acoustic data are available on the realization of reduced vowels. For example, average formant values for the stressed and unstressed vowels of Italian from a study by Albano Leoni et al (1995) are plotted in Fig. 10. It can be seen that the unstressed vowels occupy a compressed space compared to the stressed vowels. In particular, the lowest unstressed vowel is significantly higher than its stressed counterpart. Similar raising of unstressed low vowels is observed in other languages with phonological vowel reduction, including Catalan (Herrick 2003), Russian (Padgett and Tabain 2003), Bulgarian (Lehiste and Popov 1970), and Brazilian Portuguese (Fails and Clegg 1992). It is interesting to observe that unstressed [i] in Fig. 10 has a slightly higher F1 than its stressed counterpart. This is not predicted by the model proposed here, but could be a consequence of vowel-to-vowel coarticulation with non-high vowels, which would be a natural extension of the present model.

---

[6] The pattern of vowel reduction described for Shimakonde by Liphola (2001) appears to be a counterexample to the predicted relationship between reduction and duration. Shimakonde has five vowels [i, e, a, o, u], but the mid vowels optionally neutralize with [a] in pretonic syllables. Reduction never applies to stress vowels, even if they are short, and reduction can apply to unstressed long vowels derived from coalescence of vowel sequences. The pattern of reduction is anomalous in a number of other respects, including the fact that it can only apply to a vowel that occurs later in a stem if it also applies to all mid vowels that occur earlier in the word (pp. 168ff.). So it is possible that Shimakonde reduction differs in motivation from the other cases of stress conditioned vowel reduction that are discussed here.

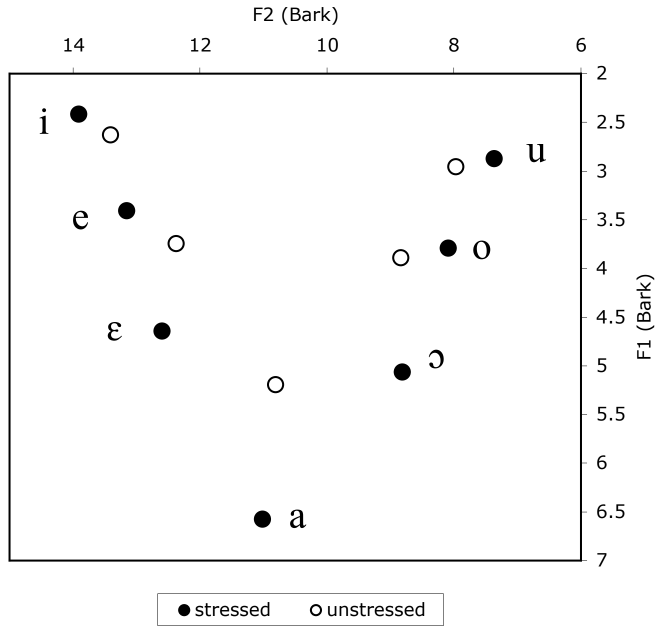


Fig. 10. Italian vowels (data from Albano Leoni et al 1995).

A final prediction that follows from relating vowel reduction to undershoot is that vowel reduction could arise in other contexts where vowel duration is short enough to give rise to significant undershoot. As noted in section 2, segmental factors such as a following voiceless stop or geminate commonly condition shorter vowel durations. The first point to make is that the model predicts that neutralizing reduction should only occur where vowel duration is particularly short, it does not predict a general correlation between vowel duration and inventory size. This follows from the nature of undershoot as discovered by Lindblom (1963): undershoot increases exponentially with decreasing vowel duration, so shortening vowels significantly below the duration of a typical stressed short vowel can result in neutralization, but increasing the duration above this point does not result in much lower undershoot. So the model does not predict that we should generally find more contrasts among long vowels compared to full short vowels. In fact long vowel inventories may be either larger or smaller than corresponding short vowel inventories (Maddieson 1984:128ff.), but this variation must be accounted for by constraints unrelated to undershoot. The same considerations imply that any segmental effects would have to result in very short vowel durations in order to give rise to vowel reduction.

There is some indication that non-stress related variation in duration influences vowel reduction. Barnes (2002:143ff.) discusses languages which have vowel reduction in unstressed syllables, but where reduction is blocked in phrase final position, a context in which vowels tend to be lengthened. In section 9, we discuss the case of Northwest Caucasian languages which contrast short and extra-short vowels, with a reduced inventory of extra-short vowels. This pattern can be analyzed as a form of duration conditioned vowel reduction. However, the typology of these kinds of effects of contextual variation in vowel duration are not well studied, so this prediction is in need of further investigation.

7. Vowel reduction primarily neutralizes height contrasts

As observed in section 2, vowel reduction primarily neutralizes height contrasts, not backness and rounding contrasts. The most common patterns of vowel reduction eliminate mid vowels, or the contrasts between lower mid and higher mid vowels, but do not neutralize front-back contrasts. Backness and rounding contrasts can be neutralized in reduction, but only under certain conditions. First, backness contrasts between low vowels are subject to neutralization as in Chamorro where the contrast between [æ] and [ɑ] is neutralized in unstressed syllables (Topping and Dungca 1973, Chung 1983). Backness and rounding contrasts between non-low vowels are generally only neutralized when most or all height contrasts are neutralized as well, as in English reduction to schwa where all vowel quality contrasts are neutralized. It is particularly striking that vowel reduction does not eliminate front rounded or non-low central vowels although these are typologically marked vowel qualities.

Our model of vowel reduction formalizes an explanation for these restrictions on the neutralization of backness and rounding contrasts based on Lindblom's (1963) observation that the strongest effect on duration reduction is to make low vowels more difficult to produce. Examining the results of Lindblom's study of vowel reduction, it is apparent that the most consistent effect of reduced duration across consonant contexts is raising of low vowels, as illustrated in Fig. 1. In the context between velar stops, vowel duration had little effect on F2 of non-low vowels, and the largest effect was on the F1 of low [a]. There was some variation depending on consonant context: there was more F2 undershoot in the context of dental stops, and less F1 and F2 undershoot in the context of bilabial stops, but overall the largest observed effect was raising of the low vowels. So undershoot primarily compresses the vowel space in the F1 dimension and has much less effect on the F2 dimension. Given these observations, we can understand the fact that vowel reduction primarily neutralizes F1 contrasts as resulting from the fact that raising low vowels leaves less room for realizing distinct F1 contrasts, whereas the scope for realizing distinct F2 contrasts is relatively unaffected by vowel duration, except where vowel duration becomes extremely short.

Van Son and Pols (1990, 1992) also found evidence that the strongest effect of vowel duration is on the F1 of low vowels. In a study of Dutch read speech, they found that F1 of low vowels was positively correlated with vowel duration, that is shorter low vowels were raised more. Correlations between F1 and duration in non-low vowels were weak, as were correlations between F2 and vowel duration, even when only vowels preceded and followed by coronals were considered (Van Son and Pols 1992). So the most consistent effect of reducing vowel duration was to lower the F1 of low vowels, as observed in Lindblom (1963).

As is apparent from Lindblom's analysis of his data, this pattern results from the fact that most consonants can substantially assimilate to the F2 of adjacent vowels, but cannot assimilate to the F1 of these vowels. In articulatory terms, assimilation in F2 primarily involves assimilation in tongue body and lip position, whereas assimilation in F1 involves assimilation in degree of constriction. Most consonants can assimilate to the lip position of an adjacent vowel, and lip-rounding gestures have been found to begin well before the acoustic onset of a rounded vowel (e.g. Benguerel and Cowan 1974, Perkell and Matthies 1992). Labial consonants are compatible with a wide range of tongue body shapes, and it has been found that much of the tongue body movement to the following vowel is completed before the release of a labial consonant (Löfqvist and Gracco 1999). Similarly, the precise place of articulation of a velar typically follows the tongue body position of adjacent vowels (Öhman 1966, Houde 1967). Coronals are generally somewhat more constrained in tongue body position because the tongue tip and blade ride on the tongue body so it is easier to form a coronal constriction if the tongue body moves cooperatively, but substantial anticipation of vowel tongue body position is still possible (Öhman 1966). Consequently, F2 adjacent to consonants of all places of articulation varies depending on the adjacent vowel, as observed by Lindblom (1963) and many subsequent studies (e.g. Krull 1987) because the positions of the relevant articulators in the vowel are generally anticipated during a preceding consonant and continue to influence the articulation of a following consonant.

On the other hand, assimilating to the F1 of a vowel is liable to result in radical changes in the manner of a consonant. This is clearest in the case of stops. Forming a complete closure at any place of articulation lowers F1 to its minimum frequency, so the higher F1 of a vowel cannot be anticipated while making a stop closure – any increase in F1 would imply loss of the stop closure. More generally, any consonant constriction in the upper half of the vocal tract lowers F1, so F1 is low in the context of all consonants other than pharyngeals and glottals (Stevens 1999). Assimilating to the higher F1 of a vowel would imply forming a less narrow constriction, i.e. failure to produce a consonantal constriction. In other words, the opening movement for a vowel can only begin at the offset of the preceding consonant, and the closing movement must be completed by onset of the following consonant, but many of the tongue body and lip positions required to realize vowel F2 can be anticipated during the preceding consonant and continue into the following consonant, so a wider range of vowel F2 values can be realized without necessitating excessively rapid articulator movements.

The conclusion that low vowels are particularly difficult to produce at short durations is supported by the Lehiste's (1970) observation that low vowels tend to be longer than higher vowels. Lehiste hypothesizes that this greater duration is required in order to make the substantial movement from a consonant to a low vowel and back again. Westbury and Keating (1980) provide data supporting this theory: they found that vowels with lower jaw positions had longer durations in a study of English. The absence of any comparable cross-linguistic effects of backness and rounding on vowel duration can be understood in terms of the observation that most consonants can assimilate to these aspects of adjacent vowels, so the magnitude of the movements between consonant and vowel are smaller.

The differential effects of vowel duration on F1 and F2 undershoot is incorporated into the model of vowel reduction presented in the previous section in terms of the factor c in equation (13) that reduces assimilation of the vowel towards the target F2 of the consonant context. This factor represents the fact that there is mutual assimilation between consonant and vowel, so the notional F2 target of the consonant is not generally realized, while the low F1 associated with consonants is consistently realized regardless of vowel context. The effects of this factor can be observed in Fig. 4, which shows that reduced vowel duration compresses the vowel space towards lower values of F1, but makes much less difference to the range of possible F2 values. Broadly speaking, reduced duration does not make it much more difficult to maintain most F2 contrasts: if an F2 contrast meets the minimum distance requirement in stressed syllables, it can do so in unstressed syllables also, unless unstressed vowels are very much shorter than stressed vowels. On the other hand, reduced duration does make it more difficult to maintain F1 contrasts, so these contrasts are liable to neutralization in unstressed syllables.

However the shape of the vowel space creates a conflict between the realization of height and backness contrasts in the lower region of the vowel space, and this can result in the elimination of F2-based contrasts between low vowels, as observed in Chamorro. The conflict arises because the range of possible F2 values narrows as F1 increases, giving the vowel space the familiar approximately triangular shape shown in Fig. 2 above. So the distinctiveness of F2 contrasts is improved by selecting vowels with lower F1, but this is liable to reduce the distinctiveness of F1 contrasts with higher vowels. For example, the F2 difference between [ɛ] and [ɔ] is larger than that between [æ] and [ɑ], but the former vowels are less distinct from high and mid vowels.

Given the nature of undershoot, the triangular shape of the vowel space is preserved under reduction in duration (as can be observed in Fig. 4), so the need to maintain the distinctiveness of F1 contrasts as the F1 range is reduced can put F2 contrasts between low vowels under pressure, and it can be optimal to give up the F2 contrast in order to realize a vowel with maximal F1. This is essentially what is observed in the vowel systems in Fig. 9 (a) and (c), where stressed vowel inventories with the same number of height distinctions between front and back vowels are reduced to inventories with a single lowest vowel. So the model is able to account for the neutralization of the [æ]-[ɑ] contrast in Chamorro. The same considerations are also relevant to understanding why a stressed vowel inventory [i, e, a, o, u] is often reduced to [i, ə, u], as in Russian, rather than [i, e, o, u]. Due to the triangular shape of the reduced vowel space, the latter inventory is not actually possible at short durations because the mid vowels must either be centralized, reducing the distinctiveness of the F2 contrast between them, or raised, resulting in less distinct F1 contrasts with [i] and [u] respectively. So it is often preferable to select the triangular unstressed inventory [i, ə, u], which maximizes distinctiveness of the contrasts between the high and non-high vowels.

The simulations reported in the previous section demonstrate that the model generates the commonly attested pattern in which only height contrasts are neutralized (Figs. 7 and 9b, c)), and a pattern in which a contrast between low front and back vowels is neutralized (Fig. 9a, d). We will now see that it can also derive patterns of reduction like those of Scots Gaelic (4) and Mantuan Italian (5) where central or front rounded vowels are retained while height contrasts are neutralized, and see that the nature of the distinctiveness constraint also plays a key role in accounting for this aspect of the typology of vowel reduction.

To model vowel systems like Scots Gaelic and Mantuan Italian, it necessary to derive inventories with non-peripheral vowels. This is achieved by increasing the weight attached to F2 differences, a, in the distinctiveness measure (6). This derives central vowels rather than front rounded vowels, since the model as currently formulated predicts that non-peripheral vowels should be near the middle of the possible F2 range, maximizing their distance from both front and back vowels. Derivation of front rounded vowels would require modeling the role of F3 in distinguishing front rounded and unrounded vowels. However, the current model is sufficient to demonstrate the general restrictions on elimination of F2 contrasts in vowel reduction.

In the simulations reported here, a is increased to 0.35, but all other parameters are as above ($k_1 = 1.5$, $\beta_1 = 0.008$, $k_2 = 1.5$, $c = 0.27$, $\beta_2 = 0.01$, $F2_l = 1400$ Hz). Stressed vowels are 160 ms in duration while unstressed vowels are 90 ms. The optimal vowel systems were determined as before.

A pattern of reduction comparable to Scots Gaelic is derived with $w_n = 6.6$. The resulting system is shown in Fig. 11, and consists of 9 vowels in stressed syllables and 6 vowels in unstressed syllables. Both inventories contain central vowels, although height contrasts are neutralized in unstressed position, including the contrast between high and mid central vowels.

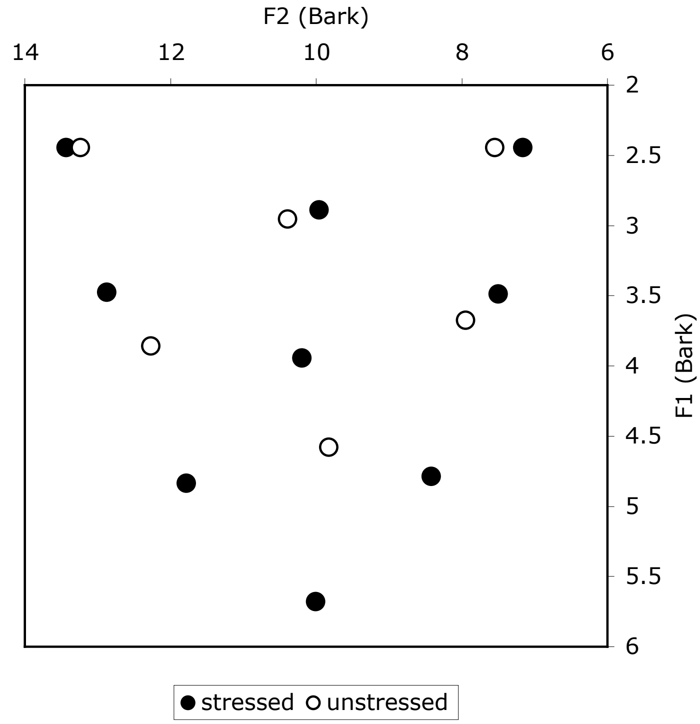
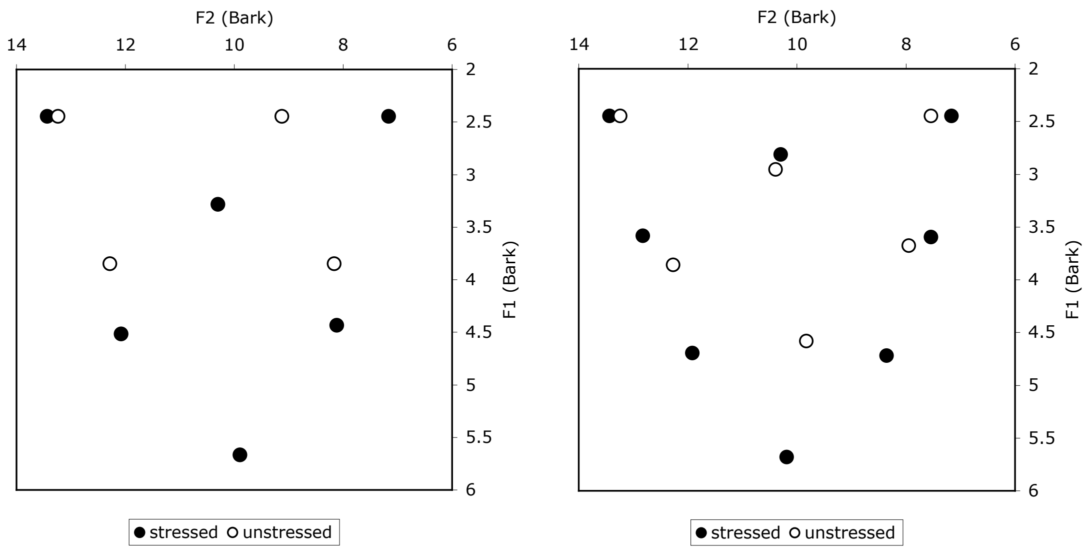


Fig. 11. Vowel system derived with $a = 0.35$ and $w_n = 6.6$ (see text for details).

The same parameter settings yield two additional systems with stressed central vowels when w_n is varied over the range from 2.6 to 6.6 (lower values derive systems without central vowels since central vowels are generally only included in larger inventories, and higher values result in more unstressed vowels than are generally attested in vowel reduction). These systems are shown in Fig. 12.



(a) (b)

Fig. 12. Additional vowel systems with stressed central vowels (see text for details).

Neither system directly corresponds to an attested system as far as I know, although (b) is comparable to the Scots Gaelic pattern in that a non-low central vowel is preserved while height contrasts are reduced. The vowel system in Fig. 12(a) shows that the model does not simply predict that central vowels are preserved in reduction. The pattern shown there could be characterized as reduction from [i, ɛ, a, ɔ, u, ə] in stressed syllables to [i, e, o, u] in unstressed syllables, eliminating the central vowel. But the central vowel [ə] in Fig. 12(a) is differentiated from [i, ɛ, ɔ, u] by substantial differences in F1 as well as a difference in F2. These F1 differences would be reduced significantly by undershoot, so these contrasts are predicted to be vulnerable to neutralization in short, unstressed syllables.

A more precise statement of the predictions of the model is that contrasts should be subject to elimination in unstressed syllables if their distinctiveness would be significantly reduced by reduced duration. As we have seen, reduced duration primarily results in raising of vowels with higher F1 (greater than 375 Hz, according to Lindblom (1963)), and has much less effect on vowel F2, so contrasts that are based purely on F1 are most vulnerable to reduction. But the F1 component of contrasts that are distinguished by differences on both dimensions are subject to significant reduction, so these 'mixed' contrasts are subject to neutralization also.

These predictions are consistent with the observation that contrasts that are based primarily on F2 are resistant to vowel reduction. It is less clear whether the prediction regarding mixed contrasts is confirmed. I am not aware of any pattern of vowel reduction resembling Fig. 12 (a). The most common types of contrast that involve significant differences in both F1 and F2 are probably between low central vowels and front and back mid vowels. These contrasts are commonly neutralized in unstressed syllables, but by eliminating the mid vowels (and raising the low vowel), as in reduction from [i, e, a, o, u] to [i, ə, u]. As discussed above, we don't generally find reduction to [i, e, o, u], eliminating the central vowel, due to the triangular shape of the vowel space. However, there is no evidence against the prediction that mixed contrasts are subject to neutralization in reduction.

8. Comparing distinctiveness constraints

We have argued that differences between undershoot in F1 and F2 provide the basis for explaining the generalization that vowel reduction primarily neutralizes height contrasts. However, it is important to realize that the measure of distinctiveness based on the minimum distance between vowels plays a crucial role in linking the generalization about undershoot to the generalization about phonological reduction. In the model developed here, the distinctiveness cost of a system of stressed and unstressed vowels is based on the minimum distance between any pair of contrasting vowels, so a contrast may be eliminated from unstressed syllables if undershoot would significantly reduce its distinctiveness compared to the stressed position. If the distance to neighboring vowels is not significantly reduced, it is preferable to retain the contrast to maximize the number of contrasting vowels. It is even worth tolerating a small reduction in distinctiveness as long as the increase in cost is less than the cost incurred by losing a contrast. Given that undershoot primarily affects F1, this implies that it is F1 contrasts that are prone to neutralization, as shown in detail above.

It is essential to this explanation that the distinctiveness cost only depends on the distances between vowels and their nearest neighbors, but not all measures of distinctiveness have this property. For example, L&L's measure of distinctiveness is more global - the cost incurred by a vowel depends on its distance from all of the other vowels in the inventory, not just the distance to the closest vowel. We will see that this means that the raising of low vowels can make a high central vowel very costly even if the distance to its nearest neighbor is not reduced much, so incorporating this measure of distinctiveness into the model of vowel reduction leads to the problematic prediction that non-low central vowels could be eliminated even when most height contrasts are preserved.

In L&L's model, the distinctiveness cost of a vowel inventory is the sum of the reciprocals of the squared distances between each pair of vowels (15) – in other words, the distance between each pair of vowels contributes to the distinctiveness cost of a vowel inventory.

(15) $$E = \sum_{i=1}^{n-1} \sum_{j=0}^{i-1} \frac{1}{d_{ij}^2}$$

This measure must be modified before it can be incorporated into the model of vowel reduction because the number of vowel pairs contributing to the sum, E, in an inventory of n vowels is $n(n\text{-}1)/2$, so dispersion cost tends to increase rapidly with inventory size simply as a result of the rapidly increasing number of pairs over which the sum in (15) is calculated. Accordingly, this measure does not provide a good basis for comparing the distinctiveness of vowel inventories of different sizes, because it would result in a strong preference for small inventories independent of the distances between the vowels. This was not an issue in L&L's model because they only compared inventories of the same size, but it is problematic in the present context because the analysis of reduction requires the comparison of larger and smaller vowel inventories. This problem can be rectified by normalizing the distinctiveness measure in (15), dividing it by the number of vowel pairs, $n(n\text{-}1)/2$. This normalized measure is essentially the average distinctiveness cost of a pair of vowels in the inventory, so it grows only as a function of increased crowding (i.e. decreased distances between vowels).

The problem with this cost function is that it penalizes vowels that are reasonably close to many other vowels because the sum of the costs $1/d^2$ for all of the pairs including that vowel can be large even if none of these distances is the smallest found in the inventory. This property leads to the strong dispreference observed in L&L's simulations for vowels like mid-central or mid front rounded vowels that lie in the interior of the vowel space. Such interior vowels are close to all of the peripheral vowels and thus contribute disproportionately to distinctiveness cost. So one strength of the minimum distance measure of dispersion is that it allows for the derivation of interior vowels, as in Figs. 11 and 12 above[7].

The global character of a distinctiveness measure like that proposed by L&L presents specific problems for the analysis of vowel reduction because it can result in particularly high costs for central vowels in unstressed positions, with the problematic consequence that it is predicted that central vowels can be eliminated in unstressed syllables, contrary to the generalizations discussed in section 2. This prediction results because reduction raises the lowest vowels in an inventory. As the lower vowels move closer to the high vowels, high central vowels become very costly because they are then close to all of the other vowels in the inventory. Consequently it is easy to derive unattested reduction patterns in which high central vowels are eliminated while three vowel heights are preserved, as in Fig. 13. This pattern of reduction was derived by substituting the square of the normalized L&L dispersion cost for the minimum distance cost employed above. The normalized dispersion cost had to be squared to derive a cost function that increases exponentially as the number of vowels increases, so that neither number cost nor distinctiveness cost can completely dominate the other. All parameters are as in the previous simulations, except the factor, *a*, by which F2 differences are reduced in the measure of distinctiveness (6) is set at 0.3. The pattern shown in Fig. 13 is optimal with $w_n = 14$.

---

[7] This is not to imply that minimum distance-based measure of distinctiveness is superior to L&L's measure in other respects. This measure also has weaknesses, for example, it can derive inventories in which the high vowels have higher F2 than their mid counterparts, a pattern which does not seem to be attested, although it is reminiscent of some vowel systems with Advanced Tongue Root distinctions (Ladefoged and Maddieson 1996:305). So, unsurprisingly, neither measure is perfect, but the minimum distance measure has definite advantages in the analysis of vowel reduction, as demonstrated below.

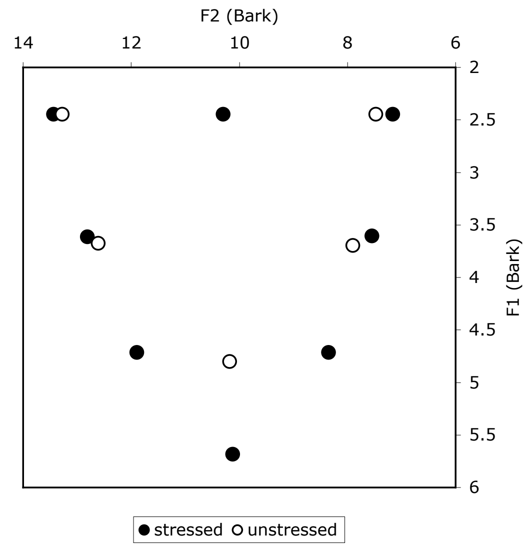


Fig. 13. An unattested pattern of vowel reduction derived using a function for distinctiveness cost based on L&L (see text for details).

It can be seen that the unstressed vowels closest to the high central vowel are not shifted much closer to it in the unstressed system, so given a minimum distance-based distinctiveness cost, a inventory of six vowels, retaining the high central vowel, would be superior to the 5 vowel unstressed inventory in Fig. 13. But with the global distinctiveness cost, adding a high central vowel to the unstressed inventory incurs a high cost due to its relative proximity to all five of the other vowels. This demonstrates that the observation that undershoot primarily affects F1 is not sufficient in itself to explain the observation that F1 contrasts are more prone to neutralization in vowel reduction, the nature of the dispersion constraint plays an essential role also. An appropriate model of vowel inventories is required to relate the observations about formant undershoot to generalizations about systems of vowel contrasts.

9. Further predictions of the model

In this section we explore further predictions of the model, considering variation in two parameters that have so far been held more or less constant: the duration of unstressed vowels, and the parameter *c* that governs the extent that consonants assimilate to the F2 of adjacent vowels.

If vowel duration is very short then contrasts of all types become difficult to maintain. Ultimately, this can result in complete neutralization of vowel quality contrasts as in English reduction to schwa. Similar patterns of reduction are found in some Southern Italian dialects. although sometimes only in post-tonic syllables (Maiden 1995). The undershoot-based model predicts that complete neutralization should only arise where vowel duration is very short so there is significant undershoot in both F1 and F2, making it impossible to maintain any adequately distinct contrasts. This is certainly the case in English where Kondo (2000) found that the reduced 'schwa' vowel had an average duration of 34 ms. The undershoot model further predicts that the single vowel quality in this type of reduction should be contextually variable, being fully assimilated to its context (not the mid central vowel that is implied by transcribing these vowels as [ə]). This prediction is confirmed for English by Kondo (1994). The same kind of systematic variability has been observed in the realization of Dutch schwa, which is a similar short, reduced vowel, although the nature of vowel reduction in Dutch is rather different than in English (van Bergem 1994, Koopmans-van Beinum 1994).

As for variation in *c*, we have seen in the previous section that low values of *c* are typical, and that this provides the basis for the explanation of the typological restrictions on neutralization of F2 contrasts in vowel reduction. However, the model predicts that if there are circumstances under which *c* adopts higher values, then neutralization of F2 contrasts should be more prevalent in vowel reduction under these conditions. We will see that there are consonants that resist assimilation in F2 (and thus have a high value of *c*), but the predicted patterns of reduction have not been observed, although similar patterns do arise in languages with systems of extra short vowels such as Kabardian.

We would expect that consonants would resist assimilation in F2 where F2 transitions are important to the realization of consonant contrasts. This is the case with contrastively palatalized consonants, for example, and Richey (2000) finds that F2 at the release of Russian palatalized consonants varies little with vowel context. Russian is also a language with vowel reduction, and thus provides a suitable test case for the model. There is a clear effect of palatalization on the pattern of vowel reduction: standard Russian has five vowels in stressed syllables, [i, e, a, o, u], and three ([i, ə, u]) in most unstressed syllables, but only two ([i, u]) appear following palatalized consonants. However, this pattern is likely to be due as much to the strong raising effect of palatalized consonants as to their fronting effect. That is, extreme raising of the low vowel reduces the F1 difference between the central vowel and [i, u] to the point where the contrasts are no longer tenable. The fronting effect of palatals would also contribute to the difficulty of maintaining three distinct vowels, but on its own cannot derive an inventory with two high vowels in the model proposed here. If the primary effect of palatalized consonants were on F2 of adjacent vowels, we would expect some difference in height between the two remaining vowels. It is plausible that palatalized consonants should have a stronger raising effect than plain consonants since palatalization involves a high front tongue body constriction which is sustained after the release of the primary constriction.

The effects of secondary articulations like palatalization and velarization might be more apparent in languages with front rounded or non-low central vowels – e.g. we might expect neutralization of these contrasts adjacent to palatalized consonants, in unstressed syllables. However, testing this prediction would involve finding a number of languages with contrastive secondary articulations, front rounded or non-low central vowels and vowel reduction – a combination that is not expected to be frequent, and is not attested as far as I know. The predicted effect of F2 neutralization among extra short vowels is observed in Northwest Caucasian languages such as Kabardian, and Shapsug, but not in the context of stress-conditioned vowel reduction.

Kabardian (Kuipers 1960, Colarusso 1992, Choi 1991) and Shapsug (Smeets 1984) have consonant inventories that include contrasts between an unusually large number of places of articulation, including some secondary articulations, so it is plausible that F2 transitions should be particularly important in realizing these distinctions, and accordingly resistant to assimilation. Although these languages do not have stress conditioned vowel reduction, they do distinguish short and extra-short vowels. Each has five short vowels [i, e, a, o, u] (Kuipers 1960:23f., Smeets 1984:123), and two extra-short vowels, which can be transcribed broadly as [ɨ, ə], although the F2 of these vowels is actually dependent on the consonant context (Choi 1991). The absence of F2 contrasts among extra short vowels can be analyzed as a consequence of high levels of F2 undershoot, due to the high values of *c* hypothesized to characterize Northwest Caucasian consonants (cf. Flemming 2004). Comparable patterns of F2 neutralization are observed among short medial vowels in Marshallese, a language with an extensive system of palatalization and velarization contrasts (Bender 1968, Choi 1992, Flemming 2004).

10. Reduced vowel inventories in non-initial syllables

Finally, we return briefly to the analysis of languages with richer inventories of vowel contrasts in initial syllables. As noted in section 2, this pattern can give rise to the appearance of anomalous systems of vowels reduction where stress also falls on the initial syllable. For example, Estonian has nine vowel qualities [i, e, æ, y, ø, a, o, u, ɤ] in initial stressed syllables, but does not allow mid back unrounded [ɤ] in non-initial syllables. This is superficially a counterexample to the restrictions on F2 neutralization analyzed above, because the only contrast that is neutralized in unstressed syllables is a rounding contrast. However, as noted by Steriade (1993) and Beckman (1998, chapter 2), there is evidence that initial position is a location in which we find larger numbers of vowel contrasts, independently of stress, so Estonian can be analyzed as a case of non-initial reduction, not unstressed vowel reduction.

The key evidence for this line of analysis is the existence of languages like Turkish and Shona in which stress is non-initial, but the full range of vowel contrasts is only observed in initial syllables. In Turkish, the non-high rounded vowels [ø, o] can only appear in initial syllables, but stress is usually final (Inkelas and Orgun 2003). It is also clear that the distributional restrictions are not related to vowel duration, since vowels in initial syllable are only very slightly longer than vowels in second syllables (Barnes 2002: chapter 4). In Shona, mid vowels only arise in non-initial syllables through vowel harmony (Beckman 1998), but stress is penultimate (Stevick 1965). The prediction is then that the appearance of anomalous F2-neutralizing patterns of vowel reduction should only involve languages with initial stress. This is the case in Estonian, and in the Mongolian and Turkic languages, most of which only allow a full set of vowel contrasts in initial syllables, and many of which have initial stress (Barnes 2002).

A number of researchers have suggested that the realization of a greater number of contrasts in initial syllables reflects the importance of initial syllables for lexical access (Beckman 1998, MacEachern 1997:148f., Smith 2002, see Barnes 2002 for a dissenting view). Maximizing the number of contrasts early in a word means that words can be differentiated more quickly, facilitating rapid word recognition (MacEachern 1997:148f., Nooteboom 1981). This reasoning directly motivates realizing more contrasts in the first syllable of a word – in the terms of the model proposed here, this implies the possibility of a higher value of the weight w_n, favoring maximization of the number of contrasts, in word initial syllables. This would derive the possibility of selecting a larger vowel inventory in this position, but crucially the difference in inventory size is not motivated by undershoot, and so the patterns of 'reduction' in non-initial syllables are not expected to be the same as in stress-conditioned reduction. We expect the vowels that are excluded from non-initial positions to be the kinds of vowels that are excluded from inventories of that size cross-linguistically. This is the case in Estonian, where elimination of the back unrounded vowel yields an eight vowel inventory essentially the same as the full vowel inventory of Finnish. However, the analysis of vowel distribution in Estonian, Shona, and in the Mongolian and Turkic languages would require additional constraints beyond those considered here since these languages also have vowel harmony processes.

11. Conclusions

We have seen that the phenomenon of vowel reduction can be derived from three independently motivated constraints on sound systems: maximize the distinctiveness of contrasts, maximize the number of contrasts, and minimize effort. This has been demonstrated in the context of an explicit model based on fairly direct formulations of these fundamental constraints. Constraints of this kind have been employed in previous models of vowel inventories, beginning with the work of Liljencrants and Lindblom (1972), but the key innovations required for the analysis of vowel reduction are to model aspects of the prosodic and segmental context of vowels, and to allow for different vowel inventories in different contexts. Contextual restrictions on the distributions of sounds are the core of phonology, so it is necessary to extend models of sound systems in this direction in order to address a wider range of phonological phenomena.

For the analysis of vowel reduction, we have argued that it is necessary to consider how vowel duration varies as a function of vowel context and the consequences for vowel production, in particular formant undershoot. Relating phonological vowel reduction to phonetic undershoot helps us to explain the typological generalization that vowel reduction primarily neutralizes height contrasts, and only neutralizes backness and rounding contrasts under restricted conditions.

Providing an explicit formalization of the model of vowel reduction makes it clear how all components of the model contribute to deriving the observed generalizations. In particular, we have seen that the specific form of the distinctiveness constraints and the shape of the space of possible vowels play important roles in explaining the restrictions on neutralization of backness and rounding contrasts. The role of the distinctiveness constraints was not obvious from an informal presentation of the explanation.

The distinctiveness constraint proposed here is novel in that it measures the overall distinctiveness of a system of vowels in terms of the distance between the closest pair of vowels in the system. This measure is inspired by work in the theory of digital communications systems but differs from most previous work on vowel systems which has followed L&L in summing costs related to the distances between all pairs of vowels. We have seen that the minimum distance measure has advantages in the analysis of vowel reduction, and can derive interior vowels more easily than the L&L measure, so this type of distinctiveness constraint merits further exploration.

Acknowledgements

Many thanks to audiences at the ZAS Conference on the Phonetics-Phonology Interface, UCLA, Stanford, and MIT for comments on this research at various stages in its development, and to Donca Steriade for helpful comments on this paper.

References

Albano Leoni, F., M.R. Caputo, L. Cerrato, F. Cutugno, P. Maturi, and R. Savy (1995). Il vocalismo dell'Italiano. Analisi di un campione televisivo. *Studi Italiani di Linguistica Teorica e Applicata* 24, 405-411.

Anderson, John B. (1999). *Digital transmission engineering*. IEEE Press, New York.

Atal, B.S., Chang, J.J., Mathews, M.V., & Tukey, J.W. (1978). Inversion of articulatory-to-acoustic transformation in the vocal tract by a computer-sorting technique. *Journal of the Acoustical Society of America* 63, 1535-1555.

Barker, M.A.R. (1964). *Klamath grammar*. University of California Publications in Linguistics 32, University of California Press, Berkeley.

Barnes, Jonathan (2002). *Positional Neutralization: A Phonologization Approach to Typological Patterns*. PhD dissertation, University of California, Berkeley.

Beckman, Jill (1998). *Positional Faithfulness*. PhD dissertation, University of Massachusetts, Amherst.

Bender, Byron W. (1968). Marshallese phonology. *Oceanic Linguistics* 7, 16-35.

Benguerel, A. P., and H. A. Cowan (1974). Coarticulation of upper lip protrusion in French. *Phonetica*, 30, 41—55.

Berinstein, Ava (1979). A cross-linguistic study on the perception and production of stress. *UCLA working papers in phonetics 47*, University of California, Los Angeles.

Borgstrom, C. (1940). *A linguistic survey of the Gaelic dialects of Scotland*, vol. 1*: The dialects of the Outer Hebrides*. Norsk Tidsskrift for Sprogvidenskap, Suppl. Bind 1.

Bosch, L.F.M. ten, L.J. Bonder, and L.C.W. Pols (1987) Static and dynamic structure of vowel systems. *Proceedings of the 11th international congress of phonetic sciences,* Vol.1, 235-238.

Card, Elizabeth (1983). *A phonetic and phonological study of Arabic emphatics*. PhD Dissertation, Cornell University, Ithaca, NY.

Chen, Matthew (1970). Vowel length variation as a function of the voicing of consonant environment. *Phonetica* 22, 129-159.

Choi, John D. (1991). An acoustic study of Kabardian vowels. *Journal of the International Phonetic Association* 21, 4-12.

Choi, John D. (1992). *Phonetic Underspecification and Target Interpolation: An Acoustic Study of Marshallese Vowel Allophony (UCLA Working Papers in Phonetics 82)*. Ph.D. dissertation, University of California, Los Angeles.

Chung, Sandra (1983). Transderivational relationships in Chamorro phonology. *Language* 59, 35-66.

Clements, George N. (2003). Feature economy in sound systems, *Phonology* 20, 287-333.

Colarusso, John (1992). *A Grammar of the Kabardian Language*. University of Calgary Press, Calgary.

Crosswhite (1999). *Vowel reduction in Optimality Theory*. Ph.D. dissertation, UCLA.

Crosswhite (2001). Vowel reduction. To appear in B. Hayes, R. Kirchner, and D. Steriade (eds.) *Phonetically-Based Phonology*. CUP.

de Boer, Bart (2001). *The Origins of Vowel Systems*. Oxford University Press, Oxford.
Diehl, Randy L., Björn Lindblom, and Carl P. Creeger (2003). Increasing realism of auditory representations yields further insights into vowel phonetics. *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona.*
Everett, Daniel, and Keren Everett (1984). On the relevance of Syllable Onsets to Stress Placement. *Linguistic Inquiry* 15, 705-711.
Fails, Willis C., and J. Halvor Clegg (1992). A spectrographic analysis of Portuguese stressed and unstressed vowels. Donald P. Macedo and Dale A. Koike (eds.) *Romance Linguistics: The Portuguese Context*. Bergin and Garvey, Westport, 31-42.
Flemming, Edward (2001). *Auditory Representations in Phonology*. Routledge, New York.
Flemming, Edward (2004). Contrast and perceptual distinctiveness. Bruce Hayes, Robert Kirchner, and Donca Steriade (eds.) *Phonetically-Based Phonology*. Cambridge University Press, 232-276.
Greenberg, Marc L. (2000). *A Historical Phonology of the Slovene Language.* Universitätsverlag C. Winter, Heidelberg.
Hagiwara, Robert (1995). Acoustic realizations of American /r/ as produced by women and men. *UCLA Working Papers in Phonetics* 90.
Halle, Morris (1959). *The Sound Pattern of Russian*. Mouton, The Hague.
Herrick, Dylan (2003), An Acoustic Analysis of Phonological Vowel Reduction in Six Varieties of Catalan. PhD dissertation, University of California, Santa Cruz.
Houde, Richard A. (1967*). A study of tongue body motion during selected speech sounds.* Ph.D. dissertation, University of Michigan.
Inkelas, Sharon, and Orhan Orgun (2003). Turkish stress: a review. *Phonology* 20, 139-161.
Klatt, Dennis H. (1987). Review of text-to-speech conversion for English. *Journal of the Acoustical Society of America* 82, 737-793.
Kondo, Yuko (1994). Targetless schwa: is that how we get the impression of stress timing in English? *Proceedings of the Edinburgh Linguistics Department Conference '94*, 63-76.
Kondo, Yuko (2000). Production of schwa by Japanese speakers of English: an acoustic study of shifts in coarticulatory strategies from L1 to L2. Michael Broe and Janet Pierrehumbert (eds.) *Papers in Laboratory Phonology 5: Acquisition and the Lexicon.* Cambridge University Press, Cambridge, 29-39.
Koopmans-van Beinum, Florien (1994). What's in a schwa? *Phonetica* 51, 68-79.
Krull, Diana (1987). Second formant locus patterns as a measure of consonant-vowel coarticulation. *Phonetic Experimental Research at the Institute of Linguistics, University of Stockholm* 5, 43-61.
Kuipers, Aert H. (1960). *Phoneme and Morpheme in Kabardian (Eastern Adyghe)*. Janua Linguarum, series minor, no. 8. Mouton, The Hague.
Ladefoged, Peter (2001). *Vowels and Consonants: An introduction to the sounds of languages*. Blackwell, Oxford.
Ladefoged, Peter, and Ian Maddieson (1996). *The Sounds of the World's Languages.* Blackwell, Oxford.
Lehiste, Ilse (1970). *Suprasegmentals*. MIT Press, Cambridge.
Lehiste, Ilse, and K. Popov (1970). Akustiche Analyse bulgarischer Silbenkerne. *Phonetica* 21, 40-48.
Lencek, Rado L. (1982). *The Structure and History of the Slovene Language*. Slavica, Columbus, Ohio.

Liljencrants, Johan, and Björn Lindblom (1972). Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language* 48, 839-62.
Lindblom, Björn (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America* 35, 1773-1781.
Lindblom, Björn (1986). Phonetic universals in vowel systems. J.J. Ohala and J.J. Jaeger (eds) *Experimental Phonology*. Academic Press.
Lindblom, Björn (1990a). On the notion of "possible speech sound". *Journal of Phonetics* 18, 135-152.
Lindblom, Björn (1990b). Phonetic content in phonology. *PERILUS* 11, 101-118.
Lindblom , Björn (1998). Systemic constraints and adaptive change in the formation of sound structure. James R. Hurford, Michael Studdert-Kennedy and Chris Knight (eds.) *Approaches to the Evolution of Language: Social and Cognitive Bases*. Cambridge University Press, Cambridge, 242-264.
Liphola, Marcelino M. (2001). *Aspects of Phonology and Morphology of Shimakonde.* PhD dissertation, Ohio State University.
Löfqvist, Anders, and Vincent Gracco (1999). Interarticulator programming in VCV sequences: Lip and tongue movements. *Journal of the Acoustical Society of America* 105, 1864-1876.
MacEachern, Margaret (1997). *Laryngeal cooccurrence restrictions*. PhD dissertation, University of California, Los Angeles.
Maddieson, Ian (1984). *Patterns of Sounds*. Cambridge University Press, Cambridge.
Maddieson, Ian (1985). Phonetic cues to syllabification. Victoria A. Fromkin (ed.) Phonetic Linguistics. Academic Press, New York, 203-221.
Maiden, Martin (1995). Vowel systems. Martin Maiden and Mair Parry (eds.) *The Dialects of Italy*. Routledge, London, 7-14.
Majidi, Mohammad-Reza (1991). Illustrations of the IPA: Persian: Farsi. *Journal of the International Phonetic Association* 21, 96-98.
Mattoso Camara, Joaquim (1972). *The Portuguese Language*. University of Chicago Press, Chicago.
Mazzola, Michael L. (1976). *Proto-Romance and Sicilian*. Peter de Ridder, Lisse.
McDonough, Joyce (2003). *The Navajo sound system*. Kluwer, Dordrecht.
Miglio, Viola (1996). Mantuan vowel shifts. *University of Maryland Working Papers* 4, 103-124.
Moon, Seung-Jae. and Björn Lindblom (1994). Interaction between duration, context, and speaking style in English stressed vowels. *Journal of the Acoustical Society of America* 96, 40-55.
Nelson, W.L. (1983). Physical principles for economies of skilled movement. *Biological Cybernetics* 46, 135-147.
Nooteboom, Sieb G. (1981). Lexical retrieval from fragments of spoken words: beginnings vs. endings. *Journal of Phonetics* 9, 407-424.
Ohala, John J. (1980). Moderator's introduction to Symposium of Phonetic Universals in Phonological Systems and their Explanation. *Proceedings of the Ninth International Congress of Phonetic Sciences (Vol. 3)*. Institute of Phonetics, Copenhagen, 181-185.
Öhman, S.E.G. (1966). Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America* 39, 151–168.

Padgett, Jaye, and Marija Tabain (2003). Adaptive dispersion theory and phonological vowel reduction in Russian. Ms., UC Santa Cruz and Macquarie University, Sydney.

Palková, Zdena (1994). *Fonetika a Fonologie Cestiny*. Universita Karlova, Prague.

Payne, David L. (1981). *The phonology and morphology of Axininca Campa*. Summer Institute of Linguistics, Arlington, Texas.

Perkell, Joseph S., and Melanie L. Matthies (1992). Temporal measures of anticipatory coarticulation for the vowel /u/: Within- and cross-subject variability. *Journal of the Acoustical Society of America* 91, 2911-2925.

Perkell, Joseph S., Majid Zandipour, Melanie L. Matthies, and Harlan Lane (2002). Economy of effort in different speaking conditions. I. A preliminary study of intersubject differences and modeling issues. *Journal of the Acoustical Society of America* 112, 1627-41.

Podesva, Robert, and Niken Adisasmito-Smith (1999). Acoustic analysis of the vowel systems of Buginese and Toba Batak. *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco, 535-538.

Recasens, Daniel (1991). *Fonètica Descriptiva del Català*. Institut D'Estudis Catalans, Biblioteca Filològica XXI, Barcelona.

Richey, Colleen (2000). Cues for place and palatalization of Russian stops. Ms, Stanford University.

Schwartz, Jean-Luc; Louis-Jean Boë,; Nathalie Vallée and Christian Abry (1997). The dispersion-focalization theory of vowel systems. *Journal of Phonetics* 25, 255-286.

Smeets, Rieks (1984). Studies in West Circassian Phonology and Morphology. Hakuchi Press, Leiden.

Smith, Jennifer L. (2002). *Phonological augmentation in prominent positions*. PhD dissertation, University of Massachusetts, Amherst.

Steriade, Donca (1993). Positional neutralization. Talk presented at NELS 24, University of Massachusetts, Amherst.

Stevens, Kenneth N. (1999). *Acoustic Phonetics*. MIT Press, Cambridge.

Stevick, Earl (1965). *Shona; Basic Course*. Department of State, Washington D.C.

Topping, Donald, and Bernadita Dungca (1973). *Chamorro Reference Grammar*. University of Hawaii Press, Honolulu.

Vallée, Nathalie, Jean-Luc Schwartz and Pierre Escudier (1999). Phase spaces of vowel systems. A typology in the light of the Dispersion-Focalization Theory (DFT). *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco, 333-336

Van Bergem, Dick R. (1994). A model of coarticulatory effects on the schwa. *Speech Communication* 14, 143-162.

Van Son, R.J.J.H., and Louis C.W. Pols (1990). Formant frequencies of Dutch vowels in a text, read at normal and fast rate. *Journal of the Acoustical Society of America* 88, 1683-1693.

Van Son, R.J.J.H., and Louis C.W. Pols (1992). Formant movements of Dutch vowels in a text, read at normal and fast rate. *Journal of the Acoustical Society of America* 92, 121-127.

Westbury, John, and Patricia A. Keating (1980). Central representation of vowel duration. *Journal of the Acoustical Society of America* 67, S37A.

Wouters, Johan, and Michael W. Macon (2002). Effects of prosodic factors on spectral dynamics. I. Analysis. . *Journal of the Acoustical Society of America* 111, 417-427.