

To appear in B.Hayes, R.Kirchner, D.Steriade (eds)
The Phonetic Bases of Phonological Markedness
Cambridge, University Press

Contrast and Perceptual Distinctiveness

Edward Flemming
Stanford University

1 Introduction

Most ‘phonetically-driven’ or functionalist theories of phonology propose that two of the fundamental forces shaping phonology are the need to minimize effort on the part of the speaker and the need to minimize the likelihood of confusion on the part of the listener. The goal of this paper is to explore the perceptual side of this story, investigating the general character of the constraints imposed on phonology by the need to minimize confusion.

The need to avoid confusion is hypothesized to derive from the communicative function of language. Successful communication depends on listeners being able to recover what a speaker is saying. Therefore it is important to avoid perceptually confusable realizations of distinct categories; in particular distinct words should not be perceptually confusable. The phonology of a language regulates the differences that can minimally distinguish words, so one of the desiderata for a phonology is that it should not allow these minimal differences, or contrasts, to be too subtle perceptually. In Optimality Theoretic terms, this means that there are constraints favoring less confusable contrasts over more confusable contrasts.

There is nothing new about the broad outlines of this theory (cf. Lindblom 1986, 1990, Martinet 1955, Zipf 1949, among others), but it has important implications for the nature of phonology. First, it gives a central role to the auditory-perceptual properties of speech sounds since distinctiveness of contrasts is dependent on perceptual representation of speech sounds. This runs counter to the articulatory bias in phonological feature theory observed in Chomsky and Halle (1968) and its successors. Substantial evidence for the importance of perceptual considerations in phonology has already been accumulated (e.g. Boersma 1998, Flemming 1995, Jun this volume, Steriade 1995, 1997, Wright this

volume; see also Hume and Johnson forthcoming pp.1-2 and references cited there). This paper provides further evidence for this position, but the focus is on a second implication of the theory: the existence of constraints on contrasts. Constraints favoring distinct contrasts are constraints on the differences between forms rather than on the individual forms themselves. We will see that paradigmatic constraints of this kind have considerable implications for the architecture of phonology.

The next section discusses why we should expect perceptual markedness to be a property of contrasts rather than individual sounds and previews evidence that this is in fact the case. Then constraints on contrast will be formalized within the context of a theory of phonological contrast. The remainder of the paper provides evidence for the key prediction of the theory: the markedness of a sound depends on the sounds that it contrasts with.

2 Perceptual markedness is a property of contrasts

The nature of the process of speech perception leads us to expect that any phonological constraints motivated by perceptual factors should be constraints on contrasts, such as the contrast between a back unrounded vowel and a back rounded vowel, not constraints on individual sounds, such as a back unrounded vowel. Speech perception involves segmenting a speech signal and categorizing the segments into a pre-determined set of categories such as phonetic segments and words. The cues for classification are necessarily cues that a stimulus belongs to one category as opposed to another. So we cannot talk about cues to a category, or how well a category is cued by a particular signal without knowing what the alternatives are. For example, it is not possible to say that a back unrounded vowel presents perceptual difficulties without knowing what it contrasts with. It is relatively difficult to distinguish a back unrounded vowel from a back rounded vowel so if a language allows this contrast the back unrounded vowel can be said to present perceptual difficulties, and the same can be said of the back rounded vowel. But if it is known that a back unrounded vowel is the only vowel which can appear in the relevant context, then all the listener needs to do is identify that a vowel is present as opposed to a consonant, which is likely to be unproblematic.

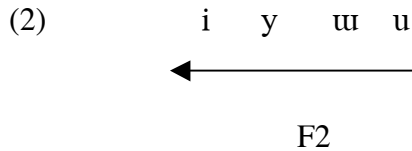
Perceptual difficulty is thus very different from articulatory difficulty. Articulatory difficulty can be regarded as a property of an individual sound in a particular context because it relates to the effort involved in producing that sound. There is no analogous notion of effort involved in perceiving a sound – perceptual difficulties don't arise because particular speech sounds tax the auditory system, the difficulty arises in correctly categorizing sounds. Thus it does not seem to be possible to provide a sound basis in perceptual phonetics for constraints on the markedness of sounds independent of the contrasts that they enter into. This point is assumed in Liljencrants and Lindblom's models of how perceptual factors shape vowel inventories (Liljencrants and Lindblom 1972, Lindblom 1986), and similar considerations are discussed in Steriade (1997).

The difference between regarding perceptual markedness as a property of contrasts rather than sounds can be clarified through consideration of alternative approaches to the analysis of correlations between backness and lip rounding in vowels. Cross-linguistically, front vowels are usually unrounded whereas non-low back vowels are usually rounded. This is true of the common five vowel inventory in (1), and in the UPSID database as a whole, 94.0% of front vowels are unrounded and 93.5% of back vowels are rounded (Maddieson 1984).

- (1) i u
 e o
 a

The perceptual explanation for this pattern is that co-varying backness and rounding in this way maximizes the difference in second formant frequency (F2) between front and back vowels, thus making them more distinct. In general front and back vowels differ primarily in F2, with front vowels having a high F2 and back vowels having a low F2. Lip-rounding lowers F2 so the maximally distinct F2 contrast is between front unrounded and back rounded vowels (Liljencrants and Lindblom 1972, Stevens, Keyser and Kawasaki 1986). This is illustrated in (2) which shows the approximate positions of front

and back rounded and unrounded vowels on the F2 dimension¹. It can be seen that the distinctiveness of contrasts between front and back rounded vowels, e.g. [y-u], or between front and back unrounded vowels, e.g. [i-ʊ], is sub-optimal.



The standard phonological analysis of this pattern of covariation is to posit feature co-occurrence constraints against front rounded vowels and back unrounded vowels (3).

- (3) *[-back, +round]
 *[+back, -round]

This analysis does not correspond to the perceptual explanation outlined above. The constraints in (3) imply that front rounded vowels and back unrounded vowels are marked sounds, whereas the perceptual explanation implies that it is the contrasts involving front rounded vowels and back unrounded vowels that are dispreferred because they are less distinct than the contrast between a front unrounded vowel and a back rounded vowel. In Optimality Theoretic terms, there is a general principle that contrasts are more marked the less distinct they are, which implies a ranking of constraints as in (4), where *X-Y means that words should not be minimally differentiated by the contrast between sounds X and Y. (More general constraints which subsume these highly specific constraints will be formulated below).

- (4) *y-ʊ >> *i-ʊ, *y-u >> *i-u

¹ Positioning of vowels is based on modeling of the effects of lip-rounding reported in Stevens (1999:291-3). There is significant cross-linguistic variation in the F2 of high front rounded vowels (Schwartz, Beautemps, Abry, & Escudier 1993) depending on differences in the precise position of the tongue body and degree of lip-rounding/protrusion, and the same probably applies to back unrounded vowels.

These two accounts make very different predictions: Constraints on the distinctiveness of contrasts predict that a sound may be marked by virtue of the contrasts it enters into. If there are no constraints on contrasts, then the markedness of contrasts should depend simply on the markedness of the individual sounds, and should be insensitive to the system of contrasts. We will see a range of evidence that markedness of sounds is indeed dependent on the contrasts that they enter into – i.e. that there are markedness relations over contrasts as well as over sounds – and that the relative markedness of contrasts does correspond to their distinctiveness.

For example, the dispreference for front rounded vowels and back unrounded vowels extends to other vowels with intermediate F2 values, such as central vowels. Most languages contrast front and back vowels, and if they have central vowels, they are in addition to front and back vowels. The same explanation applies here also: since central vowels like [ɨ] fall in the middle of the F2 scale in (2), contrasts like [i-ɨ] and [i-u] are less distinct than [i-u] and consequently dispreferred. But we will see in 4.1 that in the absence of front-back contrasts, vowels with intermediate F2 values, such as central vowels, are the unmarked case in many contexts. A number of languages, including Kabardian (Kuipers 1960, Choi 1991), and Marshallese (Bender 1968, Choi 1992), have short vowel inventories which lack front-back contrasts. These so-called ‘vertical’ vowel systems consist of high and mid, or high, mid, and low vowels whose backness is conditioned by surrounding consonants, resulting in a variety of specific vowel qualities, many of which would be highly marked in a system with front-back contrasts, e.g. central vowels, back unrounded vowels, and short diphthongs. Crucially there are no vertical vowel inventories containing invariant [i] or [u], vowels which are ubiquitous in non-vertical inventories. I.e. there are no vowel inventories such as [i, e, a] or [u, o, a].

This pattern makes perfect sense in terms of constraints on the distinctiveness on contrasts: as already discussed central vowels are not problematic in themselves, it is the contrast between front and central or back and central vowels which is marked. In the absence of such F2-based contrasts, distinctiveness in F2 becomes irrelevant, and minimization of effort becomes the key factor governing vowel backness. Effort

minimization dictates that vowels should accommodate to the articulatory requirements of neighboring consonants.

These generalizations about vertical vowel systems show that the markedness of sounds depends on the contrasts that they enter into because sounds such as central vowels, which are marked when in contrast with front and back vowels, can be unmarked in the absence of such contrasts. The same pattern is observed in vowel reduction: when all vowel qualities are neutralized in unstressed syllables, as in English, the result is typically a ‘schwa’ vowel – a vowel type which is not permitted in stressed syllables in the same languages. This type of contrast-dependent markedness cannot be captured in terms of constraints on individual sounds. Ní Chiosáin and Padgett (1997) succinctly formulate the problem for theories without constraints on contrast as follows: the cross-linguistic preference for front unrounded and back rounded vowels over central vowels suggests a universal ranking of segment markedness constraints as shown in (5), which implies any language with [ɨ] will have [i, u] also. But this would imply that if only one of these vowels appears it should be [i] or [u], and certainly not a central vowel. More generally, this approach incorrectly predicts that if a sound type is unmarked, it should be unmarked regardless of the contrasts it enters into.

(5) *ɨ >> *u, *i

According to the contrast-based analysis proposed here, the dispreference for central vowels is more accurately a dispreference for the sub-maximally distinct contrasts between central and front or back vowels, i.e. constraints of the form shown in (6). These constraints are simply irrelevant where no such contrast is realized, so vowel markedness is determined by other constraints – in this case minimization of effort. This analysis is developed in 4.1, below.

(6) *i-i, *i-u >> *i-u

The focus of this paper is these constraints on the distinctiveness of contrasts, and their implications for phonology. However it is also essential to consider general constraints, such as effort minimization, that limit the distinctiveness of contrasts since actual contrasts are less than maximally distinct. So the first step is to situate constraints on the distinctiveness of contrasts within the context of a theory of phonological contrast. This is the topic of the next section. This model will then be applied to the analysis of particular phenomena, demonstrating the range of effects of distinctiveness constraints, and the difficulties that arise for models that do not include constraints on contrasts.

3 The dispersion theory of contrast

Constraints on the distinctiveness of contrasts are formalized here as part of a theory of contrast dubbed the ‘dispersion theory’ after Lindblom’s (1986, 1990) ‘Theory of Adaptive Dispersion’, which it resembles in many respects. The core of the theory is the claim that the selection of phonological contrasts is subject to three functional goals:

- i. Maximize the distinctiveness of contrasts
- ii. Minimize articulatory effort
- iii. Maximize the number of contrasts

As noted above, a preference to maximize the distinctiveness of contrasts follows from language’s function as a means for the transmission of information. This tendency is hypothesized to be moderated by two conflicting goals. The first is a preference to minimize the expenditure of effort in speaking, which appears to be a general principle of human motor behavior not specific to language. The second is a preference to maximize the number of phonological contrasts that are permitted in any given context in order to enable languages to differentiate a substantial vocabulary of words without words becoming excessively long.

These ideas are not new. They have antecedents in the work of Passy (1891) and Zipf (1949), for example, and have been developed in detail by Martinet (1952, 1955) and Lindblom (1986, 1990). The latter has developed quantitative models of contrast

selection based on the principles of maximization of distinctiveness and minimization of effort (but not maximization of the number of contrasts).

The conflicts between these goals can be illustrated by considering the selection of contrasting sounds from a schematic two dimensional auditory space, shown in figure 1. Figure 1a shows an inventory which includes only one contrast, but the contrast is maximally distinct, i.e. the two sounds are well separated in the auditory space. If we try to fit more sounds into the same auditory space, the sounds will necessarily be closer together, i.e. the contrasts will be less distinct (fig. 1b). Thus the goals of maximizing the number of contrasts and maximizing the distinctiveness of contrasts inherently conflict. Minimization of effort also conflicts with maximizing distinctiveness. Assuming that not all sounds are equally easy to produce, attempting to minimize effort reduces the area of the auditory space available for selection of contrasts. For example, if we assume that sounds in the periphery of the space involve greater effort than those in the interior, then, to avoid effortful sounds it is necessary to restrict sounds to a reduced area of the space, thus the contrasts will be less distinct, as illustrated in fig. 1c. Note that while minimization of effort and maximization of the number of contrasts both conflict with maximization of distinctiveness, they do not directly conflict with each other.

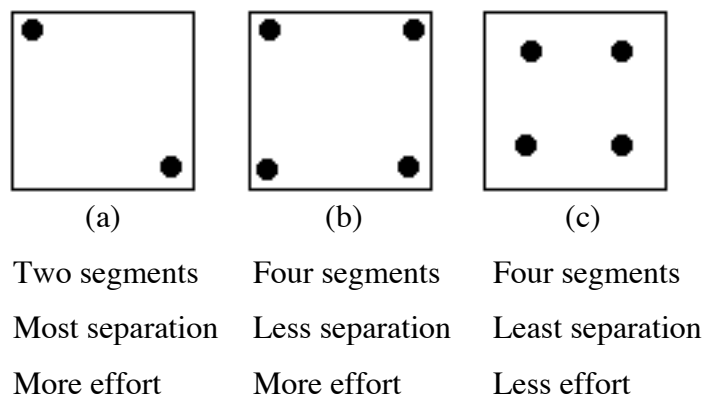


Fig 1. Selection of contrasts from a schematic auditory space.

Given that the three requirements on contrasts conflict, the selection of an inventory of contrasts involves achieving a balance between them. A source of cross-linguistic

variation is variation in the compromise that given languages adopt. The next section presents a preliminary formalization of the dispersion theory in terms of Optimality Theory (Prince and Smolensky 1993). Optimality theory is suitable for this purpose, because it provides a system for specifying the resolution of conflict between constraints.

3.1 Formulation of the constraints on contrast

Optimality Theoretic models achieve optimization without numerical calculation by adhering to a requirement of strict constraint dominance, i.e. where two constraints conflict, the higher-ranked constraint prevails (Prince and Smolensky 1993:78). In the Dispersion Theory, assigning complete dominance to any one of the proposed fundamental constraints yields inappropriate results. For example, if maximization of the number of contrasts dominates, the result will be a huge number of very fine contrasts. The essence of the dispersion theory is that the conflicting goals are balanced against each other.

The balancing of conflicting scalar constraints can be modeled in terms of strict dominance by decomposing the scalar constraints into a ranked set of sub-constraints. This technique is adopted by Prince and Smolensky (1993) in the analysis of syllable structure, where a general constraint requiring a syllable nucleus to be maximally sonorous is decomposed into a set of constraints against particular segments being in the nucleus, with the sub-constraints being ranked according to the sonority of the segments. The sub-constraints corresponding to the scalar constraints can then be interleaved, resulting in a balance between them. This strategy will be followed here.

3.1.1 Maximize the distinctiveness of contrasts

Given the considerations outlined in section 2 above, the measure of distinctiveness which is predicted to be relevant to the markedness of a contrast between two sounds is the probability of confusing the two sounds. Our understanding of the acoustic basis of confusability is limited, so any general model of distinctiveness is necessarily tentative. To allow the precise formulation of analyses, a fairly specific view of distinctiveness will be presented, but many of the details could be modified without affecting the central claims advanced here.

In psychological work on identification and categorization it is common to conceive of stimuli (such as speech sounds) as being located in a multi-dimensional similarity space where the distance between stimuli is systematically related to the confusability of those stimuli – i.e. stimuli which are closer together in the space are more similar, and hence more confusable (e.g. Shepard 1957, Nosofsky 1992). This conception is adopted here. The domain in which we have the best understanding of perceptual space is vowel quality. There is good evidence that the main dimensions of the similarity space for vowels correspond well to the frequencies of the first two formants (Delattre, Liberman, Cooper, and Gerstman 1952, Plomp 1975, Shepard 1972), and less clear evidence for a dimension corresponding to the third formant (see Rosner and Pickering 1994:173ff. for a review).

A coarsely quantized three-dimensional vowel space, adequate for most of the analyses developed here, is shown in (7a-c) (cf. Liljencrants and Lindblom 1972). Sounds are specified by matrices of dimension values, e.g. [F1 1, F2 6, F3 3] for [i]. That is, dimensions are essentially scalar features so standard feature notation is used with the modification that dimensions take integer values rather than +/- . The locations of different vowel qualities are indicated as far as possible using IPA symbols. In some cases there is no IPA symbol for a particular vowel quality (e.g. the unrounded counterpart to [u] which might occupy [F1 2, F2 2]), while in many cases more than one vowel could occupy a given position in F1-F2 space due to the similar acoustic effects of lip rounding and tongue backing. Some examples are shown in (7c). Also, the IPA low back unrounded vowel symbol [ɑ] is used for a wide range of vowel qualities in transcriptions of English dialects and could have been used to symbolize [F1 7, F2 2]. Similarly, [y] could also have been used for [F1 1, F2 5].

(7) a.

						F2						
						6	5	4	3	2	1	
						i	i	y	i	ɯ	u	1
							ɪ	ʏ			ʊ	2
							ɛ	ø		ʎ	ɔ	3
							e	ø	ə	ʏ	o	4 F1
								ɛ	ɜ	ʌ	ɔ	5
								æ	ɐ		ɑ	6
									a	ɑ		7

b.

			F3			
			3	2	1	
			i		y,ɯ,u	1
				ɪ	ʏ,ʊ	2
				ɛ	ø,ʎ,ɔ	3
				e	ø,ʎ,o	4 F1
				ɛ	ʌ,ɔ	5
					ɑ	6
					a	7

c.

						F2						
						6	5	4	3	2	1	
						i				ɯ	u	1
							ɪ				ʊ	2
							ɛ				ɔ	3
							e			ə	o	4 F1
								ɛ	æ		ɔ	5
								æ			ɑ	6
									a			7

The distinctiveness of a pair of vowel qualities should then be calculated from the differences on each of these three dimensions. For example, it might correspond to the euclidean distance between the two vowels. We will return to the issue of overall distinctiveness below, but our most confident judgements involve the relative distinctiveness of differences on individual dimensions. For example [i-u] is more distinct than [i-ɯ] because all the vowels have similar F1, but the first contrast involves a greater difference in F2. For this reason, and for ease of exposition, we will start by considering the restricted case of selecting a set of contrasting sounds along one perceptual dimension, specifically selection of vowels contrasting in F1 ('vowel height').

The requirement that the auditory distinctiveness of contrasts should be maximized can be decomposed into a ranked set of constraints requiring a specified minimal auditory

distance between contrasting forms (8). Since we are initially restricting attention to individual auditory dimensions, it is only necessary to consider constraints in which the specified minimum distance is on a single dimension. The required distance is indicated in the format *Dimension:distance*, e.g. ‘MINDIST = F1:2’ is satisfied by contrasting sounds that differ by at least 2 on the F1 dimension.

(8) MINDIST = F1:1 >> MINDIST = F1:2 >>... >> MINDIST = F1:4

To encode the fact that auditory distinctiveness should be maximized, MINDIST = $D:n$ is ranked above MINDIST = $D:n+1$, i.e. the less distinct the contrast, the greater the violation.

1.1.2 Maximize the Number of Contrasts

The requirement that the number of contrasts should be maximized can be implemented in terms of a positive constraint, MAXIMIZE CONTRASTS that counts the number of contrasts in the candidate inventory. The largest inventory or inventories are selected by this constraint, all others are eliminated. Of course the largest candidate inventories will usually have been eliminated by higher-ranked constraints, so this constraint actually selects the largest viable inventory.

1.1.3 Balancing the requirements on contrasts

The language-specific balance between these first two constraints on contrasts is modeled by specifying the language-specific ranking of the constraint MAXIMIZE CONTRASTS in the hierarchy of MINDIST constraints. Effectively, the first MINDIST constraint to outrank MAXIMIZE CONTRASTS sets a threshold distance, and the optimal inventory is the one that packs the most contrasting sounds onto the relevant dimension without any pair being closer than this threshold.

Continuing the example of the selection of F1 contrasts, the conflict between the two constraints on contrasts is illustrated in the tableau in (18). This tableau shows inventories of contrasting vowel heights and their evaluation by MINDIST and MAXIMIZE contrasts

constraints. We are considering constraints on contrasts so the candidates evaluated here are sets of contrasting forms rather than outputs for a given input.

MINDIST constraints assign one mark for each pair of contrasting sounds which are not separated by at least the specified minimum distance. For example, candidate (b) violates MINDIST = F1:4 twice because the contrasting pairs [i-e] and [e-a] violate this constraint while [i-a] satisfies it, being separated by a distance of 6 on the F1 dimension. (Note that the number of violations will generally be irrelevant for MINDIST constraints ranked above MAXIMIZE CONTRASTS because it will always be possible to satisfy the MINDIST constraint by eliminating contrasts).

MAXIMIZE CONTRASTS is a positive scalar constraint, according to which more contrasts are better, so evaluation by this constraint is indicated using one check mark (✓) for each contrasting sound – more check marks indicate a better candidate according to this constraint. The conflict between the two constraint types is readily apparent in (9): sets of vowel height contrasts which better satisfy MAXIMIZE F1 CONTRASTS incur worse violations of the MINDIST constraints.

(9)

	MINDIST = F1:1	MINDIST = F1:2	MINDIST = F1:3	MINDIST = F1:4	MINDIST = F1:5	MAXIMIZE CONTRAST S
a.	i-a					✓✓
b.	i-e-a			**	**	✓✓✓
c.	i-ɛ-ε-a		***	***	*****	✓✓✓✓
d.	i-I-ɛ-ε-a	**	*****	*****	*****	✓✓✓✓✓

The effect of ranking MAXIMIZE CONTRASTS at different points in the fixed hierarchy of MINDIST constraints is illustrated by the tableaux in (10) and (11). The ranking in (10) yields three distinct vowel heights – i.e. the winning candidate is (b). This candidate violates MINDIST = F1:3, but any attempt to satisfy this constraint by improving distinctiveness, as in candidate (a), violates higher-ranked MAXIMIZE CONTRASTS by selecting only two contrasting vowel heights. It is not possible to fit three contrasting

vowels with a minimum separation of 3 features on the F1 dimension. Candidate (c) better satisfies MAXIMIZE CONTRASTS than (b), maintaining four contrasting vowel heights, but (c) violates higher-ranked MINDIST = F1:2 since [e-ε] and [ε-a] each differ by only 1 on the F1 dimension.

(10)

	MINDIST = F1:2	MINDIST = F1:3	MAXIMIZE CONTRASTS	MINDIST = F1:4	MINDIST = F1:5
a. i-a			✓✓!		
b. i-e-a			✓✓✓	**	**
c. i-ε-e-a		*!***	✓✓✓✓	***	*****

Thus the particular balance achieved here between maximizing the number of contrasts and maximizing the distinctiveness of the contrasts yields three contrasting heights. Altering the ranking of MAXIMIZE CONTRASTS results in a different balance. For example, if less weight is given to maximizing the number of contrasts, ranking MAXIMIZE CONTRASTS below MINDIST = F1:3, the winning candidate has just two contrasting vowel heights, differing maximally in F1. It is apparent that the maximally distinct F1 contrast [i-a] is preferred over any sub-maximal contrast, such as [i-æ] (which violates MINDIST = F1:6), although this comparison is not included in the tableau.

(11)

	MINDIST = F1:2	MINDIST = F1:3	MINDIST = F1:4	MAXIMIZE CONTRASTS	MINDIST = F1:5
a. i-a				✓✓	
b. i-e-a			*!*	✓✓✓	**
c. i-ε-e-a		*!***	***	✓✓✓✓	*****

Not all conceivable rankings of MAXIMIZE CONTRASTS correspond to possible languages. The balance between maximization of the number of contrasts and maximization of the distinctiveness of contrasts is determined by the ranking of

MAXIMIZE CONTRASTS relative to the MINDIST constraints. Allowing all definable rankings predicts the existence of languages which value the number of contrasts very highly, resulting in a huge number of very fine contrasts, and languages which value distinctiveness very highly, resulting in a handful of maximally distinct contrasts. Neither of these extremes is attested. It seems that there is a lower bound on the distinctiveness required for a contrast to be functional, and that there is an upper bound beyond which additional distinctiveness provides a poor return on the effort expended. This could be implemented by specifying that certain MINDIST constraints, referring to the smallest acceptable contrastive differences, are universally ranked above MAXIMIZE CONTRASTS, and that MAXIMIZE CONTRASTS is in turn universally ranked above another set of MINDIST constraints which make ‘excessive’ distinctiveness requirements. However it would be desirable to derive these bounds from general considerations of perceptibility and communicative efficiency rather than simply stipulating them.

Note that the need to place limits on possible constraint rankings is not novel to the dispersion theory. The same issue arises with respect to standard faithfulness constraints: If all faithfulness constraints are at the top of the ranking then all inputs will surface as well-formed outputs, i.e. this ranking would yield an unattested language with no restrictions on the form of words. Conversely, if all faithfulness constraints were at the bottom of the ranking then all inputs would be mapped to a single, maximally well-formed output (presumably the null output, i.e. silence).

1.1.4 Minimization of effort

The analyses above do not include effort minimization constraints. No general account of the effort involved in speech production will be proposed here, instead specific constraints such as ‘Don’t voice obstruents’ and ‘Don’t have short low vowels’ will be motivated as they become relevant. If a sound violates an effort constraint which outranks MAXIMIZE CONTRASTS, it will not be employed even if it would allow more contrasts or more distinct contrasts.

1.2 Some effects of MINDIST constraints

1.2.1 Dispersion

The most basic consequence of the distinctiveness constraints (MINDIST constraints) is a preference for distinct contrasts. This gives rise to dispersion effects whereby contrasting sounds tend to be evenly distributed over as much auditory space as effort constraints will allow (cf. Liljencrants and Lindblom 1972, Lindblom 1986). This effect has already been demonstrated above in relation to F1 contrasts, and the preference for front unrounded and back rounded vowels discussed in section 2 is another instance of this tendency, applied to contrasts on the F2 dimension. The acoustic effects of lip-rounding mean that the maximal F2 difference is between front unrounded vowels and back rounded vowels (12), so if maximization of distinctiveness of F2 contrasts outranks maximizing the number of contrasts, these are the vowels that will be selected (13). F2 contrasts involving central vowels are necessarily sub-maximal, and thus are dispreferred. Of course, the appearance of central vowels non-peripheral vowels may be motivated by the desire to maximize contrasts – i.e. if MAXIMIZE CONTRASTS is ranked above MINDIST = F2:3.

(12)

						F2
6	5	4	3	2	1	
i	i	y	i	u	u	

(13)

	MINDIST = F2:3	MINDIST = F2:4	MAXIMIZE CONTRASTS	MINDIST = F2:5
a. i-u			✓✓	
b. i-u			✓✓	*!
c. y-u			✓✓	*!
d. i-i		*!	✓✓	*
e. i-i-u	*!	**	✓✓✓	**

This notion of dispersion of contrasting sounds is also closely related to the concept of ‘enhancement’, a term coined by Stevens, Keyser, and Kawasaki (1986). Stevens et al observe that ‘basic’ distinctive features are often accompanied by ‘redundant’ features which ‘strengthen the acoustic representation of distinctive features and contribute additional properties which help the listener to perceive the distinction’ (p.426). They regard the relationship between [back] and [round] in vowels as one of enhancement: [round] enhances distinctive [back]. In terms of the dispersion theory, this can be reformulated as the observation that independent articulations often combine to yield more distinct contrasts.

1.1.2 Neutralization

A second basic effect of MINDIST constraints, in interaction with the other dispersion theoretic constraints, is neutralization of indistinct contrasts. In dispersion theory, neutralization of a contrast results when constraints prevent it from achieving sufficient distinctiveness in some environment. That is, in a ranking of the form shown in (14) where *EFFORT is an effort minimization constraint penalizing some articulation, a contrast will be neutralized in some context if it cannot be realized with a distinctiveness of d without violating *EFFORT.

(14) MINDIST = d , *EFFORT >> MAXIMIZE CONTRASTS

The distinctiveness that can be achieved for a given degree of effort varies across contexts. Some cues to contrasts are simply unavailable in certain contexts, for example release formant transitions are not available as a cue to consonant place if the consonant is not released into an approximant. In addition the articulatory effort involved in realizing a cue is generally highly context-dependent, for example voicing of an obstruent is more difficult following a voiceless sound than following a voiced sound because it is more difficult to initiate voicing than to sustain it (Westbury and Keating 1986). So the possibility of realizing a contrast that satisfies MINDIST = d without violating *EFFORT depends on context, and consequently a given type of contrast may be selected as optimal

in some contexts and not in others – i.e. the contrast is neutralized in those other contexts. For example, consonant place contrasts may be permitted before sonorants, but neutralized before obstruents, where stop bursts and release transitions are not available. Thus dispersion theory provides an account of Steriade’s (1995, 1997) generalization that contrasts are neutralized first in environments where ‘the *cues* to the relevant contrast would be diminished or obtainable only at the cost of additional articulatory maneuvers’ (Steriade 1997:1).

It is important to note that the ranking of other constraints will typically be crucial in making the realization of a distinct contrast more effortful in a particular context – e.g. stop bursts will only be absent before obstruents if some constraint requires the stop closure to overlap with the following consonant. In the example we will consider here metrical constraints on unstressed vowel duration make distinct vowel contrasts more difficult to realize in unstressed syllables.

The analysis of neutralization will be exemplified with analyses of two common patterns of vowel reduction: Reduction from a seven vowel inventory (15i) in primary stressed syllables to a five vowel inventory (15ii) in other syllables, as in Central Italian dialects (Maiden 1995), and reduction from a five vowel inventory (15ii) in primary stressed syllables to a three vowel inventory (15iii) elsewhere, as in Southern Italian dialects (Maiden 1995) and Russian (Halle 1959).

(15)	(i)	(ii)	(iii)
	i u	i u	i u
	e o	e o	a
	ε ɔ	a	
	a		

The Central Italian pattern is exemplified in (16) with data from standard Italian (as described in dictionaries). The pairs of words on each line are morphologically related so the parenthesized forms illustrate alternations that arise when stress is shifted off a vowel which cannot appear in an unstressed syllable.

(16)	<u>stressed vowels</u>		<u>unstressed vowels</u>	
	[i]	víno ‘wine’	vinífero	‘wine-producing’
	[e]	péska ‘fishing’	peskáre	‘to fish’
	[ɛ]	bél:o ‘beautiful’	(bel:íno	‘pretty’)
	[a]	máno ‘hand’	manuále	‘manual’
	[ɔ]	mól:e ‘soft’	(mol:eménte	‘softly’)
	[o]	nóme ‘name’	nomináre	‘to name, call’
	[u]	kúra ‘care’	kuráre	‘to treat’

The Southern Italian pattern is exemplified by the dialect of Mistretta, Sicily (Mazzola 1976) (17).

(17)	<u>stressed vowels</u>		<u>unstressed vowels</u>	
	[i]	vín:i ‘he sells’	vin:ímu	‘we sell’
	[e]	véni ‘he comes’	(vin:ímu	‘we come’)
	[a]	ávi ‘he has’	avíti	‘he has’
	[o]	móri ‘he dies’	(mur:ímu	‘we die’)
	[u]	uʝ:i ‘he boils’	uʝ:ímu	‘we boil’

These patterns of reduction involve neutralization of F1 contrasts only. According to the analysis of neutralization outlined above, this implies that it is more difficult to produce distinct F1 contrasts in unstressed positions. The most likely source of that difference in difficulty is the difference in duration between primary stressed and other vowels in these languages. So the proposed analysis is that producing low vowels is increasingly difficult as vowel duration is reduced, and this motivates raising of short low

vowels, leaving a smaller range of the F1 dimension for distinguishing F1 contrasts. This in turn can result in the selection of a smaller number of contrasts².

The most direct evidence for a relationship between vowel duration and the ability to achieve a high F1 comes from Lindblom's (1963) finding that the F1 of Swedish non-high vowels decreases exponentially as vowel duration decreases. It is also well established that low vowels are longer than high vowels, other things being equal (Lehiste 1970). These effects are commonly attributed to the greater articulator movement involved in producing a low vowel between consonants: low vowels require an open upper vocal tract to produce a high F1, whereas all consonants (other than pharyngeals and laryngeals) require upper-vocal tract constrictions, so producing a low vowel between consonants requires substantial opening and closing movements. Westbury and Keating (1980, cited in Keating 1985) provide evidence that vowel duration differences are indeed related to distance moved: they found that vowels with lower jaw positions had longer durations in a study of English. Thus producing a low vowel with the same duration as a higher vowel will typically require faster, and consequently more effortful, movements

Given these considerations, the direct consequence of shorter vowel duration is that it becomes more difficult to produce a high F1, so the same effort required to produce a high F1 in a longer vowel results in a lower F1 in a shorter vowel. Reduction of low /a/ to [ɐ] or [ə] in unstressed syllables is accordingly commonly reported both impressionistically, and in experimental studies such as Lindblom (1963). This correlation between duration and raising of low vowels has been observed in Central Italian also: a study of vowels in the speech of five male Italian television news readers (Albano Leoni et al 1995) found that /a/ in a primary stressed syllable was twice as long as medial unstressed /a/, and the mean F1 of /a/ was 750 Hz in primary stressed syllables, but 553 Hz in medial unstressed syllables (close to the F1 of a stressed lower-mid

² Crosswhite (this volume) suggests that vowel raising is also desirable in unstressed syllables because it lowers the sonority of the vowel, resulting in a better correspondence between stress and vowel intensity.

vowel)³. So the inventory in unstressed positions is more accurately transcribed as [i, e, ɜ, o, u], where [ɜ] is a lower-mid central vowel.

While the direct effect of vowel shortening is on the difficulty of producing low vowels, this has obvious consequences for the selection of F1 contrasts: if the lowest vowel in an inventory is lower-mid ([F1 5] in the terms used here) this leaves less room for distinguishing F1 contrasts than in stressed syllables where the lowest vowel is truly low ([F1 7]), so it is not possible to maintain the same number of height contrasts with the same distinctiveness. Consequently three vowel heights are selected in unstressed syllables, and four in longer, stressed syllables.

This analysis can be formalized in terms of the constraint ranking in (18). Reduction affects height contrasts only, so we need only consider the F1 dimension, repeated in (19).

- (18) UNSTRESSED VOWELS ARE SHORT,
 *SHORT LOW V,
 MINDIST = F1:2 >>
 MAXIMIZE CONTRASTS >>
 MINDIST = F1:3

- (19) F1 dimension

1	i
2	ɪ
3	e̞
4	e, ə
5	ɛ, ɜ
6	ɐ
7	a

³ Word-final unstressed syllables were more variable in duration, probably because duration in this position is dependent on phrase-final lengthening effects. F1 of final /a/ was correspondingly more variable. The greater duration of phrase final vowels does not lead to a larger vowel inventory in this position – this is probably a ‘uniformity’ effect (Steriade 1997, 2000), i.e. it allows words to have a more consistent realization across phrasal positions.

UNSTRESSED VOWELS ARE SHORT is a place-holder for whatever constraints require unstressed vowels to be shorter than stressed vowels. *SHORT LOW V is a constraint against expending the effort to produce a short low vowel. This should properly be derived from a general model of articulatory effort (cf. Kirchner 1998), but for present purposes we will formalize it as a constraint that penalizes short vowels with F1 of greater than 5 on the scale in (19).

In stressed syllables, the first two constraints are irrelevant, so the ranking yields four vowel heights, each separated by F1:2, as shown in (20). However, in unstressed syllables, high-ranking UNSTRESSED VOWELS ARE SHORT requires short vowels, so *SHORT LOW V is applicable also. This effort minimization constraint penalizes low vowels, so the candidate [i-ɛ-ε-a] is now ruled out because [a] has [F2 7] (21a). Attempting to maintain four contrasts while avoiding low vowels, as in candidate (b), results in violations of MINDIST = F1:2 because [ε-ɜ] don't differ in F1. The winning candidate has three vowel heights, and so is evaluated as worse by MAXIMIZE CONTRASTS, but satisfies the higher-ranked minimum distance requirement.

(20) Central Italian – Vowels in primary stressed syllables.

	*SHORT LOW V	MINDIST = F1:2	MAXIMIZE CONTRASTS	MINDIST = F1:3
a.	í-á		✓✓!	
b.	í-é-á		✓✓✓!	
c.	☞ í-ɛ-é-á		✓✓✓✓	***

(21) Central Italian – Vowels in unstressed syllables.

	*SHORT LOW V	MINDIST = F1:2	MAXIMIZE CONTRASTS	MINDIST = F1:3
a.	i-ɛ-ε-a	*!	✓✓✓✓	***
b.	i-ɛ-ε-ɜ		✓✓✓✓	***
c.	☞ i-ɛ-ɜ		✓✓✓	**

A similar ranking (22) derives the Southern Italian pattern in which three vowel heights are contrasted in primary stressed syllables, but only two elsewhere. The only difference is that both MINDIST constraints are ranked above MAXIMIZE CONTRASTS – i.e. distinctiveness requirements are more demanding.

(22) UNSTRESSED VOWELS ARE SHORT,

*SHORT LOW V,

MINDIST = F1:2 >>

MINDIST = F1:3 >>

MAXIMIZE F1 CONTRASTS

This is the same ranking of MINDIST constraints used to derive three contrasting vowel heights in (10) above, and the same derivation applies in primary stressed syllables, where the top-ranked constraints are irrelevant. In unstressed syllables, *SHORT LOW V is applicable again, so the lowest vowel possible is [ɜ], and it is not possible to fit a vowel between [i] and [ɜ] while satisfying MINDIST = F1:3 (23b), so only two vowel heights are selected (23c).

(23) Southern Italian – Vowels in unstressed syllables.

	*SHORT LOW V	MINDIST = F1:2	MINDIST = F1:3	MAXIMIZE CONTRASTS
a.	i-e-a	*!		✓✓✓
b.	i-ɛ-3		*!*	✓✓✓
c.	☞ i-3			✓✓

This analysis is based on the assumption that the difference between the two patterns of reduction lies in the ranking of MINDIST constraints, but there may also be differences in the characteristic durations of the unstressed vowels. It has already been noted that the difficulty of producing a low vowel increases as vowel duration decreases, so if Southern Italian unstressed vowels are shorter than Central Italian unstressed vowels, then more raising of low vowels may occur, making reduction to a two-height system more desirable. Good evidence that different degrees of shortening can result in different degrees of reduction in this way is provided by Brazilian Portuguese. Brazilian Portuguese combines the two patterns of reduction: the seven vowel system (15i) is permitted in primary stressed syllables, the five vowel system (15ii) in syllables preceding the stress, and the three vowel system (15iii) in unstressed final syllables (stress is generally penultimate) (Mattoso Camara 1972). Since both patterns of reduction occur in the same language, they cannot be accounted for in terms of differences in the ranking of MINDIST constraints. However, Major (1992) found that final unstressed syllables are substantially shorter than pre-stress syllables (which are in turn shorter than stressed syllables)⁴, so the same degree of effort should result in higher vowels in this

⁴ Moraes (1998) found relatively small differences in duration between pre-tonic and final unstressed vowels, but he only measured high vowels which tend to be short in any case. Major measured low vowels, which are more relevant here. Moraes (1998) also shows that the duration difference can be eliminated by phrase-final lengthening of final unstressed syllables. As in Italian, it appears to the phrase-medial characteristics that are relevant to neutralizing vowel reduction. It is also interesting to note that Moraes found that final unstressed vowels have much lower intensity than vowels in other positions, and that this remains true even with final-lengthening (this should not be a consequence of vowel raising, since all vowels were high). Intensity should play some role in the perceptibility of vowel contrasts, but this factor is not analyzed here.

position. The ‘low’ vowel is indeed higher in this position, as indicated by the standard impressionistic transcription of the final unstressed vowel system as [i, ə, u] (e.g. Mattoso Camara 1972). Acoustic data reported by Fails and Clegg (1992) shows a progressive decrease in F1 of the lowest vowel from primary stressed, to pre-stress, to final unstressed.

So the two degrees of vowel reduction observed in Brazilian Portuguese can be attributed to the difference in vowel duration in the final vs. pre-stress position. This analysis makes it necessary to distinguish two degrees of vowel shortening, which we will refer to as ‘short’ and ‘extra short’. It is also necessary to recognize two forms of gradience in the relationship between effort and vowel duration. Given that faster movements involve greater effort, then the effort involved in producing a low vowel increases as vowel duration decreases:

(24) *EXTRA-SHORT LOW V >> *SHORT LOW V

By the same reasoning, the lower a vowel is, the more effortful it is, for a given vowel duration (25). The EXTRA-SHORT counterpart of each of these constraints is universally ranked higher, although the way in which the two hierarchies should interleave depends upon the precise interpretation of ‘short’ and ‘extra-short’. For the analysis of Brazilian Portuguese, the crucial rankings are as in (26), i.e. short [ɐ] involves comparable effort to extra-short [ɜ].

(25) *SHORT a >> *SHORT ɐ >> *SHORT ɜ >> *SHORT ə >> ... >> *SHORT i

(26) *SHORT ɐ, *EXTRA-SHORT ɜ >> *SHORT ɜ

The full constraint ranking is in (27). Again, only the relationship between vowel duration and reduction is analyzed here, so the first two constraints await replacement by a proper analysis of the relationship between vowel position and duration.

- (27) UNTRESSED VOWELS ARE SHORT,
 FINAL UNSTRESSED VOWELS ARE EXTRA-SHORT,
 *SHORT ɐ ,
 *EXTRA-SHORT ɜ ,
 MINDIST = F1:2 >>
 MAXIMIZE CONTRASTS >>
 *SHORT ɜ ,
 MINDIST = F1:3

In primary stressed syllables, only the MINDIST constraints and MAXIMIZE CONTRASTS are relevant, and these are ranked exactly as in Central Italian (18, above), so they yield four degrees of vowel height, as in (20, above). In pre-stress position, vowels are required to be short, so the effort minimization constraints become relevant, resulting in selection of three vowel heights, as in Central Italian (28). In final unstressed syllables, vowels must be extra-short (transcribed with a breve: [ǎ]), so high-ranked *EXTRA-SHORT ɜ makes the lowest possible vowel a mid [ə] (29a-c). Consequently, only two vowel heights are possible without violating MINDIST = F1:2 (29d).

- (28) Brazilian Portuguese – Vowels in pre-stress syllables.

	*SHORT ɐ	*EXTRA-SHORT ɜ	MINDIST = F1:2	MAXIMIZE CONTRASTS	*SHORT ɜ	MINDIST = F1:3
a. i- ɐ - ɛ -a	*!			✓✓✓✓	*	***
b. i- ɐ - ɛ - ɜ			*!	✓✓✓✓	*	***
c. ɨ i- ɐ - ɜ				✓✓✓	*	**
d. i- ɜ				✓✓!	*	

(29) Brazilian Portuguese – Vowels in final unstressed syllables.

	*SHORT e	*EXTRA-SHORT ɜ	MINDIST = F1:2	MAXIMIZE CONTRASTS	*SHORT ɜ	MINDIST = F1:3
a.	ĩ-ẽ-ě-ă	**		✓✓✓✓	*	***
b.	ĩ-ẽ-ě-ɜ	*!	*	✓✓✓✓	*	***
c.	ĩ-ẽ-ɜ	*!		✓✓✓	*	**
d.	ĩ-ẽ-ă		*!	✓✓✓	*	**
e.	ĩ-ă			✓✓		

These analyses illustrate the basic approach to neutralization within dispersion theory – i.e. the effort involved in making a contrast distinct can vary across contexts, and neutralization becomes optimal in contexts where adequate distinctiveness cannot be achieved with an acceptable degree of effort. In these cases, positional constraints on vowel duration result in variation in the difficulty of producing low vowels, which has consequences for the distinctiveness of height contrasts. It is also of interest that ‘phonetic’ reduction is central to this analysis of neutralizing vowel reduction. It is usual to abstract away from partial raising of low vowels in unstressed syllables, e.g. to [e] or [ɜ], in analyses of neutralizing reduction, so the low vowel is treated as having the same feature representation in all contexts. In the analysis proposed here, neutralization of height contrasts is results from the difficulty of maintaining distinctiveness in the face of partial raising of low vowels.

We will see that duration-based neutralization is also central to some of the case studies that provide more direct evidence for distinctiveness constraints (4.1), but before turning to those cases, we will consider some additional issues in the formulation and application of dispersion theory. The next section briefly addresses the analysis of complete words and morphophonological alternations, while the following section will discuss the generalization of perceptual distinctiveness to contrasts that differ in more than one perceptual dimension.

1.1.3 Analysis of words and alternations

The analysis of vowel reduction raises two important issues concerning analyses using dispersion theory. First, we have only considered the selection of inventories of contrasting sounds but a phonology must characterize the set of well-formed possible words in a language. The implication of dispersion theory is that words must be evaluated with respect to paradigmatic constraints in addition to the familiar syntagmatic markedness constraints, such as effort minimization and metrical constraints. That is, words must be sufficiently distinct from other minimally contrasting possible words (MINDIST), and there must be a sufficient number of such contrasting words (MAXIMIZE CONTRASTS). Deriving inventories of sounds in particular contexts is an important step towards the analysis of complete words because for a word to be well-formed, each sound in that word must be a member of the optimal inventory for its particular context. We will see in section 5 that developing this basic idea fully raises some technical issues, but we will postpone that discussion until we have more thoroughly motivated the constraints on contrast.

The second issue raised by the analysis of vowel reduction is how morpheme alternations should be analyzed. The analysis in 3.2.2 derives the distributional fact that a smaller inventory of vowels is permitted in short, unstressed syllables, but says nothing about the alternations that occur when stress falls on different syllables in a morpheme. For example, in Sicilian Italian, the vowel of the stem ‘come’ alternates between mid when stressed, as in [véni] ‘he comes’, and high when unstressed, as in [vinímu] ‘we come’. The analysis proposed here correctly allows the attested forms, and rules out unreduced *[venímu], but says nothing about the alternations observed in the realization of particular morphemes. In standard OT, the analysis of alternations centers on faithfulness to the underlying representation of a morpheme. The analyses considered here have no input underlying representation or faithfulness constraints, but we will see that it is not possible to analyze alternations by simply adding faithfulness constraints to the analysis of vowel reduction, rather, a fundamental revision of the analysis of alternations is required which probably makes inputs unnecessary.

Standard faithfulness constraints essentially require that the output be as similar as possible to the input – i.e. segments should not be inserted or deleted, feature values should not be changed, etc. These constraints play two conceptually distinct roles: They play a central role in determining the distribution of contrasts in a language, and they also serve to regulate alternations.

The role of faithfulness constraints in determining contrasts is discussed by Prince and Smolensky (1993, ch.9) and Kirchner (1997), and can be illustrated as follows. Essentially, a constraint IDENT F, where F is a feature, favors preserving underlying differences - i.e. if the input contains [+F], the output should contain [+F], if the input contains [-F], the output should contain [-F]. So, if IDENT F is satisfied, an underlying difference between [+F] and [-F] is preserved on the surface and we have a contrast in F. In the dispersion theory, this function of faithfulness constraints is taken over by MAXIMIZE CONTRASTS, so faithfulness constraints are not needed to play this role.

However, MAXIMIZE CONTRASTS cannot simply replace faithfulness constraints because faithfulness constraints also regulate alternations, requiring allomorphs of a given morpheme to be similar. The standard analysis of similarities between allomorphs involves proposing a unique underlying form of the morpheme from which all surface allomorphs are derived, as exemplified in (30). A second component of the analysis must be some requirement that outputs be similar to inputs, otherwise an output need bear no resemblance to the input, and derivation from a common underlying form would in no way guarantee allomorphic similarity. Faithfulness fulfills this function.

(30)	/ætəm/ ‘atom’	/-z/ ‘(pl.)’
	↙ ↘	↙ ↓ ↘
	[æɾəm] [ətʰám-ɪk]	[-s] [-z] [-ɪz]

It is also not possible to combine dispersion constraints with the faithfulness-based account of allomorphic similarity because the two are fundamentally incompatible. This is illustrated in (31) which repeats the ranking used in (10) to derive three contrasting vowel heights [i-e-a], with the addition of a top-ranked faithfulness constraint IDENT

[F1], which requires that output segments have the same [F1] value as the corresponding input segment – i.e. input values of [F1] must be preserved in the output. The problem arises where the input contains a vowel which is not part of the inventory derived by the dispersion constraints. The tableau in (31) shows such a case – the input contains the vowel [ɪ]. Although the input is a complete word, we only consider the evaluation of F1 (vowel height) – i.e. the contrast constraints are evaluated relative to a set of forms differing only in vowel height. The underlined form is the selected output, whereas the other forms are the set of contrasting forms required for the evaluation of constraints on contrast.

The inclusion of faithfulness constraints subverts the intended effect of the MINDIST and MAXIMIZE CONTRASTS constraints, because it makes the selected inventory of vowel height contrasts dependent on the input under consideration – the same constraints that are supposed to derive three vowel heights, as in (10), yield two [ɪ-a] in (31) because faithfulness to the input F1 forces inclusion of [ɪ] in an output form.

(31)

	/pɪt/	IDENT [F1]	MINDIST = F1:3	MAXIMIZE CONTRASTS	MINDIST = F1:4
a.	<u>pit</u> -pet-pat	*!		✓✓✓	**
b.	<u>pɪt</u> -pet-pat		*!	✓✓✓	**
c.	☞ <u>pɪt</u> -pat			✓✓	
d.	<u>pit</u> -pat	*!		✓✓	

It is essential to the proper operation of the constraints on contrast that they evaluate the contrasting sounds that can be used to distinguish words – i.e. they generalize across possible words, and cannot be evaluated relative to a single word as in the standard OT model.

So given the dispersion theory of contrast, we need an account of allomorphic similarity which is independent of the theory of contrast, rather than inextricably combined with it as with faithfulness to an input. In Flemming (1995) it is proposed that

allomorphy should be analyzed in terms of a direct requirement of similarity between the surface forms of a morpheme, i.e. ‘output-output correspondence’ or ‘paradigm uniformity’ constraints. These constraints have been used to account for cyclicity and related effects (e.g. Benua 1997, Burzio 1998, Kenstowicz 1996, Steriade 1997, 2000), but, as Burzio (1998) observes, they can naturally be extended to account for all similarity relations between realizations of a morpheme including those observed in allomorphy, eliminating any role for an input (cf. also Hayes 1999). However most of the cases considered here concern distribution rather than alternations, so we will not pursue this line further here.

1.3 Further issues in perceptual distinctiveness

So far we have only considered distinctiveness on individual dimensions relevant to vowel quality. For the most part, additional dimensions will be introduced as they become relevant, but we will consider here some of the general issues involved in extending the model to the analysis of consonant contrasts and to contrasts which are realized on more than one dimension.

1.3.1 The perceptual space of consonants

Consonant contrasts are more problematic than vowel quality contrasts for two reasons: the relevant perceptual space is probably of much higher dimensionality than the vowel quality space, and cues to consonantal contrasts are often temporally distributed. For example place contrasts among intervocalic stops are cued by preceding and following formant transitions, and the stop burst, among other cues (see Wright this volume for a review). To represent contrasts of this kind, we will adopt the standard assumption that time is divided into discrete segments, but will make use of smaller segments than are usually assumed, mainly to accommodate the representation of stop bursts and formant transitions. A CV sequence is minimally represented by three segments: C constriction, release transition, and vowel. Each segment is represented as a matrix of values on the various dimensions (32).

(32)

<i>closure</i>	<i>release transition</i>	<i>vowel</i>
[intensity 0]	[intensity 7] [F2 2]	[intensity 7] [F2 6]

The formant frequencies following consonant release are specified in the release transitions segment. Similarly, a VC sequence consists of vowel, closure transitions, and consonant constriction. A sequence of a stop followed by an approximant also contains a burst segment immediately following closure, corresponding to the brief period of frication that arises as the articulators move apart after release of the stop closure.

1.3.2 Contrasts on multiple dimensions

In general perception and production studies give us some grasp of the dimensions of perceptual space, and as already noted, it is often possible to make judgements of relative distinctiveness on a dimension with some confidence, but it is much more difficult to assess the relative distinctiveness of contrasts on different dimensions or, more generally, to determine the overall distinctiveness from differences on multiple dimensions. For this reason we will generally avoid analyses which depend on these aspects of distinctiveness. But it is worth considering how multiple dimensions can be considered in an analysis in order to obtain a more general picture of the workings of dispersion constraints, so we will present a simple example of selection of a vowel inventory in two dimensions, F1 and F2.

Selection of a vowel inventory is analyzed, as above, as the selection of vowels from the space in (33) so as to best satisfy MINDIST constraints and MAXIMIZE CONTRASTS, but it is now necessary to address how differences on two dimensions, F1 and F2, combine to yield a total perceptual distance between two sounds.

(33)

		F2					
6	5	4	3	2	1		
i	i	y	i	ɯ	u	1	
	ɪ	ʏ			ʊ	2	
	e̞	ø̞		ɤ	ɔ̞	3	
	e	ø	ə	ɤ	o	4 F1	
		ɛ	ɜ	ʌ	ɔ	5	
		æ	ɐ		ɑ	6	
			a	ɑ		7	

One simple approach is to adopt the common assumption that the perceptual space for a property like vowel quality is euclidean (Shepard 1972, 1991), so the distance d between vowel a with formants $F1_a$ and $F2_a$, and vowel b with formants $F1_b$ and $F2_b$, is given by the familiar formula:

$$(34) \quad d = \sqrt{(F1_a - F1_b)^2 + (F2_a - F2_b)^2}$$

MINDIST constraints can then refer to total perceptual distance rather than distance on a particular dimension. For example, MINDIST =4 is satisfied by [ɔ-u] which differ by 4 in F1, [y-u] which differ by 4 in F2, or [u-ø] which differ by 3 in F1 and 3 in F2, yielding a total distance of 4.2.

A sample analysis is shown in (35). For readability, evaluations by MAXIMIZE CONTRASTS are given as numbers rather than long lists of check marks. The winning candidate is the canonical five vowel system (b), essentially because it is the largest inventory that can be packed into the vowel space while keeping every pair of vowels separated by a distance of at least 3 – i.e. all contrasts satisfy the MINDIST constraint ranked immediately above MAXIMIZE CONTRASTS. Candidates (c)-(e) have more contrasts, but consequently violate high-ranked MINDIST constraints. Candidate (f) contains 5 contrasting vowels, like the winning candidate, but they are less well dispersed than they could be because [i-ɯ] and [e-ɯ] are less distinct than [i-u] and [e-u].

(35)		MINDIST = 2	MINDIST = 3	MAXIMIZE CONTRASTS	MINDIST = 4	MINDIST = 5
a.	i u a			3!		
b.	i u e o a			5	***	*****
c.	i u e o ε ɔ a		*!*****	7	*****	*****
d.	i u e o a ̤	*!	*	6	*****	*****
e.	i i u e o a		*!	6	*****	*****
f.	i ʉ e o a			5	***	*****!*

This approach to overall perceptual distance is probably too constrained. If the perceptual distinctiveness of contrasts is completely fixed across languages, as implied above, then it is hard to explain the observed range of cross-linguistic variation in vowel systems. For example, why do inventories larger than five vowels sometimes add F1 contrasts, while others add front rounded vowels? It is possible that other constraints might explain the preference for one vowel system over another. For example, it might be argued that some of these vowel qualities are more effortful than others, so variation in the relative ranking of effort minimization and distinctiveness constraints results in variation in vowel inventories. Such explanations should not be ruled out, but it should be borne in mind that the problem is multiplied when we turn to consonant inventories, where cross-linguistic variation is even greater.

Thus it seems likely that perceptual distinctiveness can vary to some extent cross-linguistically (cf. Lindblom 1986). This could result from differences in the distribution

of attention between dimensions (or portions of dimensions) of an otherwise universal perceptual space. Such differences in the allocation of attention have been hypothesized to result from perceptual learning in general (e.g. Nosofsky 1986, Goldstone 1994) and in acquisition of speech perception in particular (e.g. Jusczyk 1993, Guenther, Husain, Cohen, & Shinn-Cunningham 1999), and are often modeled as stretching and shrinking the dimensions of a perceptual space (e.g. Goldstone, Steyvers, Spencer-Smith, & Kersten 2000; Guenther et al 1999; Livingston, Andrews, and Harnad 1998; Nosofsky 1992).

The extent of this variation must be restricted otherwise we would expect even less cross-linguistic consistency (Lindblom 1986), but we can hypothesize that differences in allocation of attention to F3 vs. F1 accompanies differences between including rounding contrasts between front vowels vs. a larger set of height contrasts, for example.

In any case, the example analysis illustrates two basic principles that should govern MINDIST constraints if they reflect distance in a metric perceptual space: (i) A contrast can be made more distinct by increasing differences between sounds along any dimension, so $F1:2+F2:2 > F1:1+F2:2$, regardless of the relative weights of these dimensions (where '+' is used to represent the combination of differences on distinct dimensions). (ii) Equal distances on different dimensions must be joined disjunctively in MINDIST constraints. E.g. if $F1:3 = F2:3$, as above, then there must be a MINDIST constraint which is satisfied by either difference. This constraint can be written 'MINDIST = $F1:3$ OR $F2:3$ ', in the absence of a general formula for overall perceptual distance.

4 Evidence for constraints on contrasts

Now we have laid out the basics of a theory which incorporates constraints on the distinctiveness of contrasts, the theory will be applied in the analysis of phenomena which illustrate the effects of these constraints, and which are problematic for theories which do not include constraints on contrasts. In general terms, the case studies provide evidence for the central prediction that the markedness of a sound depends on the sounds that it contrasts with. Without constraints on contrasts, markedness is predicted to be purely a property of sounds (or phonological structures), so the markedness of a sound should be independent of the nature of the sounds that it contrasts with.

4.1 F2 contrasts and vowel dispersion

The first case study concerns the preference for front unrounded and back rounded vowels discussed in section 2. This pattern has already been analyzed as a result of the preference for maximally distinct contrasts, i.e. the MINDIST constraints (section 3.2.1): the maximal F2 difference is between front unrounded vowels and back rounded vowels, so if maximization of distinctiveness of F2 contrasts outranks maximizing the number of contrasts, these are the vowels that will be selected. F2 contrasts involving non-peripheral vowels (central vowels, front rounded vowels, etc) are necessarily sub-maximal, and thus are dispreferred.

The novel prediction made by this analysis is that front unrounded and back rounded vowels should only be preferred where there are F2 contrasts. If there are no vowel contrasts that are primarily realized in terms of F2 differences, other constraints are predicted to govern backness and rounding of vowels, the most general of which are effort minimization constraints. It is unusual for all vowel F2 contrasts to be neutralized, but there are two circumstances in which this happens: in ‘vertical’ vowel inventories, and in fully neutralizing vowel reduction in unstressed syllables, as in English reduction to ‘schwa’. In both cases the predictions of the dispersion theoretic analysis are confirmed: we do not find front unrounded or back rounded vowels in most contexts, rather backness and rounding are governed by minimization of effort. This means that they are realized as smooth transitions between preceding and following consonants, which frequently yields central or centralized vowel qualities.

We will first consider vertical vowel inventories, then total vowel reduction, providing evidence that they do pattern as predicted, and clarifying why F2 contrasts are neutralized in these instances.

4.1.1 Vertical vowel inventories

Vowel inventories which lack front-back contrasts are found in Marshallese (Bender 1968, Choi 1992), Northwest Caucasian languages (Colarusso 1988) including Kabardian (Kuipers 1960, Colarusso 1992) and Shapsug (Smeets 1984), and some Ndu languages of Papua New Guinea including Iatmul (Laycock 1965, Staalsen 1966). These languages are typically described as having only central vowels, however this is a claim about the

underlying vowel inventory posited as part of a derivational analysis, not an observation about the surface vowels. On the surface, all of these languages distinguish short vowels from longer vowels, with conventional F2 contrasts among the longer vowels, but no F2 contrasts among the extra-short vowels. In the best documented cases, it is clear that the short vowels are not consistently central, rather they consist of a transition from the articulation of the preceding consonant to the articulation of the following consonant, deviating only to realize F1 (height) distinctions. Consequently these vowels are highly variable in quality, and frequently perceived as diphthongal.

For example, the Northwest Caucasian languages Kabardian and Shapsug have a standard system of five normal length vowels [i, e, a, o, u] (Kuipers 1960:23f., Smeets 1984:123), and a ‘vertical’ system of two short vowels, which can be transcribed broadly as [i, ə]⁵. However all sources indicate that the precise backness and rounding of these vowels depends of their consonantal context. Colarusso (1988) states that in NW Caucasian languages, ‘The sequence C₁iC₂ means “go from 1 to 2, letting your tongue follow the shortest path that permits an interval of sonorant voicing.” C₁əC₂ means “go from 1 to 2... but at the same time imposing on this trajectory an articulatory gesture which pulls the tongue body down and back.”’⁶ (p.307). Similarly, Smeets (1984) says of Shapsug that ‘the timbre of the vowels is determined by the character of the adjacent consonants’ (p.120) and that ‘vowels between consonants that color in a different way have swift-moving polyphthongal variants’.

Marshallese has a long vowel inventory with F2 contrasts between front, central (or back) unrounded, and back rounded vowels. Medial short vowels contrast in height only, and high, mid, and low vowels /i, ə, a/ are distinguished (Bender 1968). Again the backness and rounding of these vowels is dependent on consonant context: Choi (1992) shows that the F2 trajectory of these vowels is a nearly linear interpolation between the F2 values for the preceding and following consonants.

⁵ Kuipers actually transcribes the Kabardian high vowel as [ə], the mid-vowel as [a], and the ‘long’ low vowel as [ā], and Colarusso (1988) follows him in this, but their descriptions, Colarusso’s phonetic transcriptions, and acoustic data in Choi (1991) all indicate that the vowels are actually high and mid respectively.

⁶ The transcription of vowels has been altered in accordance with conventions adopted here.

So, while it is not accurate to say that these languages lack F2 contrasts, it is plausible that F2 is not important in contrasts involving the short vowels because they do not contrast with each other in F2, and it seems likely that they are primarily distinguished from full vowels by duration rather than quality.

The transitional quality of the short vowels is plausibly analyzed as the least effort articulation for a vowel. Although articulatory effort is not well understood, basic considerations imply that higher velocity movements should be more effortful (Kirchner 1998, Nelson 1983, Perkell 1997), and velocity of movement in a vowel is minimized by adopting a linear trajectory between preceding and following consonants. The absolute least effort trajectory between consonants will typically not yield a vowel because all supralaryngeal consonants involve narrower constrictions than vowels. In addition there are height distinctions among the short vowels, so some movement in terms of vowel height is necessary, but backness and rounding can be interpolated between preceding and following consonants, producing the near linear F2 movements observed by Choi.

So these vertical vowel systems are what we expect given the analysis of F2 dispersion above – i.e. where F2 contrasts are neutralized backness and rounding of vowels should be determined by effort minimization, not maximization of distinctiveness. As a result the vowel qualities that surface in the absence of contrast are often central, back unrounded, front unrounded, or short diphthongs involving these qualities, all vowel types which would be highly marked in the presence of F2 contrasts. That is, the markedness of vowel qualities depends on the contrasts that they enter into.

This conclusion holds even more clearly if we follow Choi (1992) in analyzing these vowels as being phonetically underspecified for [back] and [round] – i.e. these vowels lack specifications for these features in the output of the phonology, and the specific contextual allophones are generated through a process of phonetic interpolation. Such unspecified vowels only occur in the absence of F2 contrasts, so they are not just marked in the presence of F2 contrasts, they are unattested.

If there are no constraints on contrast, then this type of pattern should not arise. I.e., if markedness were a property of sounds but not of contrasts, then the same vowel qualities should be unmarked regardless of contrastive status. This point is demonstrated in more detail below, but first it is desirable to address the question why vertical vowel

inventories exist at all. That is, we have seen that the dispersion theoretic analysis correctly predicts the properties of vowel inventories without F2 contrasts, but we have not yet explained why a language would forgo F2 contrasts in the first place. This question is, in a sense, secondary to the outcome of neutralization, but it might be thought that vertical vowel inventories contradict MAXIMIZE CONTRASTS by failing to exploit F2 contrasts, so it is necessary to show that this is not the case.

In outline, the analysis proposed here is that vertical vowel inventories arise from a standard pattern of neutralization (3.2.2) – i.e. F2 contrasts are neutralized in contexts where it is too difficult to realize them distinctly. Two factors contribute to creating such a context: first, very short vowel duration makes realization of F2 contrasts much more difficult because much faster movements are required to reach distinct vowel qualities. Lindblom (1963) shows that F2 at the mid-point of a vowel in a CVC where both consonants are the same tends to move closer to an F2 value characteristic of the consonant as the duration of the vowel is reduced. That is, as a vowel becomes shorter, it becomes more effortful to deviate from the least effort transition between preceding and following consonants by a significant amount, but deviation from the least effort transition is required to realize distinct F2 values for contrasting vowels. At short durations, the effort of realizing a distinct F2 contrast can be sufficient to make neutralization optimal. This difficulty is exacerbated by rich inventories of consonant place contrasts. F2 transitions play an important role in realizing these contrasts, so it is less possible to facilitate vowel contrasts by coproducing vowels with consonants.

The relevance of consonant contrasts is most apparent in Marshallese which has an extensive system of palatalization, velarization, and labio-velarization contrasts (e.g. [p^j-p^y], [k-k^w]). Sequences such as [p^jup^j] and [p^yip^y] obviously require substantial tongue body movement. The Caucasian languages contrast large sets of places of articulation, together with some secondary articulations. So to some extent it appears that vertical vowel inventories are trading vowel F2 contrasts for consonant-centered F2 contrasts. Indeed, analysts have varied between analyzing Arrernte as a vertical vowel language with extensive labio-velarization contrasts, or as a language with vowel F2 contrasts, and a smaller consonant inventory (Ladefoged and Maddieson 1996:357).

However there is good evidence that consonant contrasts alone do not give rise to neutralization of F2 contrasts, and that vowel duration plays an important role. The most direct evidence is the fact that NW Caucasian and Marshallese have F2 contrasts among longer vowels, and the vowels that lack F2 contrasts are very short. As noted above, Kabardian and Shapsug have five normal length vowels [i, e, a, o, u]. These are sometimes overlooked because the front and back vowels have been analyzed as being derived from underlying */ij, əj, əw, iw/* in all the sources cited above. The mid vowels can be slightly diphthongal, particularly in word-final position, but the diphthong movement is from mid to high (e.g. [ei, ou]), not from central to front or back – i.e. no part of the vowel is central. They can also be realized as simple monophthongs, particularly medially (Kuipers 1960:23, Smeets 1984:123). So the vowel+glide analysis is motivated primarily by a desire not to posit additional phonemes – the central vowels and glides are independently required so these authors preferred to analyze long vowels in terms of these phonemes rather than posit additional vowel phonemes. In derivational terms, it is also justified by the fact that morphological concatenation of a non-low central vowel plus a glide yields a front or back vowel. However, the fact remains that there are surface contrasts between normal front and back vowels on the surface, but there are no contrasts between short front and back vowels.

Kuipers (1960) refers to the Kabardian full vowels as ‘long’, however it is clear that this length distinction is very different from vowel length contrasts in languages like Italian or Finnish. Kuipers compares the ‘long’ vowels to Russian stressed vowels in duration, and states that the short high vowel is ‘ultrashort’ (p.24). Choi (1991) reports mean durations of 58 ms for [i] and 84 ms for [ə] in read speech, while ‘long’ vowels averaged about 140 ms. By comparison, English short (lax) vowels can exceed 140ms when accented (e.g. Erickson 2000, Peterson and Lehiste 1960). The vowels of other NW Caucasian languages are probably similar to Kabardian – e.g. Smeets describes unstressed medial [i] in Shapsug as ‘very short’ and easily deleted (p.122), and Colarusso (1988) describes inter-consonantal [i] as being commonly reduced to just a ‘full release of the first consonant’ in all NW Caucasian languages (p.349).

The Marshallese vertical vowels are also short: Choi (1992) reports mean durations of 68 ms, 82 ms, and 100 ms for high, mid, and low vowels in sentence final, accented position (the environment in which greatest duration would be expected). The long vowels averaged over 300 ms in the same environment (which suggests these are long vowels in the usual sense). Marshallese long vowels are analyzed by Bender (1968) as underlying vowel-glide-vowel sequences, which necessitates positing an abstract glide to account for long back unrounded vowels. Again this obscures the nature of the surface contrasts which are between phonetically monophthongal vowels, with no evidence of a medial glide (Choi 1992:71). It is also problematic on phonological grounds since evidence from stress and reduplication indicates that long vowels form single heavy syllables, not syllable sequences (Byrd 1992).

So vertical vowel inventories are the result of positional neutralization of F2 contrasts in extra-short vowels where the effort of realizing distinct F2 contrasts is high. Given that vowel F2 contrasts are neutralized, the realization of vowels defaults to the least-effort realization, i.e. a smooth transition between the positions for the preceding and following consonants, deviating only to realize F1 distinctions.

The constraint ranking in (36) is a partial formalization of this analysis. The constraint *HIGHEFFORT is intended to penalize particularly rapid movements – specifically, with very short vowels, it rules out anything more than small deviations from a smooth transition between tongue body and lip positions for preceding and following consonants. A general formulation of this constraint would take us too far from our primary concerns here, but see Kirchner (1998) for discussion of some of the issues involved. The MINDIST constraint imposes a substantial minimum distance for vowel contrasts in F2, and for contrasts based primarily on F2 during consonant release transitions⁷. This constraint is satisfied by contrasts between fully front and back vowels (e.g. i-u, e-o) or between palatalized and velarized consonants (e.g. p^j-p^v) (see sample F2 specifications in 37).

⁷ This is not intended to imply that differences in F2 transitions and in vowel F2 are always equivalent. It is likely that differences in the end-point of a formant movement are often less distinct than differences realized over a greater duration during a vowel.

(36) *HIGHEFFORT, MINDIST = F2:4 >> MAXIMIZE CONTRASTS

(37)

F2					
6	5	4	3	2	1
i	ị	i̥	i̜	u̜	u
	e	e̜	ə	ɤ	o
				p̣ ^y	
				p ^j	

The operation of these constraints is illustrated by the tableau in (38) which shows the selection of an inventory of CVCs with extra-short vowels, considering only secondary articulation contrasts as representatives of consonant contrasts and only F2 contrasts among vowels.

(38)

		*HIGH EFFORT	MINDIST = F2:4	MAXIMIZE CONTRASTS
a.	<p>p̣^jip̣^j p̣^jụ̈p̣^j p̣^jip̣^y p̣^jụ̈p̣^y</p> <p>p̣^yip̣^j p̣^yụ̈p̣^j p̣^yip̣^y p̣^yụ̈p̣^y</p>	*!*****		8
b.	<p>p̣^jip̣^j p̣^jᶲp̣^j p̣^jip̣^y p̣^jᶲp̣^y</p> <p>p̣^yip̣^j p̣^yᶲp̣^j p̣^yip̣^y p̣^yᶲp̣^y</p>		*!***	8
c.	<p style="padding-left: 100px;">p̣^jip̣^j p̣^jᶲp̣^y</p> <p style="padding-left: 100px;">p̣^yip̣^j p̣^yᶲp̣^y</p>			4
d.	<p style="padding-left: 100px;">p̣ip̣ p̣ụ̈p̣</p>			2!

Candidate (a) best satisfies MAXIMIZE CONTRASTS since it allows palatalization-velarization contrasts on consonants in all positions, and front-back contrasts in vowels, however CVCs such as [p̣^jup̣^j, p̣^yip̣^y] involve substantial violations of *HIGHEFFORT since they involve movement from two full front-back movements in a short duration. Candidate (b) is intended to include CVCs which barely satisfy *HIGHEFFORT, i.e. they represent the maximum allowable effort, while maintaining vowel and consonant contrasts in all positions (indicated by somewhat arbitrary transcriptions). However, with such short vowels, the maximum allowable effort results in only small deviations from a smooth transition between the secondary articulations of the consonants (indicated by somewhat arbitrary transcriptions), and consequently indistinct F2 contrasts, so candidate (b) violates the MINDIST constraint. The winning candidate, (c), satisfies *HIGHEFFORT

since it involves only transitional vowels, transcribed here with central vowel symbols. There are no vowel F2 contrasts, and the palatalization-velarization contrasts satisfy the MINDIST constraint.

Candidate (d) satisfies the *HIGHEFFORT and MINDIST constraints by neutralizing consonant contrasts rather than vowel contrasts. This candidate loses out to (c) because it realizes less contrasts. Probably other considerations contribute to this outcome – e.g. consonant place contrasts will typically be cued by a release burst or during the consonant constriction itself, as well as by formant transitions, so they may be more distinct than extra-short vowel F2 contrasts – but it seems likely that one advantage of adopting a vertical vowel system is that many consonant contrasts can be differentiated in a relatively short duration (consonant constriction plus transitions) whereas distinct vowel contrasts take longer to realize. So abandoning vowel F2 contrasts may actually be motivated by MAXIMIZE CONTRASTS rather than being in conflict with this constraint. If correct, this suggests that vertical vowels bear similarities to the transitional vowels that break up consonant clusters in some Berber and Salishan languages (Dell and Elmedlaoui 1996, Flemming, Ladefoged, & Thomason 1994).

This analysis needs development, but it serves to give context to the central claim that complete neutralization of F2 contrasts results in the realization of F2 being governed by effort minimization, as predicted by dispersion theory. This phenomenon is difficult to analyze without constraints on contrasts because it means that the vowels which are unmarked in F2 contrasts – peripheral vowels – are very different from the vowels which are unmarked in the absence of such contrasts – i.e. transitional vowels. Without constraints on contrast, the basic analysis of generalizations about vowel inventories must reside in the ranking of markedness constraints on vowels. As Ní Chiosáin and Padgett (1997) point out, constraints against non-peripheral vowels would have to be ranked above constraints against front unrounded and back rounded vowels to account for the typological preference for these peripheral vowels (39). Basic patterns of cross-linguistic variation in vowel inventories can then be derived by ranking faithfulness at different points in this hierarchy.

(39) *i, *y, *u >> *i, *u

The problem is that it is precisely these ‘marked’ qualities which surface when backness contrasts are neutralized in vertical vowel systems, but the ranking in (39) implies that the vowels [i, u] should always be preferred to central vowels, so one of these vowels should be expected to appear in cases of neutralization. Allowing *i to rank below any of *i, *u would predict unattested basic vowel inventories such as [i, a, u]. More generally, without constraints on contrast, inventories should always include the least marked sounds, no matter what the size of the inventory is. There is simply no way to directly capture contrast-dependent generalizations about markedness.

As noted above, the problem is even more severe if it is assumed that transitional, vertical vowels are simply unspecified for [back] and [round], or [F2], because such transitional vowels do not occur in F2 contrasts at all, so a constraint against such targetless vowels must be inviolable in contrast, but violable in the absence of contrast.

It is always possible to propose a re-analysis of a patterns of contrast-dependent markedness as positional markedness. For example, vertical vowel inventories seem to be restricted to extra-short vowels so it is possible to formulate a constraint against non-transitional vowel qualities among extra-short vowels and restrict the markedness constraints in (39) to apply only to longer vowels. The strategy of proposing completely different markedness hierarchies for different positions proliferates increasingly specific constraints, which should prompt us to seek more general organizing principles of exactly the kind proposed here. In any case, it is not clear that this approach can provide an adequate account of the typology of extra-short vowel systems, as will be seen following the discussion of fully neutralizing vowel reduction in the next section.

1.1.2 Fully neutralizing vowel reduction

The other situation in which we find neutralization of F2 contrasts is in vowel reduction. In languages such as English (Hayes 1995), Southern Italian dialects (Maiden 1995) and Dutch (Booij 1995) all vowel quality distinctions are neutralized in some unstressed syllables. The resulting vowels are usually transcribed as schwa [ə]. Again, the preference for peripheral vowels is suspended when F2 contrasts are suspended, as

predicted by constraints on distinctiveness of contrasts. Instead of one of the unmarked full vowels [i, u, a] we find a ‘marked’ vowel which is not permitted in stressed syllables in these languages. Furthermore schwa is contextually variable, assimilating to surrounding consonants in much the same way as vertical vowels, suggesting that its realization is largely determined by effort minimization, as expected given the dispersion theory. This point will be substantiated first, then we will briefly outline an analysis of total vowel reduction as the end point of the kinds of duration-motivated neutralization exemplified by vowel reduction (3.2.2) and vertical vowel inventories.

To see that result of full neutralization is plausibly the minimum effort vowel it is necessary to dispel some misconceptions created by the use of the schwa symbol [ə] to represent a mid central unrounded vowel in the IPA. This is not an accurate description of the vowel that results from total neutralization, at least in English or Dutch, the languages for which acoustic data are available⁸. Firstly, it is often noted that schwa is highly variable according to context. Van Bergem (1994) shows that F2 in Dutch schwa is largely predictable from phonemic context (preceding VC and following CV), and that the F2 trajectory is an almost linear interpolation between F2 at the release of the preceding consonant and F2 at the onset of the following consonant. Kondo (1994, 2000) found similar results for English, studying a smaller range of contexts. As suggested above, there is reason to believe that this type of trajectory is the least-effort realization for a vowel, and this is van Bergem’s (1994) interpretation of Dutch schwa.

Kondo (1994) and van Bergem (1994) also found that schwa is a high vowel in English and Dutch - F1 of schwa is generally comparable to that of full high vowels such as [i, u]. This is also expected if schwa is a minimal effort vowel, because all consonants, with the exception of pharyngeals and laryngeals involve constrictions in the upper vocal tract which lower F1, so the minimal deviation from surrounding consonants will result in a vowel with a low F1. There is some deviation from F1 of surrounding consonants because producing a vowel requires a more open vocal tract than surrounding consonants

⁸ Preliminary investigation of the Bari dialect of Italian, based on recordings accompanying Valente (1975), suggests that schwa is much the same as in English, at least in this dialect of Italian.

in order to produce the raised intensity and well-defined formant structure characteristic of vowels, but minimal deviation yields a high vowel.

Taken together, these results indicate that the schwa found in environments of total neutralization is essentially similar to the high vowel [i] of a vertical vowel inventory, which helps to explain the use of schwa to transcribe this vowel in Caucasian languages (e.g. Kuipers 1960, Smeets 1984)⁹. Accordingly, both phenomena are analyzed as fundamentally similar results of very short vowel duration. As discussed above, in relation to less radical patterns of vowel reduction (3.2.2), it becomes increasingly difficult to produce a high F1 as vowel becomes shorter. We have also seen that, at shorter durations it becomes difficult to realize distinct values of F2 in a given consonantal context. Schwa is even shorter than a high vertical vowel – Kondo (2000) reports a mean duration of 34 ms for English – so it is unsurprising that both F1 and F2 contrasts are neutralized in this context, even without the range of consonant place contrasts hypothesized to influence F2 neutralization in vertical vowel inventories. (In fact, it should be noted that preceding and following vowels play a greater role in determining F2 of schwa than appears to be the case with vertical vowels).

This is a situation where vowel qualities which are not even permitted where there are vowel quality contrasts (in stressed syllables), are the least marked qualities where quality contrasts are neutralized (in unstressed syllables). This phenomenon is particularly striking because less radical vowel reduction can be analyzed as preserving the markedness relationships found in stressed syllables. That is, the common patterns of reduction from (40i) to (40ii) and from (40ii) to (40iii) involve reduction to a subset of the stressed vowels (abstracting away from raising of low vowels), so it is possible to assuming a fixed scale of vowel markedness (41). Vowel reduction can be analyzed by

⁹ Many dialects of English distinguish two reduced vowels, usually transcribed as [ə] and [ɪ], which contrast in derived minimal pairs like ‘Rosa’s’ [ɹoʊzəz] vs. ‘roses’ [ɹoʊzɪz]. In my own speech these vowels are distinguished by F2 (fronting), not F1: [ɪ] has a higher F2 than [ə]. The transcription conventions used in the TIMIT database also distinguish [ə] (AX) from [ɪ] (IX) on the basis of F2 (Seneff and Zue 1988). It is also likely that word-final schwa differs from other schwas because English allows contrasts in unstressed final syllables between [ɪ], e.g. ‘pity’, [ə], e.g. ‘Rosa’, [oʊ], e.g. ‘veto’ (Hayes 1995). The contrast between ‘Rosa’s’ and ‘roses’ is probably a cyclicity effect whereby the quality of the final schwa of ‘Rosa’ is preserved in derived ‘Rosa’s’, whereas the vowel found in ‘roses’ is the fronted reduced vowel which is expected between coronals.

distinguishing higher-ranked faithfulness to height features ([high], [low], [tense]) in stressed syllables, so more marked vowels are permitted in these positions (e.g. Beckman 1998). However, this line of analysis does not extend to fully neutralizing reduction: Pushing faithfulness to the lowest reaches of the hierarchy should yield one of the least marked vowels [i, a, u], not a putatively more marked vowel such as [ɨ].

(40)	(i)	(ii)	(iii)
	i u	i u	i u
	e o	e o	a
	ɛ ɔ	a	
	a		

(41) *ɛ, *ɔ >> *e, *o >> *i, *u, *a

Positing distinct markedness hierarchies for stressed and unstressed syllables does not solve the problem either. To account for the fact that a high transitional vowel is the result of total neutralization, this has to be ranked as the least marked vowel in unstressed syllables. But this implies that the high transitional vowel should be part of any unstressed vowel inventory, and this is not the case. A common expansion beyond a single unstressed vowel seems to be to three vowels [i, ə, u], as in Brazilian Portuguese (3.2.2). (Note that [ə] is used here in the IPA sense of a mid central unrounded vowel, as opposed to the high transitional vowel which results from total neutralization. So as soon as F2 contrasts are included in an inventory, MINDIST constraints on F2 differences become relevant, and the preference for front unrounded and back rounded vowels reasserts itself.

1.1.3 Related phenomena

Dispersion theory predicts that where no contrasts are primarily realized on a given dimension, then realization on that dimension will be governed by minimization of effort, or other contextual markedness constraints. Neutralization of F2 contrasts in vertical

vowel inventories and in fully neutralizing vowel reduction are examples of this phenomenon. There are probably many other examples of this pattern, but in some cases they can be difficult to detect because the least effort realization of a sound type is similar to a sound found in contrast. For example, in many contexts, the least-effort laryngeal state for an obstruent will be voicelessness, due to aerodynamic factors discussed in the next section. However, voiceless stops also provide a distinct contrast with voiced stops, so least effort stops may be similar to contrastively voiceless stops in many contexts. Dispersion theory leads us to expect that non-contrastive voiceless stops should be more prone to partial voicing following a preceding sonorant, but the differences involved can only be identified by instrumental analysis, so we do not have relevant data for many languages (but see Hsu 1998 for evidence of this pattern in Taiwanese). However, there is good evidence for the related prediction that effortful enhancements of stop voicing should only apply where there are voicing contrasts, as shown in the next section.

Vowel nasalization provides another probable example of this type of pattern. In the absence of vowel nasalization contrasts, it is to be expected that oral vowels should be more tolerant of contextual nasalization. Again, differences in the magnitude and extent of partial nasalization can only be determined by instrumental methods. Cohn (1990) shows that contrastive oral vowels in French undergo much less contextual nasalization than English vowels preceding a nasal, but there is no obvious difference following a nasal. A clearer pattern involves partial denasalization of nasals adjacent to contrastively oral vowels. This pattern is observed in a wide variety of languages (Anderson 1976, Herbert 1986), the most striking instance being Kaingang, where nasals are pre-nasalized preceding an oral vowel (42b), post-nasalized following an oral vowel (42c), and “medio-nasalized” between oral vowels (42d):

- (42) a. $\tilde{V}m\tilde{V}$
 b. $\tilde{V}\widehat{mb}V$
 c. $V\widehat{bm}\tilde{V}$
 d. $V\widehat{bmb}V$

Herbert 1986 analyzes these patterns of partial denasalization as motivated by the pressure to maximize the distinctiveness of vowel nasalization contrasts – the only way to make sure that a vowel following a nasal is completely oral is to raise the velum before the offset of the nasal, resulting in a brief oral stop closure. So partial denasalization is the consequence of total intolerance of contextual nasalization. As might be expected, only contrastive orality conditions denasalization, i.e. this phenomenon is primarily observed in languages with vowel nasalization contrasts. The only exceptions to this generalization are found in Australian languages, including Gupapuyŋu (Butcher 1999), where nasals are optionally pre-stopped post-vocally although there are no vowel nasalization contrasts. Butcher suggests that this partial denasalization serves to ensure that the closure transitions are oral, avoiding the destructive effect of nasalization on the distinctiveness of formant patterns (Repp and Svastikula 1988, Wright 1986, see also section 4.3 below). The distinctiveness of formant transitions is particularly important because the relevant languages distinguish 4-6 places of articulation among nasals. So the exceptional cases also appear to be motivated by distinctiveness constraints.

1.2 Enhancement of stop voicing contrasts

Another example of contrast-dependent markedness is provided by the typology of laryngeal contrasts among stops. A number of languages contrast pre-nasalized or implosive stops with voiceless unaspirated stops, but do not have plain voiced stops. The preference for pre-nasalized or implosive stops over plain voiced stops is explained on the grounds that pre-nasalized and implosive stops are more distinct from voiceless stops. That is, pre-nasalization and implosion can be regarded as enhancements of the stop voicing contrast (cf. Iverson and Salmons 1996). However these sounds are also more effortful than plain voiced stops, so most languages forgo these enhancements. Crucially enhancement of stop voicing does not occur in the absence of contrast – i.e. we do not find pre-nasalization or implosivization of inter-vocally voiced stops for example. This is expected if the only reason for exerting the additional effort involved in producing these sounds is to satisfy a constraint on the distinctiveness of contrasts, but, like other contrast-dependent patterns of distribution, it is difficult to account for without constraints on contrast.

Pre-nasalized and implosive stops are often thought of as more marked than plain voiced stops. While it is true that they are less frequent than plain voiced stops, there is no implicational relationship between these sound types: a substantial number of languages have pre-nasalized or implosive stops without having plain voiced stops. E.g. San Juan Colorado Mixtec has pre-nasalized stops but no plain voiced stops (43) (Campbell, Peterson & Lorenzo Cruz 1986). This pattern is discussed by Iverson and Salmons (1996) in relation to Mixtec, and by Herbert (1986:16ff.) who cites a number of other examples, including Fijian, Lobaha, Reef Islands-Santa Cruz languages, and South Gomen. Other examples include Southern Barasano (Smith & Smith 1971) and Guaraní (Gregores & Suárez 1967).

(43) San Juan Colorado Mixtec stops:

p	t	tʰ	k
ᵐb	ⁿd	ⁿdʰ	

Languages which contrast voiceless and implosive stops but lack plain voiced stops seem to be less common (Maddieson 1984:28), but the UPSID database of phonological inventories (Maddieson 1984) includes two examples: Nyangi and Maasai (both Eastern Sudanic). The stops of Nyangi are shown in (44). In addition, Vietnamese voiced stops are often implosive (Nguyen 1970), and Ladefoged and Maddieson (1996) report that ‘fully voiced stops in many diverse languages (e.g. Maidu, Thai and Zulu) are often accompanied by downward movements of the larynx that make them slightly implosive’ (p.78).

(44) Nyangi stops:

p	t	c	k
b	d	f	g

Voiced stops are distinguished from voiceless stops by a variety of cues. One of the most important is Voice Onset Time (Lisker and Abramson 1964, Lisker 1975), but the presence of voicing during closure (indicated by periodicity and low-frequency energy) is

also significant (Stevens and Blumstein 1981). Implosive and pre-nasalized stops are more strongly voiced than plain voiced stops, and so are better distinguished from voiceless stops in this respect. It is difficult to sustain high intensity of voicing during a stop closure because pressure builds up behind the closure until there is no longer a pressure drop across the glottis. Without a sufficient pressure drop there is no airflow through the glottis, so voicing ceases (Ohala 1983, Westbury and Keating 1986). So voicing tends to decline in intensity through a voiced stop closure. Lowering the velum allows air to be vented from the vocal tract, mitigating the pressure build-up, and thus facilitating the maintenance of high intensity of voicing. In addition, radiation from the nose results in higher intensity of the speech signal than radiation through the neck, which is the only source of sound in an oral stop (Stevens et al 1986:439).

Similarly, lowering the larynx during the stop closure, as in implosive stops, expands the oral cavity, reducing the build-up of pressure. Consequently implosives are characteristically strongly voiced. Lindau (1984) found that the amplitude of voicing actually increased through the course of an implosive closure. Implosives also have very low-intensity release bursts because the intensity of the burst depends on oral pressure at release (Ladefoged and Maddieson 1996:82). Intensity of the release burst has been shown to be a significant cue to stop voicing contrasts in English (Repp 1979), so this is also likely to make implosives more distinct from voiceless stops than plain voiced stops.

Given these considerations, it seems likely that languages like Mixtec and Nyangi prefer prenasalized-voiceless and implosive-voiceless stop contrasts over the more common voiced-voiceless contrast because the former are more distinct contrasts. This analysis follows Henton, Ladefoged, and Maddieson (1992), who suggest this interpretation of pre-nasalization, and Iverson and Salmons (1996), who extend it to implosion. The conflicting constraint that leads many languages to forgo maximizing distinctiveness is probably effort minimization. Implosives involve more effort than plain voiced stops because they involve an additional larynx-lowering gesture. Pre-nasalized stops require fairly rapid lowering and raising of the velum. These sounds probably also yield less distinct manner contrasts than plain voiced stops since they have higher intensity during closure, whereas stops are generally distinguished from continuants by their low intensity. The low intensity of implosive release bursts may also make them less

distinct from continuants since the presence of a release burst also distinguishes plosives from other consonants.

This analysis can be formalized as follows. We will assume that at least two dimensions distinguish voiced and voiceless stops: VOT and strength of voicing ([voice]), which could be quantified in terms of the intensity of the periodic part of the speech signal, for example (45).

(45) VOT dimension	<u>0</u>	<u>1</u>	<u>2</u>
	d	t	t ^h
Voice dimension	<u>0</u>	<u>1</u>	<u>2</u>
	t	d	ɖ
			ⁿ d

The relevant (universal) ranking of MINDIST constraints is shown in (46). Contrasts such as [d-t] satisfy the higher-ranked constraint only, whereas [ⁿd-t] and [ɖ-t] satisfy both constraints.

(46) MINDIST = VOT:1+VOICE:1 >> MINDIST = VOT:1+VOICE:2

For present purposes the fact that pre-nasalized stops and implosives involve greater effort than plain voiced stops will be implemented as a fixed ranking of constraints against these sound types (47).

(47) *IMPLOSIVE, *PRE-NASALIZED STOP >> *VOICED STOP

Then a language like Nyangi, with implosives in place of voiced stops, is derived by the following ranking, as shown in (49).

- (48) MINDIST = VOT:1+voice:1 >>
MINDIST = VOT:1+voice:2 >>
MAXIMIZE CONTRASTS,
*PRE-NASALIZED STOP >>
*IMPLOSIVE >>
*VOICED STOP

(49)		MINDIST = VOT:1+ VOICE:1	MINDIST = VOT:1+ VOICE:2	MAXIMIZE CONTRASTS	*PRE- NASALIZED STOP	*IMPLOSIVE	*VOICED STOP
a.	t-d		*!	✓✓			*
b.	t-d̥			✓✓		*	
c.	t- ⁿ d			✓✓	*!		
d.	t			✓!			

We will assume for now that the preference for implosives over pre-nasalized stops depends purely on the relative ranking of the effort-minimization constraints against these sounds types, so the ranking in (48) derives implosives where *PRE-NASALIZED STOP >> *IMPLOSIVE (cf. 49c), while pre-nasalized stops are derived if this ranking is reversed. The more common voiced-voiceless contrast is derived if MINDIST = VOT:1 + VOICE 2 is ranked below both of these effort minimization constraints (50).

(50)		MINDIST = VOT:1+ VOICE:1	MAXIMIZE CONTRASTS	*PRE- NASALIZED STOP	*IMPLOSIVE	MINDIST = VOT:1+ VOICE:2	*VOICED STOP
a.	t-d̥		✓✓			*	*
b.	t-d̥		✓✓		*!		
c.	t- ⁿ d		✓✓	*!			
d.	t		✓!				

If a voicing contrast is not maintained, the distinctiveness of voicing contrasts is irrelevant, so voicing of stops is determined primarily by effort minimization. In many contexts effort minimization prefers devoicing of stops due to aerodynamic factors reviewed above, but in some contexts, e.g. following a nasal or in short stops between vowels, voicing appears to be easier to produce and many languages follow effort minimization, resulting in allophonically voiced stops in these contexts (Westbury and Keating 1986, Kirchner 1998). For example, stops are voiced intervocalically and following nasals in Tümpisa Shoshone (Dayley 1989, Kirchner 1998). Implosives and prenasalized stops, on the other hand, are never preferred by effort minimization constraints, so these sounds are only expected in contrast with voiceless stops.

The patterns of distribution analyzed here involve a contrast-dependent generalization: implosives and pre-nasalized stops can be preferred to voiced stops where they contrast with voiceless stops, but they are never preferred to voiced stops where there is no voicing contrast. That is, there is no post-nasal implosivization or intervocalic pre-nasalization. This situation is difficult to account for without constraints on contrasts because any simple way of deriving implosives/pre-nasalized stops in place of voiced stops without these constraints is liable to predict that these sounds could also be preferred in the absence of contrast.

In a theory without constraints on contrasts, a preference for implosives over voiced stops implies a ranking of constraints with the effect of that shown in (51). The exact formulation of *VOICED STOP and *IMPLOSIVE is not important, it is only necessary that one favors implosive stops over voiced stops, and the other effectively imposes the reverse preference. We must also assume that faithfulness to the feature that differentiates implosives from plain voiced stops (e.g. [lowered larynx]) is low-ranked throughout to explain the absence of contrasts between plain voiced and implosive stops.

(51) IDENT[VOICE], *VOICED STOP >> *IMPLOSIVE

The reverse ranking of *VOICED STOP and *IMPLOSIVE would also have to be allowed to derive the usual voiced-voiceless contrast:

(52) IDENT[VOICE], *IMPLOSIVE >> *VOICED STOP


The problem arises when these ranking possibilities are combined with rankings required to analyze allophonic variation in languages without voicing contrasts. The basic ranking for a language without stop voicing contrasts has to place IDENT[VOICE] below the effort minimization constraints:

(53) *IMPLOSIVE >> *VOICED STOP >> IDENT[VOICE]

To derive intervocalic voicing, it is necessary to differentiate the markedness of voiced stops between vowels from their markedness in other contexts. A simple approach is to posit a constraint against intervocalic voiceless stops, *VOICELESS STOP/V_V, ranked above the general constraint against voiced stops (54). But we have already seen that *VOICED STOP must be able to out-rank *IMPLOSIVE to account for languages with implosives but no plain voiced stops. So nothing prevents reversing the ranking of these constraints, as in (55-56), which derives the unattested phenomenon of intervocalic voicing implosivization, i.e. stops are implosive between vowels (55), but voiceless elsewhere (56), because this ranking makes it preferable to replace any voiced stop by an implosive.

(54) *VOICELESS STOP/V_V, *IMPLOSIVE >> *VOICED STOP >> IDENT[VOICE]

(55)

	/ata/	*VOICELESS STOP/V_V	*VOICED STOP	*IMPLOSIVE	IDENT [VOICE]
a.	ata	*!			
b.	ada		*!		*
c.	 ada			*	*

(56)

	/ad/	*VOICELESS STOP/V_V	*VOICED STOP	*IMPLOSIVE	IDENT [VOICE]
a.	at				*
b.	ad		*!		
c.	ad'			*!	

The problem is that without constraints on contrast it is not possible to express the fact that implosives and prenasalized stops are only favored because they yield more distinct contrasts with voiceless stops (or an additional contrast), so nothing favors implosives in the absence of stop voicing contrasts. Without constraints on contrasts, it is necessary to posit constraints favoring implosives and prenasalized stops independent of contrast, which then predicts that these sounds could be preferred over plain voiced stops in the absence of contrast. The preference for implosives and prenasalized stops must be strictly dependent on the presence of a contrast, which implies constraints on contrasts.

1.3 Allophonic and contrastive nasalization

The pattern observed in the previous section could be characterized as showing that allophonic stop voicing is not subject to enhancement, whereas contrastive voicing can be. I.e. it provides evidence that allophonic stop voicing behaves differently from contrastive stop voicing. Comparing the behavior of allophonic and contrastive instances of a sound type is a good way to investigate the prediction that the markedness of a sound depends on the sounds that it contrasts with, because the same sound types can be observed in different systems of contrasts. E.g. an allophonically nasalized vowel generally contrasts with other nasalized vowels, but by definition it does not minimally contrast with its oral counterpart, whereas a contrastively nasalized vowel minimally contrasts both with other nasal vowels and with its oral counterpart. In this section we will see that the markedness of an allophonically nasalized sound can differ from the markedness of the same sound where nasalization is minimally contrastive. This difference is unexpected if there are no constraints on contrasts because then a constraint

against nasalized vowels, for example, applies equally to allophonically and contrastively nasalized vowels.

In dispersion theory, a contextually nasalized vowel will generally violate more MINDIST constraints than its oral equivalent because nasalization reduces the distinctiveness of vowel quality contrasts, particularly those involving F1 (Wright 1986, Beddor 1993). The markedness of a contrastively nasalized vowel depends on this factor also, but it is more dependent on the distinctiveness of the contrast with its oral counterpart. This distinction is confirmed by the existence of mismatches between the typologies of nasalization contrasts and resistance to allophonic nasalization in nasal harmony, discussed by Ní Chiosáin and Padgett (1997).

Blocking of nasal harmony can be illustrated from the Johore dialect of Malay (Onn 1980). In this language nasality is only contrastive on stops, and spreads rightward from a nasal stop onto a sequence of vowels, glides and laryngeals (56a). All other segment types block the spread of nasalization (56b).

- (57) a. mĩnĩm ‘to drink’ baŋõn ‘to rise’
 mãʔãp ‘to pardon’ pəŋõŋãhãhã ‘central focus’
 mãʔãŋ ‘stalk (palm)’ mõnãwãwã ‘to capture (active)’
- b. mõratappi ‘to cause to cry’ pəŋãwãsan ‘supervision’
 mãkan ‘to eat’

Languages with nasal harmony vary as to which segments block the spread of nasality. Building on surveys by Schourup (1972), Piggot (1992) and Cohn (1993), Walker (1998) shows that segments can be arranged into the hierarchy shown in (58) such that if any segment type blocks nasal harmony, all segment types lower on the hierarchy block nasal harmony as well¹⁰. So in Johore Malay the top two levels of the

¹⁰ The hierarchy (58) follows Walker and Pullum (1999) in placing laryngeals at the same level as vowels. Walker (1998) does not actually include laryngeals in any statement of the hierarchy, although the data supporting its placement with vowels are reported there.

hierarchy undergo nasalization – liquids and segments lower on the hierarchy block harmony. Sundanese (Robins 1957, Cohn 1990) is similar to Malay, but glides also block nasalization.

(58) vowels, laryngeals > glides > liquids > fricatives > obstruent stops

Walker (1998) analyzes this generalization in terms of a corresponding hierarchy of constraints on the ‘compatibility’ of different segment types with nasality:

(59) *NASOBSSTOP >>
*NASFRICATIVE >>
*NASLIQUID >>
*NASGLIDE >>
*NASVOWEL >>
*NASSONSTOP

Walker does not propose any constraint on nasalized laryngeals, but remarks that they typically pattern with vowels (p.50), so *NASLARYNGEAL should presumably be ranked at the same level as *NASVOWEL in the hierarchy. Different patterns of blocking are then derived by ranking the constraint which motivates nasal harmony, SPREAD[+NASAL], at different points in this hierarchy. Segment types subject to nasality constraints ranked below SPREAD[+NASAL] undergo nasalization harmony, while those subject to higher-ranking constraints block harmony. E.g. the ranking for Johore Malay is as in (60).

- (60) *NASOBSSTOP >>
 *NASFRICATIVE >>
 *NASLIQUID >>
SPREAD[+NASAL] >>
 *NASGLIDE >>
 *NASVOWEL >>
 *NASSONSTOP

The difficulty faced by this approach is that these same markedness constraints should also account for the typology of nasalization contrasts – placing IDENT[NASAL] at different points in the hierarchy should yield the typology of nasalization contrasts. As Walker observes (p.53), the predicted pattern is broadly correct: the most common nasal sounds are nasal stops, and the next most common are nasalized vowels, while other contrastively nasalized sounds are rare (Maddieson 1984, Cohn 1993). However, Ní Chiosáin and Padgett (1997) note that the predictions concerning nasalized laryngeals are problematic. Nasalization is never contrastive on glottal stops, suggesting that *NASLARYNGEAL should be high-ranked, but on the other hand laryngeals are among the sounds most susceptible to nasal harmony, implying that *NASLARYNGEAL should be low-ranked. That is, contrastive nasalization of laryngeals is very marked, but non-contrastive nasalization is unproblematic. This situation is predicted by the contrast-based analysis of blocking and contrast: Nasalization contrasts on glottal stops are unsatisfactory because nasalized and oral glottals are acoustically identical (cf. Walker and Pullum 1999), but by the same token, contrasts between glottal stops and other consonants are unaffected by contextual nasalization of the glottal. The acoustics of [h] are also relatively unaffected by velum lowering (Ohala 1975:301). The problem with Walker’s approach is that it conflates markedness of contrastive and non-contrastive nasalization.

A faithfulness-based analysis according to which nasal harmony is blocked because it violates faithfulness to the underlying [-nasal] specification of the blocking segment fares no better. Such an account must distinguish IDENT[NASAL] constraints for different segment types, e.g. IDENT[NASAL]/FRICATIVE, IDENT[NASAL]/LIQUID, etc. The blocking

hierarchy is then derived by imposing a universal ranking on these faithfulness constraints:

- (61) IDENT[NAS]/OBSSTOP >>
IDENT[NAS]/FRICATIVE >>
IDENT[NAS]/LIQUID >>
IDENT[NAS]/GLIDE >>
IDENT[NAS]/VOWEL,
IDENT[NAS]/LARYNGEAL

Again the problem is that these faithfulness constraints are required to do double duty: they must also account for the typology of nasalization contrasts. Nasality contrasts should be derived by the ranking of a general markedness constraint, presumably *[+NASAL]. However, this makes completely inaccurate predictions concerning the typology of nasal contrasts: E.g. the high ranking of IDENT[NAS]/FRICATIVE motivated by the resistance of fricatives to nasalization implies that nasalization contrasts on fricatives should be common also. At the other extreme, the susceptibility of vowels to contextual nasalization should imply that nasalization contrasts on vowels are more marked than nasalization contrasts on liquids or glides¹¹. Again, the diagnosis is that this approach fails to distinguish markedness of contrastive nasalization from markedness of non-contrastive, contextual nasalization.

Dispersion theory predicts that nasal harmony should be blocked where nasalization would be articulatorily difficult, or where it would give rise to indistinct contrasts. Ní Chiosáin and Padgett propose a dispersion-based analysis of the behavior of laryngeals according to which nasalization contrasts do not arise on laryngeals because they would be indistinct, as above, but laryngeals do not block nasal harmony because a nasalized laryngeal is a low effort sound, implying an account of blocking in terms of articulatory

¹¹ Adopting the set of markedness constraints in (59) rather than undifferentiated *[+nasal] avoids some of the problematic predictions of a pure faithfulness account at the cost of proliferating constraints, but it still fails to account for the behavior of laryngeals: although IDENT[NAS]/LARYNGEAL is ranked low, so is *NASLARYNGEAL, so it is predicted that contrastively nasalized laryngeals can be derived while excluding all other nasalization contrasts.

compatibility with nasalization. However, there is no obvious articulatory difficulty in lowering the velum during any sound type (although it may result in a change in manner as well as nasalization). So the blocking hierarchy is more plausibly derived from distinctiveness constraints (a possibility that Ní Chiosáin and Padgett (1997) also consider (fn.25)). Specifically, most of the blocking hierarchy (57, above) can be derived from the generalization that nasal harmony is blocked where nasalization would endanger contrasts with nasal stops. That is, blocking of nasal harmony is a consequence of MINDIST constraints blocking the creation of indistinct contrasts between nasals and nasalized consonants. Since laryngeals are largely unaffected by nasalization, they remain highly distinct from voiced nasals, and consequently are least likely to block nasal harmony.

As for the rest of the hierarchy, nasalizing a voiced stop actually results in a nasal, and so would neutralize contrasts. Approximant consonants such as glides, laterals and rhotics are already similar to nasals in that these are all are sonorant consonants, and lowering the velum further reduces this difference. In general, the narrower the oral constriction of a nasalized approximant, the more similar it will be to a nasal stop. I.e. in a nasal stop, all airflow is through the nasal cavity, whereas in a nasalized approximant there is airflow through both oral and nasal cavities, but a narrower oral constriction results in less airflow through the oral cavity, and correspondingly more airflow through the nasal cavity, resulting in closer approximation to a nasal stop. Thus nasalized laterals, e.g. [l̃], and flaps [ɾ̃] are most similar to nasal stops, and nasalized glides, [w̃, j̃] somewhat less so, and nasalized vowels least of all. Nasalizing a voiced fricative is liable to lead to loss of frication because most air is then vented through the nose, leaving insufficient oral pressure to generate frication (Ohala and Ohala 1993:227f.), so the basic result is a narrowly constricted nasalized approximant, which presumably is slightly closer to a nasal than a nasalized liquid. So nasalizing sounds which are higher on the blocking hierarchy results in less distinct contrasts with nasal stops, i.e. violation of higher-ranked MINDIST constraints. So the blocking segments in a particular nasal harmony system depend on the position of SPREAD[+NASAL] in the hierarchy of MINDIST constraints. I.e. most of the constraint hierarchy in (59) can be replaced by the following

hierarchy of MINDIST constraints (where NAS 'FRICATIVE' refers to the sound that results from lowering the velum during a voiced fricative):

- (62) MINDIST = NAS'FRICATIVE'-NASSTOP >>
MINDIST = NASLIQUID-NASSTOP >>
MINDIST = NASGLIDE-NASSTOP >>
MINDIST = NASVOWEL-NASSTOP >>
MINDIST = NASLARYNGEAL-NASSTOP

Distances in (62) are expressed descriptively because it is not clear what dimensions distinguish these sounds. Indeed, nasalized vowels and approximants raise potential difficulties for the formant-based approach to spectral quality adopted above because nasal coupling introduces additional formants (resonances of the nasal cavity), so formants of nasalized sounds do not correspond straightforwardly to the formants of oral sounds. In addition nasal sounds generally include spectral zeroes whose primary effect is to reduce the intensity of nearby formants (Maeda 1993), and formant intensities are not dimensions we have considered so far.

Blocking by voiceless consonants has a different basis. Nasalizing a voiceless stop results in a voiceless nasal, which is quite distinct from a voiced nasal, but are highly dispreferred for other reasons. A fully devoiced nasal is similar to [h] and also yields indistinct place contrasts since noise is generated mainly at the nostrils, and so is the same regardless of oral place of articulation (Ohala and Ohala 1993:232). Note that contrastively voiceless nasals are actually voiced during part of the nasal closure (Ladefoged and Maddieson 1996:113). Ohala and Ohala suggest that this realization is adopted precisely to diminish the problems just outlined. Nasalizing a voiceless fricative would yield similar results, since frication would be lost, as in nasalization of a voiced fricative.

The typology of nasalization contrasts depends partly on the distinctiveness facts just outlined, but many other distinctiveness relations are significant, in particular the distinctiveness of contrasts between corresponding oral and nasal sounds, which is irrelevant where nasalization is non-contrastive. These additional considerations explain

the discrepancies between the typologies of blocking and contrast. As already noted, the acoustic near-identity between oral and nasalized laryngeals explains why these sounds are not found in contrast¹².

Nasal stops seem to provide good contrasts with oral stops and approximants, and consequently are found in almost all languages (Ferguson 1963, Maddieson 1984). This is probably because they are well differentiated from obstruents by intensity, and from other sonorants by their distinctive formant structure. Nasals have more closely spaced formants than oral sonorants because the nasal-pharyngeal tract is longer than the oral tract. Nasal formant structure is further differentiated from most oral sonorants by the presence of spectral zeroes at the resonant frequencies of the oral cavity behind the closure (Fujimura 1962, Stevens 1999:487ff.) (laterals also have spectral zeroes, but at higher frequencies than in nasals). Given the desirability of nasal stops, the undesirability of nasalized approximants follows from fact that nasalized approximants yield poor contrasts with nasal stops for the reasons outlined above in the analysis of blocking by approximants.

2 Conclusion: working with constraints on contrast

We have now seen substantial evidence that phonology includes constraints on contrasts, specifically constraints that favor maximizing the distinctiveness of contrasts (MINDIST), and a constraint that favors maximizing the number of contrasts (MAXIMIZE CONTRASTS). We have also seen that these constraints do not operate independently from more familiar syntagmatic markedness constraints, e.g. as a theory of inventories, somehow operating outside of conventional phonological analyses. The interaction between syntagmatic and paradigmatic constraints is central to the derivation of basic phenomena such as neutralization (3.2.2) and blocking in harmony processes (4.3). According to the dispersion theory, the set of well-formed words in a language represents an optimal balance between the number and distinctiveness of the contrasts between words, and constraints that define preferred sound sequences, such as effort minimization

¹² Ladefoged and Maddieson (1996:133) report the existence of contrastively nasalized [h̃] in Kwangali, but nasalization is also realized on the following vowel, so it is possible to regard this as essentially a vowel

and metrical constraints. However combining paradigmatic and syntagmatic constraints in this way does result in a system with very different properties from an OT grammar based on conventional constraints because constraints on the distinctiveness of contrasts evaluate relationships between forms. So if we want to determine whether a putative word is well-formed, we must consider whether it is sufficiently distinct from neighboring words. But these words must also be well-formed, which implies assessing their distinctiveness from neighboring words, and so on. Thus it seems that we cannot evaluate the well-formedness of a single word without determining the set of all possible words.

The analyses above avoid this problem by considering only the evaluation of inventories of contrasting sounds, or short strings of sounds, in a particular context rather than evaluating complete words. For example, evaluating vowel inventories effectively involves determining the set of contrasting sounds that are permitted in a syllable nucleus. This makes the evaluation of MINDIST and MAXIMIZE CONTRASTS straightforward since only a small number of contrasting sounds are possible in a given context. This simplification is valid given certain assumptions. First, the context must be well-formed. E.g. if we are evaluating the set of vowels that can appear before a nasal stop, it must be true that nasal stops are part of an inventory of consonant contrasts that can occur in post-vocalic position. Second, nothing outside of the specified context should be relevant. I.e. no constraint that is ranked high enough to affect the well-formedness of the inventory should refer to material outside of the specified context.

More generally, the strategy for avoiding the problem of mass comparisons is to derive generalizations about the set of possible words in a language – e.g. stressed vowels are all drawn from a certain set – rather than deriving particular words. But this strategy is not actually novel, it is the usual approach to phonological analysis. In a grammar without paradigmatic constraints, it is in principle possible to determine whether an individual word is well-formed with respect to a particular constraint ranking – it involves ‘parsing’ the word to show whether or not some input yields that word as output. But demonstrating that a grammar can derive an individual word is not usually the goal

nasalization contrast which is restricted to environments following a consonant which is highly compatible with contextual nasalization conditioned by a nasalized vowel.

of phonological analysis of a language, the goal is to devise a grammar that derives all and only the possible words of that language. The usual intermediate goal is to derive generalizations about all the possible words of the language, exactly as in the analyses here.

For example, in analyzing a language it is usual to restrict attention to a single process, e.g. place assimilation between nasals and stops, ignoring stress assignment, distribution of vowels, etc. Such an analysis may be illustrated by deriving complete words, e.g. /kanpa/ → [kampa], but in itself this is uninteresting. The real goal is to derive the generalization that nasals are always homorganic to following stops. Properly, establishing such a generalization requires showing that no contrary output is derived if all possible inputs are passed through the grammar (Prince and Smolensky 1993:91). So, with or without paradigmatic constraints, there is an important distinction between deriving individual words using a grammar and reasoning about the properties of the set of words derived by that grammar. Constraints on contrast make complete derivation of individual words difficult, but that does not preclude deriving generalizations about possible words.

This is not to imply that paradigmatic constraints do not make grammars harder to work with, only that there is no reason to believe that the problems are intractable. I hope that the arguments presented in this paper provide adequate motivation for tackling them.

3 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, Stephen R. (1976). Nasality and the internal structure of segments. *Language* 52, 326-344.
- Beckman, Jill (1998). *Positional Faithfulness*. Ph.D. dissertation, University of Massachusetts, Amherst.

- Beddor, Patrice S. (1993). The perception of nasal vowels. Marie K. Huffman and Rena A. Krakow (eds.) *Nasals, Nasalization, and the Velum, Phonetics and Phonology vol. 5*. Academic Press, San Diego, 171-196.
- Bender, Byron W. (1968). Marshallese Phonology. *Oceanic Linguistics* 7, 16-35.
- Benua, Laura (1997). *Transderivational Identity*. Ph.D. dissertation, University of Massachusetts, Amherst.
- Boersma, Paul (1998). *Functional Phonology*. PhD dissertation, University of Amsterdam.
- Booij, Geert (1995). *The Phonology of Dutch*. Oxford University Press, Oxford.
- Burgess, Eunice, and Patricia Ham (1968). Multilevel conditioning of phoneme variants in Apinayé. *Linguistics* 41, 5-18.
- Burzio, Luigi (1998). Multiple correspondence. *Lingua* 104, 79-109.
- Butcher, Andrew (1999). What speakers of Australian aboriginal languages do with their velums and why: The phonetics of the nasal/oral contrast. *Proceedings of the International Congress of Phonetic Sciences 1999*, 479-482.
- Byrd, Dani (1992). Marshallese suffixal reduplication. Jonathon Mead (ed.) *Proceedings of WCCFL XI*, 61-77.
- Campbell, Sara Stark, Andrea Johnson Peterson, and Filiberto Lorenzo Cruz (1986). *Diccionario Mixteco de San Juan Colorado*. Instituto Lingüístico de Verano, México, D.F.
- 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.
- Chomsky, Noam, and Morris Halle (1968). *The Sound Pattern of English*. New York: Harper and Row.
- Cohn, Abigail C. (1990). *Phonetic and Phonological Rules of Nasalization (UCLA Working Papers in Phonetics 76)*. Ph.D. dissertation, University of California, Los Angeles.

- Cohn, Abigail C. (1993a). A survey of the phonology of the feature [nasal]. *Working Papers of the Cornell Phonetics Laboratory* 8, 141-203.
- Colarusso, John (1988). *The Northwest Caucasian Languages: A Phonological Survey*. Garland Publishing, New York.
- Colarusso, John (1992). *A Grammar of the Kabardian Language*. University of Calgary Press, Calgary.
- Dayley, Jon P. (1989). *Tümpisa (Panamint) Shoshone grammar*. University of California Press, Berkeley.
- Delattre, Pierre C., Alvin M. Liberman, Franklin S. Cooper, and Louis J. Gerstman (1952). An experimental study of the acoustic determinants of vowel color: observations on one- and two-formant vowels synthesized from spectrographic patterns. *Word* 8, 195-210.
- Dell, François, and Mohammed Elmedlaoui (1996). On consonant releases in Imdlawn Tashlhiyt Berber. *Linguistics* 34, 357-395.
- Erickson, Molly (2000). Simultaneous effects on vowel duration in American English: A covariance structure modeling approach. *Journal of the Acoustical Society of America* 108, 2980-2995.
- 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.
- Ferguson, Charles A. (1963). Assumptions about nasals: A sample study in phonological universals. Joseph Greenberg (ed.) *Universals of Language*. MIT Press, Cambridge, 53-60.
- Flemming, Edward (1995). *Auditory Representations in Phonology*. PhD dissertation, UCLA.
- Flemming, Edward, Peter Ladefoged and Sarah G. Thomason (1994). The phonetic structures of Montana Salish. *UCLA Working Papers in Phonetics* 87, 1-34.
- Fujimura, Osamu (1963). Analysis of nasal consonants. *Journal of the Acoustical Society of America* 32, 1865-1875.
- Goldstone, R. L., Steyvers, M., Spencer-Smith, J., & Kersten, A. (2000). Interactions between perceptual and conceptual learning. E. Diettrich & A. B. Markman (eds.)

- Cognitive Dynamics: Conceptual Change in Humans and Machines*. Lawrence Erlbaum Associates, Mahwah, NJ, 191-228.
- Goldstone, Robert (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General* 123, 178-200.
- Gregores, Emma, and Jorge A. Suárez (1967). *A Description of Colloquial Guaraní*. Moton, The Hague.
- Guenther, F.H., F.T. Husain, M.A. Cohen, and B.G. Shinn-Cunningham (1999). Effects of categorization and discrimination training on auditory perceptual space. *Journal of the Acoustical Society of America* 106, 2900-2912.
- Halle, Morris (1959). *The Sound Pattern of Russian*. Mouton, The Hague.
- Hayes, Bruce (1995). *Metrical Stress Theory: Principles and Case Studies*. University of Chicago Press, Chicago.
- Hayes, Bruce (1999). Phonological restructuring in Yidiny and its theoretical consequences. Ben Hermans and Marc Oostendorp (eds.) *The Derivational Residue in Phonological Optimality Theory*. John Benjamins, Amsterdam, 175-205.
- Henton, Caroline, Peter Ladefoged, and Ian Maddieson (1992). Stops in the world's languages. *Phonetica* 49, 65-101.
- Herbert, Robert K. (1986). *Language Universals, Markedness Theory and Natural Phonetic Processes*. Mouton de Gruyter, Berlin.
- Hsu, Chai-Shune K. (1996). Voicing underspecification in Taiwanese word-final consonants. *UCLA Working Papers in Phonetics* 96, 90-105.
- Hume, Elizabeth, and Keith Johnson (forthcoming). A model of the interplay of speech perception and phonology. Elizabeth Hume and Keith Johnson (eds.) *The Role of Speech Perception in Phonology*. Academic Press, New York.
- Iverson, Gregory K., and Joseph C. Salmons (1996). Mixtec prenasalization as hypervoicing. *International Journal of American Linguistics* 62, 165-175.
- Jusczyk, Peter W. (1993). From general to language-specific capacities: the WRAPSA Model of how speech perception develops. *Journal of Phonetics* 21, 3-28.
- Keating, Patricia A. (1985). Universal phonetics and the organization of grammars. Victoria A. Fromkin (ed.) *Phonetic Linguistics*. New York: Academic Press, 115-32.

- Kenstowicz, Michael (1996). Uniform exponence and base identity. J. Durand and B. Laks (eds.) *Current Trends in Phonology*. CNRS and University of Salford Publications, 363-393.
- Kirchner, Robert (1997). Contrastiveness and faithfulness. *Phonology* 14, 83-111.
- Kirchner, Robert (1998). *An Effort-Based Approach to Consonant Lenition*. PhD dissertation, UCLA.
- 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.
- Kuipers, Aert H. (1960). *Phoneme and Morpheme in Kabardian (Eastern Adyghe)*. *Janua Linguarum*, series minor, no. 8. Mouton, The Hague.
- Ladefoged, Peter, and Ian Maddieson (1996). *The Sounds of the World's Languages*. Blackwell, Oxford.
- Laycock, D.C. (1965). *The Ndu Language Family*. Linguistics Circle of Canberra, Series C, no.1. Australian National University, Canberra.
- Lehiste, Ilse (1970). *Suprasegmentals*. MIT Press, Cambridge, MA.
- Liljencrants, Johan, and Björn Lindblom (1972). Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language* 48, 839-62.
- Lindau, Mona (1984). Phonetic differences in glottalic consonants. *Journal of Phonetics* 12, 147-155.
- 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 (1990). Phonetic content in phonology. *PERILUS* 11, 101-118.
- Lisker, Leigh (1975). Is it VOT or a first formant detector? *Journal of the Acoustical Society of America* 57, 1547-1551.

- Lisker, Leigh, and Arthur S. Abramson (1964). A cross-language study of voicing in initial stops. *Word* 20, 384-422.
- Livingston, Kenneth R., Janet K. Andrews, and Steven Harnad (1998). Categorical perception effects induced by category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24, 732-753.
- Maddieson, Ian (1984). *Patterns of Sounds*. Cambridge University Press, Cambridge.
- Maeda, Shinji (1993). Acoustics of vowel nasalization and articulatory shifts in French nasal vowels. Marie K. Huffman and Rena A. Krakow (eds.) *Nasals, Nasalization, and the Velum, Phonetics and Phonology vol. 5*. Academic Press, San Diego, 147-170.
- Maiden, Martin (1995). Vowel systems. Martin Maiden and Mair Parry (eds.) *The Dialects of Italy*. Routledge, London, 7-14.
- Major, Roy C. (1992). Stress and rhythm in Brazilian Portuguese. Donald P. Macedo and Dale A. Koike (eds.) *Romance Linguistics: The Portuguese Context*. Bergin and Garvey, Westport, 3-30.
- Martinet, Andre (1952). Function, structure, and sound change. *Word* 8, 1-32.
- Martinet, Andre (1955). *Economie des Changements Phonétiques*. Berne: Francke.
- Mattoso Camara, Joaquim (1972). *The Portuguese Language*. University of Chicago Press, Chicago.
- Mazzola, Michael L. (1976). *Proto-Romance and Sicilian*. Peter de Ridder, Lisse.
- Moraes, João A. de (1998). Brazilian Portuguese. Daniel Hirst and Albert Di Cristo (eds.) *Intonation Systems*. Cambridge University Press, Cambridge.
- Nelson, W.L. (1983). Physical principles for economies of skilled movement. *Biological Cybernetics* 46, 135-147.
- Nguyen, Dang Liem (1970). *A contrastive phonological analysis of English and Vietnamese*, vol. 4. Pacific Linguistics Series C, No. 8. Australian National University, Canberra.
- Ní Chiosáin, Máire, and Jaye Padgett (1997). Markedness, segment realization, and locality in spreading. Report no. LRC-97-01, Linguistics Research Center, University of California, Santa Cruz.

- Nosofsky, Robert M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General* 115, 39-57.
- Nosofsky, Robert M. (1992). Similarity scaling and cognitive process models. *Annual Review of Psychology* 43, 25-53.
- Ohala, John J. (1975). Phonetic explanations for nasal sound patterns. Charles A. Ferguson, Larry M. Hyman, and John J. Ohala (eds.) *Nasálfest: Papers from a Symposium on Nasals and Nasalization*. Stanford University, 289-316.
- Ohala, John J. (1983). The origins of sound patterns in vocal tract constraints. Peter F. MacNeilage (ed.) *The Production of Speech*. Springer-Verlag, New York, 189-216.
- Ohala, John J., and Manjari Ohala (1993). The phonetics of nasal phonology: Theorems and data. Marie K. Huffman and Rena A. Krakow (eds.) *Nasals, Nasalization, and the Velum, Phonetics and Phonology vol. 5*. Academic Press, San Diego, 225-250.
- Onn, Farid M. (1980). *Aspects of Malay Phonology and Morphology*. Ph.D. dissertation, Universiti Kebangsaan Malaysia, Bangi.
- Passy, Paul (1891). *Etude sur les changements phonétiques et leur caractères généraux*. Paris: Librairie Firmin-Didot.
- Perkell, Joseph S. (1997). Articulatory processes. William J. Hardcastle and John Laver (eds.) *The Handbook of Phonetic Sciences*. Blackwell, Oxford, 333-370.
- Peterson, Gordon E., and Ilse Lehiste (1960). Duration of syllabic nuclei in English. *Journal of the Acoustical Society of America* 32, 693-703.
- Piggott, Glynn (1992). Variability in feature dependency: The case of nasality. *Natural Language and Linguistic Theory* 10, 33-78.
- Plomp, Reinier (1975). Auditory analysis and timbre perception. Gunnar Fant and Michel Tatham (eds.) *Auditory Analysis and Perception of Speech*. Academic Press, New York, 7-22.
- Prince, Alan, and Paul Smolensky. (1993) *Optimality Theory: Constraint Interaction in Generative Grammar*. To appear: Cambridge, MA: MIT Press.
- Repp, Bruno H. (1979). Relative amplitude of aspiration noise as a voicing cue for syllable-initial stop consonants. *Language and Speech* 27, 173-189.
- Repp, Bruno H., and K. Svastikula (1988). Perception of the [m]-[n] distinction in VC syllables. *Journal of the Acoustical Society of America* 83, 238-247.

- Robins, R.H. (1957). Vowel nasality in Sundanese. *Studies in Linguistic Analysis*. Blackwell, London, 87-103.
- Rosner, B.S, and J.B. Pickering (1994). *Vowel Perception and Production*. Oxford University Press, Oxford.
- Schourup, Lawrence (1972). Characteristics of vowel nasalization. *Papers in Linguistics* 5, 530-548.
- Schwartz, Jean-Luc, Denis Beutemps, Christian Abry, and Pierre Escudier (1993). Inter-individual and cross-linguistic strategies for the production of the [i] vs.[y] contrast. *Jornal of Phonetics* 21, 411-425.
- Seneff, Stephanie and Victor. W. Zue (1996). Transcription and alignment of the TIMIT Database. Hiroya Fujisaki (ed.) *Recent Research Towards Advanced Man-Machine Interface Through Spoken Language*. Elsevier, New York,
- Shepard, Roger N. (1957). Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space. *Psychometrika* 22, 325-345.
- Shepard, Roger N. (1972). Psychological representation of speech sounds. Edward David and Peter Denes (eds.) *Human Communication: A Unified View*. McGraw-Hill, New York, 67-113.
- Shepard, Roger N. (1991) Integrality versus separability of stimulus dimensions: From an early convergence of evidence to a proposed theoretical basis. Gregory R. Lockhead and James R. Pomerantz, James R. (eds.) *The perception of structure: Essays in honor of Wendell R. Garner*.p. American Psychological Association, Washington, DC, 53-71.
- Smeets, Rieks (1984). *Studies in West Circassian Phonology and Morphology*. Hakuchi Press, Leiden.
- Smith, Richard, and Connie Smith (1971). Southern Barasano phonemics. *Linguistics* 75, 80-85.
- Staalsen, Philip (1966). The phonemes of Iatmul. *Papers in New Guinea Linguistics* 5 (Linguistics Circle of Canberra, Series A, No.7), 69-76.
- Steriade, Donca (1995). Neutralization and the expression of contrast. Ms, UCLA.

- Steriade, Donca (1997). Phonetics in phonology: the case of laryngeal neutralization. Ms, UCLA.
- Steriade, Donca (2000). Paradigm uniformity and the phonetics-phonology boundary. Michael Broe and Janet Pierrehumbert (eds.) *Papers in Laboratory Phonology 5: Acquisition and the Lexicon*. Cambridge University Press, Cambridge, 313-334.
- Stevens, Kenneth N. (1999). *Acoustic Phonetics*. MIT Press, Cambridge.
- Stevens, Kenneth N., and Sheila E. Blumstein (1981). The search for invariant acoustic correlates of phonetic features. Peter D. Eimas and Joanne L. Miller (eds.) *Perspectives on the study of speech*. Lawrence Erlbaum, Hillsdale, 1-38.
- Stevens, Kenneth N., S. Jay Keyser, and Haruko Kawasaki (1986). Toward a phonetic and phonological theory of redundant features. Joseph S. Perkell and Dennis H. Klatt (eds.) *Invariance and Variability in Speech Processes*. Lawrence Erlbaum, Hillsdale, 426-449.
- Valente, V. (1975). Puglia. Manlio Cortelazzo (ed.) *Profilo dei dialetti italiani vol. 15*. Pacini, Pisa.
- Van Bergem, Dick R. (1994). A model of coarticulatory effects on the schwa. *Speech Communication* 14, 143-162.
- Walker, Rachel (1998). *Nasalization, Neutral Segments, and Opacity Effects*. Ph.D. dissertation, University of California, Santa Cruz.
- Walker, Rachel, and Geoffrey K. Pullum (1999). Possible and impossible segments. *Language* 75, 764-780.
- Westbury, John, and Patricia A. Keating (1980). Central representation of vowel duration. *Journal of the Acoustical Society of America* 67, S37A.
- Westbury, John, and Patricia A. Keating (1986). On the naturalness of stop consonant voicing. *Journal of Linguistics* 22, 145-66.
- Wright, James T. (1986). The behavior of nasalized vowels in the perceptual vowel space. John J. Ohala and Jeri J. Jaeger (eds.) *Experimental Phonology*. Academic Press, Orlando, 45-68.
- Zipf, George K. (1949). *Human Behavior and the Principle of Least Effort*. Addison-Wesley, Cambridge.