The role of distinctiveness constraints in phonology

Edward Flemming
flemming@mit.edu

1. Introduction

Distinctiveness constraints favour maximization of the perceptual differences between contrasting sounds. Perceptually indistinct contrasts are disfavored because they increase the likelihood of confusion on the part of listeners. Distinctiveness constraints are required to account for the preference for more distinct contrasts that can be observed in the typology of segment inventories (Flemming 2004). For example, there is a cross-linguistic preference for front unrounded and back rounded vowels because these yield more distinct contrasts in second formant frequency than front rounded or back unrounded vowels. Distinctiveness constraints also give rise to positional neutralization effects where a contrast is neutralized in environments where it would not be sufficiently distinct. For example, stop voicing contrasts are neutralized in final position where crucial Voice Onset Time cues are not available (Steriade 1997).

However, if distinctiveness constraints interacted freely with other phonological constraints, we would expect to find other effects that are in fact unattested. For example, articulatory markedness constraints could motivate significant contextual variation in inventories of contrasts to maximize the number and distinctiveness of contrasts in the face of context-specific restrictions. Attested contextual variation in inventories is generally limited to neutralization and allophonic variation. We would also expect to find 'contextual enhancement' as a counterpart to contextual neutralization: enhancement of contrasts precisely where they would otherwise be neutralized due to insufficient distinctiveness. This phenomenon is not attested in a general form.

These limitations can be understood as indicating that distinctiveness of contrasts is not the only requirement imposed by speech perception, it is also desirable for the system of contrasts to be as consistent as possible across contexts. This minimizes the need for context-sensitive adjustment of perceptual criteria in categorizing speech sounds. These ideas are implemented in terms of a model according to which the basic role of distinctiveness constraints lies in deriving an inventory of contrasting segments which serves as the 'alphabet' from which underlying forms are constructed, much like a phoneme inventory. This process is the locus of most enhancement effects. The distinctiveness constraints evaluate contrasts between words only to check that the contrasts are adequately realized on the surface - if not, they are neutralized (giving rise to positional neutralization). However, distinctiveness constraints play no other role in the mapping from underlying to surface form. That mapping is governed by constraints requiring faithful realization of the underlying contrasts in conflict with markedness constraints (articulatory constraints, metrical constraints, etc). The faithfulness constraints favour consistent realization of the inventory of contrasts in all contexts.

The organization of the paper is as follows: sections 2-4 review the evidence for distinctiveness constraints from enhancement and positional neutralization and provide a formalization of the constraints. The evidence that distinctiveness constraints do not interact with contextual markedness constraints is discussed in section 5. These observations provide the initial motivation for the inventory-based model of contrast. Further predictions of the model are discussed in sections 6 and 7.

2. Evidence for distinctiveness constraints

The nature of distinctiveness constraints can be illustrated through an example: the typological covariation of backness and rounding in vowels. Cross-linguistically, non-low vowels are generally front and unrounded or back and rounded, as in the canonical five vowel inventory in (1).

(1)
```
i     u
e     o
   a
```

There is a straightforward perceptual explanation for this correlation based on a preference to maximize the distinctiveness of contrasts. Front and back vowels differ primarily in the frequency of the second formant (F2): front vowels have high F2 while back vowels have low F2. Rounding the lips 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 explanation implies that a language with a vowel system like (1) excludes front rounded and back unrounded vowels because the contrasts involving these vowels would be insufficiently distinct. This analysis is implemented in terms of a ranked set of constraints along the lines shown in (3) (more general formulations are introduced below). *X-Y means that words should not be minimally differentiated by the contrast between X and Y. The less distinct the contrast, the higher-ranked constraints it violates, so the optimal contrast is the most distinct.

(2)
```
  i   y  ɨ  ɯ     u
◄─────────────────
        F2
```

(3) *i-ɨ >> *i-ɯ, *y-u >> *i-u

If there were no distinctiveness constraints, the covariation of backness and rounding would have to be analyze in terms of a fixed ranking of markedness constraints as in (4).

(4) *[front, +round], *[back, -round], *[central] >> *[-back, -round], *[+back, +round]

These two analyses make rather different predictions. The segment markedness constraints imply that vowels like central [] are inherently marked whereas the distinctiveness constraints imply that it is only constraints like [i-ɨ] that are problematic because they are less than maximally distinct. If it is the contrasts that are marked then the preference for front unrounded and back rounded vowels should not apply in the absence of a backness contrast. As shown in Flemming (2004), this prediction of the distinctiveness constraint analysis is correct: although vowels with non-peripheral F2 values (ɨ, ʉ, ɯ, etc) are relatively uncommon in front-back contrasts, they are usual in the absence of such contrasts. A number of languages, including Kabardian (Kuipers 1960, Choi 1991) and Marshallese (Bender 1968, Choi 1992) have short vowel inventories that lack front-back contrasts (‘vertical’ vowel inventories). These inventories consist of vowels whose backness and rounding is conditioned by surrounding consonants, resulting in vowels that vary contextually around central qualities [ɨ, ə] (Kabardian) or [ɨ, ə, a]

(Marshallese). There are no vertical inventories containing invariant [i] or [u], so there are ni inventories such as [i, e, a] or [u, o, a] although these are otherwise very common vowel qualities cross-linguistically.

These generalizations indicate that the markedness of vowels is depends on the system of contrasts: central vowels are marked in front-back contrasts and unmarked in their absence. This sensitivity to contrastive status cannot be accounted for in terms of segment markedness constraints. The ranking in (4) that would be required to account for the pattern of preferences observed with front-back contrasts incorrectly implies that if only one of these vowels appears it should be a front unrounded vowel or a back rounded vowel, not a central vowel. On the other hand, contrast-sensitive markedness is expected given distinctiveness constraints. In the absence of F2-based contrasts, distinctiveness in F2 is irrelevant, so other constraints goern backness and rounding. In this case minimization of effort becomes the key factor. The least effort vowel plausibly involves a smooth transition between preceding and following context, which results in contextually-variable, centralized vowel qualities. The same pattern is observed in English vowel reduction: when all vowel qualities are neutralized in unstressed syllables as in English, the result ('schwa') is a vowel that varies contextually around a high central quality (Flemming 2004, 2005).

This and other examples discussed in Flemming (2004) provide evidence that phonology includes distinctiveness constraints. These constraints are transderivational in the sense that they make wellformedness of one word depend on its relations to the surface forms of independent possible words. Output-Output Correspondence constraints (Benua 1997, Kenstowicz 1996, Steriade 2000) and Anti-Homophony constraints (Crosswhite 1999) carry the same implications. All transderivational constraints raise questions about the nature and size of the set of forms that need to be compared in evaluation. These questions are particularly acute in the case of distinctiveness constraints since all possible words can be linked by chains of minimal contrasts, so there is no immediately obvious bound on the size of the comparison set. Accordingly it is important to ascertain the extent of the effects of distinctiveness constraints. To this end we will survey the range of attested effects, and identify some significant limitations on them. Before turning to this task it is necessary to provide a more precise formulation of the distinctiveness constraints.

3. The Dispersion Theory of Contrast

To formulate distinctiveness constraints, it is necessary to have representations that make it possible to measure distinctiveness. The approach adopted here is to represent sounds as being located in a multi-dimensional perceptual space where perceptual distinctiveness corresponds to distance between sounds. For example, the perceptual dimensions of the vowel quality space correspond to the frequencies of the first two or three formants (Plomp 1975, Shepard 1972, Rosner and Pickering 1994). The space of possible vowels on the F1 and F2 dimensions is shown in (5), using a coarse quantization of the dimensions. The perceptual specification of a sound is its values on the various dimensions, e.g. [i] is [F1 *1,* F2 *5*].

(5)

F2	5	4	3	2	1		
	i	y	ɨ	ʉ	u	1	
	ɪ	ʏ			ʊ	2	
	e̝	ø̝			o̝	3	F1
	e	ø	ə		o	4	
		ɛ			ɔ	5	
		æ				6	
			a	ɑ		7	

The preference to maximize distinctiveness is implemented in terms of a ranked set of distinctiveness constraints requiring a specified minimum perceptual distance between contrasting sounds as in (6). For example, MINDIST = F2:4 requires a difference of 4 on the F2 dimension, which is only satisfied by [i] vs. [u]

(6) MINDIST = F2:1 >> MINDIST = F2:2 >>... >> MINDIST = F2:4

Contrasts are not in general maximally distinct so some constraints must conflict with the MINDIST constraints. A very general conflicting constraint is the preference to maximize the number of contrasting sounds. This serves to maximize the information conveyed by each sound. For example, if a language contrasts three vowels then uttering a vowel can distinguish up to three words, whereas if it contrasts seven vowels, then a single vowel can distinguish up to seven words. This preference is implemented as a positive constraint, MAXIMIZE CONTRASTS, that selects the largest inventory of contrasts.

The language-specific balance between these two kinds of constraints is determined by their ranking as illustrated in (X). This tableau shows a ranking that derives a contrast between front unrounded and back rounded vowels. Since the MINDIST constraints evaluate the distinctiveness of contrasts, the candidates are sets of contrasting vowels. For simplicity, we only consider contrasts along the F2 dimension. Candidate (a) wins because it satisfies both MINDIST constraints. Any other combination of backness and rounding in the vowels yields a less distinct contrast, and trying to better satisfy MAXIMIZE CONTRASTS by distinguishing three vowels (candidate e) results in insufficiently distinct contrasts – the contrasts [i-ɨ] and [ɨ-u] are only separated by 2 on the F2 dimension.

(7)

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

So one of the basic effects of MINDIST constraints is dispersion of contrasts (cf. Lindblom and Engstrand 1989), that is a tendency to spread contrasting sounds evenly over as large a perceptual space as possible. Dispersion is closely related to the notion of enhancement introduced by Stevens, Keyser and Kawasaki (1986). Enhancement essentially involves the combination of independent articulations to yield a more distinct contrast, so in these terms lip rounding is said to enhance the backness contrast in the winning inventory in (7). MINDIST constraints provide a general account of enhancement effects, and of the fact that enhancement only applies to contrasts. As discussed above, rounding is only used to enhance backness constrasts, and the same contrast-dependency applies to other examples of enhancement discussed in Stevens et al (1986) and Flemming (2004).

For example pre-nasalization of voiced stops enhances voicing contrasts (Stevens et al 1986). It is normally difficult to sustain strong voicing during a stop because build up of pressure in the oral cavity makes it difficult to sustain a sufficient pressure drop across the glottis to generate vocal fold vibration. Lowering the velum during the stop closure allows airflow through the nose, so oral pressure rises much less, facilitating voicing. Prenasalization is observed as an enhancement of voicing in languages like Fijian where prenasalized [^{m}b, ^{n}d, ^{ŋ}g] contrast with voiceless [p, t, k] (Schütz 1985), but it never applies to non-contrastive voicing, e.g. we do not find pre-nasalization of intervocalically voiced stops (Flemming 2004).

4. Distinctiveness constraints apply to the surface realizations of contrasts in context

The examples of enhancement in the previous section can be seen as operating at the level of the segment inventory in the sense that they are essentially context-free effects that serve to define the inventory of contrasting sounds in a language. So it might be thought that distinctiveness constraints apply only to phoneme inventories, assuming that phoneme inventories can be given some status in an Optimality Theoretic context, but do not actually evaluate contrasts between complete words. We will argue that inventories of basic contrasts do have a theoretical status in phonology and that distinctiveness constraints do play a central role in shaping these inventories, but they must also evaluate the surface realizations of contrasts in context because they play a central role in accounting for restrictions on the distribution of contrasts. The evidence comes from the phenomenon of positional neutralization (Steriade 1995, 1997, 1999). As Steriade observes, different types of contrasts have different characteristic contexts of neutralization. For example, in many languages, obstruent voicing contrasts are only permitted before sonorants, that is they are neutralized word-finally and before obstruents (e.g. German, Russian, Sanskrit). Major place contrasts (labial vs. coronal vs. dorsal) are often restricted to pre-vocalic position (e.g. Japanese, Luganda, Selayarese). On the other hand, contrasts between retroflex and apical alveolar consonants are often restricted to post-vocalic position, being neutralized word-initially and after consonants (e.g. Gooniyandi, Miriwung, Walmatjari).

Steriade argues that the generalization that unifies these diverse patterns of neutralization is that contrasts are neutralized first in environments where ‘the cues to the relevant contrast would be diminished’ (Steriade 1997). This generalization receives a natural formulation in terms of Dispersion Theory: contrasts are neutralized in contexts where they cannot satisfy a MINDIST constraints that ranks above MAXIMIZE CONTRASTS. This line of analysis can be illustrated with respect to neutralization of obstruent voicing, reformulating part of the analysis of Steriade (1997). In this case the key perceptual dimension distinguishing voiced and voiceless obstruents in pre-sonorant position is Voice Onset Time (VOT); voiced stops have short VOT ([VOT *0*]) while voiceless stops have longer VOT ([VOT *1*])[1]. However, in final position all stops lack VOT specifications because there is no onset of voicing after the stop and hence no VOT.

The tableau in (8) and (9) show a constraint ranking that derives final devoicing. (8) shows that in pre-vocalic context a VOT difference can be realized, satisfying MINDIST = VOT:1, so the contrast is realized. But in final position (9), there is no VOT difference between voiced and voiceless so a contrast would violate the MINDIST constraint. Accordingly the contrast is neutralized. The neutralized stop is voiceless due to a low-ranked constraint against voiced osbtruents.

(8)

	_V	MINDIST = VOT:1	MAXIMIZE CONTRASTS	*[+voice, -son]
a.	☞ dV-tV		✓✓	*
b.	dV		✓!	*
c.	tV		✓!	

(9)

	V_#	MINDIST = VOT:1	MAXIMIZE CONTRASTS	*[+voice, -son]
a.	Vd#-Vt#	*!	✓✓	*
b.	Vd#		✓	*!
c.	☞ Vt#		✓	

This example shows that distinctiveness constraints must evaluate contrasts in context, based on their surface realizations. That is, the surface lack of VOT in final position is crucial to the incidence of neutralization in that context. If MINDIST constraints only applied to segment inventories then they could not play a role in deriving patterns of positional neutralization.

On the basis of evidence of this kind, I concluded in earlier work that MINDIST and MAXIMIZE CONTRASTS are output constraints that interact freely with other markedness constraints to derive the well-formed words in a language (Flemming 1995, 2002, 2004). I will refer to this type of model as the 'free interaction' model. The free interaction model presents analytical difficulties because MINDIST constraints make the wellformedness of a word depend on whether it is adequately distinct from its neighbors. For example, the well-formedness of a candidate word [pad] might depend on whether or not [pat] is also a possible word. But to determine whether [pat] is a possible word, we have to determine whether it satisfies MINDIST constraints requiring it to be adequately distinct from its neighbors, and so on. It becomes unclear how to evaluate the wellformedness of a single word without effectively determining the entire set of grammatical words. This might be possible, but it certainly isn't easy. In any case we will see that the free interaction model predicts unattested phenomena, in particular contextual reorganization of inventories, and positional enhancement as a counterpart to positional neutralization. The absence of these phenomena indicates that distinctiveness constraints play a more restricted role in phonology. We will develop an alternative model that accounts for the observed limitations and that appears to be more tractable.

[1] The VOT of voiceless stops can vary, most notably between aspirated and unaspirated stops, so further levels of VOT need to be differentiated, but all that matters for present purposes is that voiceless stops have longer VOT than voiced stops.

5. Languages have segment inventories

If MINDIST and MAXIMIZE CONTRASTS constraints interact freely with contextual markedness constraints to derive well-formed words, then contrasts should be optimized for their context. In each context, a language should select the maximum number of sufficiently distinct contrasts, given restrictions imposed by contextual constraints. Accordingly, we would expect to find substantial contextual variation in segmental contrasts, making any notion that a language has a single, coherent inventory of contrasts a poor approximation to reality. In fact languages seem to be well characterized as adopting an inventory of contrasts. Of course there is contextual neutralization of contrasts and allophonic variation in their realization, but we do not find more radical restructuring of inventories according to context.

An example of the kind of contextual variation in vowel inventories that can be derived in the free interaction model is a language that allows front rounded vowels in most contexts, but adjacent to labials front rounded vowels are excluded and central vowels appear in their place. This hypothetical pattern is one in which central vowels appear to be allophones of front rounded vowels, conditioned by labial consonants. While labials can affect rounding of adjacent vowels they never condition unrounding and retraction of front rounded vowels. To see why this unattested pattern is predicted by the free interaction model it is helpful to consider a related attested pattern in Cantonese. Cantonese has front rounded vowels in its inventory (10), but does not allow these vowels to occur adjacent to labials (Kao 1971, Yip 1988), so sequences such as *pø, *my, *fø, *yp, *øp are excluded. But in this case the front rounding contrast is simply neutralized adjacent to labials, no other contrast takes it place.

(10) Cantonese, Finnish

i	y	u
e	ø	o
	a	

If MAXIMIZE CONTRASTS and MINDIST constraints freely interact with contextual markedness constraints of the kind that exclude front rounded vowels adjacent to labials then we predict the existence of the unattested variant of Cantonese because MINDIST constraints makes the markedness of vowels depend on what other vowels they contrast with. In this case, the absence of central vowels in Cantonese is explained in part by the fact that front rounded and central vowels are acoustically similar, so the contrast between them (y-ɨ) would be less distinct than contrasts between front rounded and unrounded (i-y), or front and back rounded (y-u). However, if front rounded vowels are excluded from a particular context, adjacent to labials, then the problem of contrasting front rounded and central vowels does not arise, and central vowels may be able to emerge, as long as they are adequately distinct from front unrounded and back rounded vowels. MAXIMIZE CONTRASTS favors realizing central vowels if they satisfy higher-ranked distinctiveness constraints. The resulting inventory (11) is attested as the general vowel inventory of a language, e.g. in Romanian, so it is undoubtedly a viable set of contrasts, but it is not found in complementary distribution with an inventory like (10). This analysis is developed more formally below.

(11) Romanian:

i	ɨ	u
e	ə	o
	a	

The essence of the problematic prediction made by the interactive model is that the inventories of contrast found in different positions can be quite independent. That is because the dispersion constraints mandate realizing the maximum number of adequately distinct contrasts in every context. Combined with contextual markedness constraints, this can lead to the selection of substantially different inventories in different contexts, as in the example just discussed. Actual languages do not behave in this way – they actually behave as if they have coherent inventories of contrasts that are only modified according to context in two basic ways. First, a subset of the inventory may appear in a given context. There is generally a basic division between the contexts where vowels and consonants may appear, but within these classes contrasts can also be neutralized in a particular context. This is the actual response to the restriction against front rounded vowels occuring adjacent to labials in Cantonese: only front unrounded and back rounded vowels contrast in this context. Second, the precise realization of the contrasts may vary according to context (allophonic variation), for example vowels are often nasalized preceding nasal consonants without neutralizing any vowel quality contrasts. We do not observe more radical restructuring of a language's inventory to maximize the number or distinctiveness of contrasts in each context, contrary to the predictions of a model in which MINDIST and MAXIMIZE CONTRASTS interact freely with contextual markedness constraints.

Another class of unattested phenomena that is predicted if MINDIST constraints and MAXIMIZE CONTRASTS interact with contextual markedness constraints involve 'push chains'. A markedness constraint that affects the realization of one vowel should be able to trigger a chain shift, rearranging the vowels to maintain distinct contrasts, motivated by MINDIST constraints. For example laxing of high vowels is attested in Canadian French, where it applies mainly in closed syllables (Walker 1984) and in Cantonese, where it applies before velars (Kao 1971). Laxing of high vowel shifts them closer to mid vowels, so MINDIST constraints could motivate a lowering chain shift to maintain the distinctiveness of height contrasts. Given a vowel inventory as in (12a) in open syllables, and a requirement to satisfy MINDIST = F1:2, laxing of high vowels could trigger lowering of mid vowels and elimination of the low central vowel (12b) in order to maximize the number of contrasts while satisfying MINDIST. We never observe this kind of reorganization of a vowel system in response to a change that is conditioned in only one or two vowels.

(12)	(a)	i		u	(b)	ɪ	ʊ
		e̞		o̞		e	o
		ɛ		ɔ		æ	ɒ
			a				

F2: 5	4	3	2	1	F1
i	y	ɨ	ʉ	u	1
ɪ	ʏ			ʊ	2
e̞	ø̞			o̞	3
e	ø	ə		o	4
	ɛ			ɔ	5
	æ			ɒ	6
		a	ɑ		7

These observations do not indicate a problem with distinctiveness constraints per se. As seen in the previous sections, distinctiveness constraints are essential for the analysis of dispersion effects and the typology of positional neutralization, but the non-existence of patterns of the kinds just reviewed shows that the interaction of distinctiveness constraints with contextual markedness constraints are limited. I will argue that the proper account of the role of distinctiveness constraints gives a central role to the segment inventory. The idea is that a language has an inventory of basic contrasting sounds that are used to construct words, similar to a phoneme inventory. Distinctiveness constraints and MAXIMIZE CONTRASTS constrain inventory selection. Distinctiveness constraints also evaluate the surface realization of contrasts to check that candidate contrasts are adequately distinct – this is necessary to account for the observation that contrasts are liable to be neutralized in contexts where they would be insufficiently distinct (Steriade 1997, 2001). However we will argue that their interaction with other constraints is severely limited. In particular they do not interact directly with positional markedness constraints and thus cannot motivate reorganization of the inventory of contrasts, they can only motivate neutralization of contrasts.

Specifically these proposals are implemented by dividing phonology into three subcomponents: Inventory, Realization and Evaluation of Surface Contrasts. The basic division of labor is that the Inventory component derives specifications for a set of contrasting segments, the Realization component maps strings of segments drawn from the inventory onto their phonetic realizations, while the Evaluation of Surface Contrasts assesses the distinctiveness of candidate contrasts as realized in context and forces the neutralization of those that would be insufficiently distinct. We will see that these components are distinguished because there is limited interaction between them, but they are not organized in a serial derivation. This model will be illustrated with reference to an analysis of the front rounding contrasts adjacent to labials as in Cantonese. This will allow us to see how the model avoids the problematic predictions of a free interaction model with respect to context-dependent reorganization of segment inventories.

The inventory of contrasts is not an arbitrary imposition on the phonology, it actually reflects important perceptual considerations that are not captured by distinctiveness constraints alone. An inventory can be thought of as a division of the perceptual space of speech sounds into categories. If the system of contrasts varies according to context, then the boundaries between categories have to be altered according to context. This makes accurate perception dependent on correctly identifying context. For example, in the hypothetical language described above with front rounded vowels in complementary distribution with central vowels, listeners do not know whether to try to categorize a vowel as front unrounded vs. back rounded or as central vs. back unless they know whether the preceding or following consonant is a labial. As a result perceptual errors can propagate: misperception of consonant place makes misperception of vowels more likely. So it is desirable for the set of contrasting sounds to remain constant across contexts. In the proposed model, the inventory represents the preferred perceptual sound categories, and faithfulness constraints favor the consistent realization of these sounds in all contexts.

The Inventory component selects a basic inventory of contrasting segment types. It operates in terms of a ranking of MINDIST, MAXIMIZE CONTRASTS and segment-internal articulatory markedness constraints, but no contextual markedness constraints. The contrasting segments are specified in terms of perceptual targets. It is important for the analysis of contextual neutralization that these targets specify cues for distinguishing one segment from another, so if no contrasts are realized on a particular dimension then sounds lack perceptual targets on that dimension. For example, if there are no voicing contrasts, then stops lack targets for voicing-related dimensions or if there are no F2 contrasts then vowels lack targets for F2. We will return to this point in section 7.

The tableau in (14) illustrates part of the derivation of the vowel inventory of Cantonese, specifically the backness and rounding contrasts. Two dimensions are relevant here, F2 and F3. Front and back vowels primarily differ in F2 whereas the primary difference between front unrounded and rounded vowels is in F3. Front unrounded vowels have high F3 whereas both back and front unrounded vowels have low F3, and central vowels have intermediate F3 (13). To analyze contrasts on these two dimensions, MINDIST constraints must specify what constitutes a sufficient distance on each dimension, so MINDIST=F2:2 OR F3:2 is satisfied by a distance of 2 on either dimension.

Given the ranking shown in (14), we derive contrasts between front unrounded, front rounded and back rounded vowels (candidate b). Candidates like (e) with fewer backness/rounding contrasts are eliminated because MAXIMIZE CONTRASTS is sufficiently highly ranked. Front rounded vowels are preferred over central vowels because the contrast [y-u] is more distinct than [ɨ-u] (F2:3 vs. F2:2), while [i-y] and [i-ɨ] are comparably distinct (both satisfy MINDIST= F2:2 or F3:3). It is not possible to contrast central and front rounded vowels because these vowels only differ by 1 unit on the F2 and F3 dimensions, violating the top-ranked MINDIST constraint.

(13)

F2:	5	4	3	2	1
	i	y	ɨ	ʉ	u

F3:	4	3	2	1
	i	ɨ	y, u	ɹ

(14)

		MINDIST= F2:2 or F3:2	MAXIMIZE CONTRASTS	MINDIST= F2:3 or F3:3	MINDIST= F2:4
a.	i ɨ u		✓✓✓	**!	**
b.	☞ i y u		✓✓✓	*	**
c.	i y ɨ u	*!	✓✓✓✓	***	*****
d.	i y ɯ u	*!	✓✓✓✓	***	*****
e.	i u		✓✓!		

Given an inventory of contrasting segments, the goal is to realize the full inventory of contrasts in all contexts. So possible underlying forms consist of all sequences of segments from the inventory and the goal is to realize these underlying forms faithfully. In some cases underlying contrasts are neutralized because they cannot be realized with sufficient distinctiveness in a particular context, as in final neutralization of obstruent voicing contrasts. Following Flemming (2002), we will analyze the restriction on front rounded vowels in Cantonese in similar terms. Coarticulation with an adjacent labial renders [i] too similar to front rounded [y], so the contrast is neutralized in this context. In both cases the evaluation of distinctiveness must apply to the surface realizations of the contrasts in order to take contextual effects such as labial coarticulation into account. However, MINDIST and MAXIMIZE CONTRASTS must not interact freely with contextual markedness constraints if we are to account for the relative stability of inventories across contexts. Reconciling these two generalizations motivates the distinction between Realization and Evaluation of Surface Contrasts (ESC). The Realization component maps an input string onto its phonetic realization while the ESC assesses the distinctiveness of contrasts based on these phonetic realizations. The Realization component incorporates the contextual markedness constraints that motivate contextual variation in the realization of contrasts but does not include MINDIST constraints. The MINDIST constraints evaluate the outputs of Realization in ESC but cannot directly influence the realization of a given input.

The application of this model to neutralization of front rounding contrasts is illustrated by the tableaux in (15) and (16). It is easier to se ethe overall structure of the analysis by considering ESC first and then turn to the details of Realization.

To evaluate the distinctiveness of contrasts it is necessary to consider a target input in relation to a set of minimally contrasting inputs. This set must at least contain all input forms that differ from the target by changing, inserting or deleting a single segment, but generally must contain additional forms as will be discussed further below. Here we are interested in the input /pyn/ with a front rounded vowel adjacent to a labial. The contrast set for this input includes inputs that differ in vowel quality, /pin, pun/. The candidates in ESC specify the fates of the members of the contrast se, as illustrated in tableau (X). Each input form from the contrast set may be realized or neutralized with a neighboring form. For example, /pin/ and /pyn/ can remain distinct, or /pyn/ can be mapped onto /pin/. Neutralizations are indicated in the tableau by showing in each candidate which input forms are distinguished, with subscripts indicating which forms are merged (if any). The tableau shows three candidate realizations of the contrast set. In (a) all members remain distinct while in (b) and (c) the contrast between /pin/ and /pyn/ is neutralized. In candidate (b) both are mapped onto /pin/ and in (c) both are mapped onto /pyn/.

The constraint *MERGE penalizes neutralizations (i.e. mergers of input forms), but neutralization can be forced if a contrast would violate a higher ranked MINDIST constraint. The MINDIST constraints evaluate the distinctiveness of the contrasts between surface forms, shown in the second row of each candidate. The surface forms are supplied by the Realization component, so evaluating the candidates in ESC involves referring to the Realization to derive the surface forms in each candidate. The realizations are determined in a separate optimization for each input form so the distinctiveness constraints cannot influence realization, they can only evaluate its results. Consequently the only possible response to an indistinct contrast is to neutralize it – the inventory cannot be rearranged in more radical ways because this would require altering the results of the Realization.

(15)

F3:	4	3	2	1
	i	ɨ	y, u	ɹ
		i$^{\beta}$	ɨ$^{\beta}$	

F2:	5	4	3	2	1
	i	y	ɨ	ʉ	u

(16)

	/pyn$_1$, pin$_2$, pun$_3$/	MINDIST= F2:2 or F3:2	*MERGE	MINDIST= F2:3 or F3:3	MINDIST= F2:4
a.	/pyn$_1$, pin$_2$, pun$_3$/ pyn$_1$ pi$^{\beta}$n$_2$ pun$_3$	*!		*	**
b.	☞ /pin$_{1,2}$, pun$_3$/ pi$^{\beta}$n$_{1,2}$ pun$_3$		*		
c.	/pyn$_{1,2}$, pun$_3$/ pyn$_{1,2}$ pun$_3$		*		*!

The Realization component maps a string of segments drawn from the inventory onto its phonetic realization. This includes the assignment of stress, syllabification, gestural overlap, coarticulatory effects etc. The mapping is based on a ranking of correspondence constraints that require faithful realization of the perceptual targets of the input segments and markedness constraints, such as articulatory effort constraints and metrical constraints. Distinctiveness constraints do not apply in Realization so they do not directly interact with contextual markedness constraints. In Realization, distinctiveness of contrasts is preserved via faithfulness to the targets specified in the inventory rather than by direct application of MINDIST constraints.

According to the proposed analysis of Cantonese, the contrast between front unrounded and rounded vowels is neutralized adjacent to labials because coarticulation with the labial makes the unrounded vowels too similar to their rounded counterparts. That is, unrounded vowels are produced with a constriction at the lips due to partial assimilation to an adjacent labial consonant. Labial constrictions have similar acoustic effects whether they are due to lip rounding or not, so labial coarticulation makes the front unrounded vowels sound similar to their rounded counterparts, and neutralization results. This labial coarticulation is derived in Realization.

For present purposes we adopt a rather specific constraint, LABIAL COARTICULATION, which requires a vowel adjacent to a labial to be produced with approximation of the lips or lip-rounding. This is presumably an instance of a more general constraint or family of constraints against rapid articulatory movements. This particular constraint penalizes movement directly from closed lips to fully open lips between adjacent segments, but is satisfied by the smaller movement from closed to partially open or rounded. The acoustic effects of an unrounded labial constriction are similar to lip rounding but lesser in degree since both are labial constrictions, but rounding involves protrusion of the lips as well as constriction. In general labial constriction and protrusion both lower the frequencies of all formants, but in front vowels the lowering effect on F3 is greatest since this formant is primarily associated with the cavity immediately behind the lips. So we will assume that unrounded labial constriction lowers the F3 value of front and central vowels (15)[2]. Vowels with an unrounded labial constriction are transcribed with a superscript [$^{\beta}$].

The tableau in (17) shows that LABIAL COARTICULATION outranks the correspondence constraints IDENT(F2) and IDENT(F3), which require faithful realization of input F2 and F3 specifications. So an unrounded vowel adjacent to a labial must be realized with labial constriction even though this results in deviation from its target F3. Consequently the realization of /pin/ is [pi$^{\beta}$n], and this is reflected in the candidates in ESC (16), repeated below as (18). The arrows indicate the relationship between candidate (a) in this tableau, and the Realization shown in (Y): candidate (a) includes an underlying form /pin/ which is realized as [pi$^{\beta}$n]. The mapping between UR and surface form is shown in the realization tableau (Y). Each UR shown in tableau (X) is mapped onto its corresponding surface form by a similar process. Rounded vowels adjacent to a labial do not violate LABIAL COARTICULATION and so are realized faithfully.

Tableau (18) shows that it is not possible to realize all the members of the contrast set (candidate a) because the contrast between [pyn] and [pi$^{\beta}$n] is insufficiently distinct since it involves differences of only 1 unit on the F2 and F3 dimensions, violating the top-ranked MINDIST constraint. Candidates (b) and (c) represent the two ways in which the offending contrast could be neutralized, preserving either the unrounded or the front rounded vowel. The unrounded vowel (candidate b) is preferred because it is more distinct from back [u], so the contrast [pi$^{\beta}$n - pun] satisfies MINDIST = F2:4.

[2] This analysis is simplified in a number of respects. It would be more realistic to analyze the result of labial coarticulation as a vowel in which lip aperture gradually increases through the vowel but that would require more detailed representations than we have adopted so far. In any case, the main point is that labial coarticulation shifts the acoustics of a front unrounded vowel in the direction of a front rounded vowel, and that remains true on either analysis. Also lip constriction should lower F2 as well as F3, but the effect on [i] is too small to be represented in the terms proposed here since we are assuming that [i] and [y] only differ by one unit in F2. The effect on a central vowel might be larger, but, as we will see, all that matters to the discussion below is that it is possible to produce a labially coarticulated vowel that has the specifications [F2 *3*, F3 *2*], whether this vowel is strictly central or slightly more fronted.

(17) Realization:

	/pin/	*LABIAL COARTICULATION	IDENT(F2)	IDENT(F3)
a.	pin	*!		
b.	☞ pi$^{\beta}$n			*
c.	pɨ$^{\beta}$n		**!	***

(18) ESC:

	/pyn$_1$, pin$_2$, pun$_3$/	MINDIST= F2:2 or F3:2	*MERGE	MINDIST= F2:3 or F3:3	MINDIST= F2:4
a.	/pyn$_1$, pin$_2$, pun$_3$/ pyn$_1$ pi$^{\beta}$n$_2$ pun$_3$	*!		*	**
b.	☞ /pin$_{1,2}$, pun$_3$/ pi$^{\beta}$n$_{1,2}$ pun$_3$		*		
c.	/pyn$_{1,2}$, pun$_3$/ pyn$_{1,2}$ pun$_3$		*		*!

The ESC employs the same ranking of MINDIST constraints as the Inventory – there is no re-ranking of constraints between components, although only certain classes of constraints apply in each component. The initial hypothesis is that *MERGE also occupies the same place in the ranking as MAXIMIZE CONTRASTS. This implies that the minimum level of distinctiveness that is acceptable for a contrast in the inventory should also be the threshold below which a surface contrast is neutralized. In the analysis above, the smallest acceptable F3 contrast must have a distinctiveness of F3:2 since MINDIST = F3:2 ranks just above MAXIMIZE CONTRASTS (14). If *MERGE occupies the same position as MAXIMIZE CONTRASTS in the ranking of MINDIST constraints, any contextual influence that would reduce the distinctiveness of an F3 contrast below F3:2 should result in neutralization of that contrast, as above. We will see some evidence that *MERGE may be allowed to rank higher than MAXIMIZE CONTRASTS which would mean that the level of distinctiveness required for a contrast to be included in the inventory is higher than the level that is demanded of surface contrasts. This allows some leeway for contrasts to fall below their canonical level of distinctiveness without being neutralized. This might well be necessary to account for cases in which contrasts are retained even in environments where their distinctiveness seems to be significantly reduced. Padgett and Tabain (2005) provide evidence that this is true of Russian vowel quality contrasts: unstressed vowels are less distinct in F1-F2 space than stressed vowels. Another possible case is discussed in section 8 below.

The inventory-based model avoids the problematic prediction that inventories should be highly contextually-variable. Taking the example of the preceding analysis, we can see that these constraints cannot derive the unattested variant pattern according to which central vowels are substituted for front rounded vowels adjacent to labials. Non-low central vowels are nor part of the basic inventory so they cannot appear in underlying representations. Realization will not map an input front rounded vowel onto a central vowel because that would involve a violation of faithfulness that is not motivated by an markedness constraint that is active in Realization (cf. 17). In ESC, violation of MINDIST constraints can only be avoided by neutralizing a contrast, not by reorganizing contrasts, so the contrast between [i] and [y] can be neutralized adjacent to labials if it would be insufficiently distinct, but it cannot be replaced by a contrast between [i] and [ɨ].

In the free interaction model, vowel contrasts after labials would be selected subject to the contextual markedness constraint LABIAL COARTICULATION. This would make it possible to derive the problematic pattern where central vowels can emerge wherever front rounded vowels are excluded. This is illustrated in (19). The same ranking that derives front rounded vowels in most contexts, as in (14), derives central vowels adjacent to labials because labial coarticulation makes the front rounding contrast untenable. Simply neutralizing the contrast (candidate d) is dispreferred because it yields fewer contrasts.

(19)

		LABIAL COARTIC.	MINDIST= F2:2 or F3:2	MAXIMIZE CONTRASTS	MINDIST= F2:3 or F3:3
a.	pi py pu	*!		✓✓✓	*
b.	pi$^{\beta}$ py pu		*!	✓✓✓	*
c.	!☞ pi$^{\beta}$ pɨ$^{\beta}$ pu			✓✓✓	**
d.	pi$^{\beta}$ pu			✓✓!	

1.1. A note on Richness of the Base

The proposed model involves a modification of the principle of 'Richness of the Base', adopted in most OT models of phonology. This principle states that all languages have the same set of inputs (Prince and Smolensky 2004), whereas it is proposed here that inputs are constrained to be constructed from a language-specific inventory of segments. This is a well-motivated and theoretically harmless modification to Richness of the Base. In standard OT, Richness of the Base is a corollary of the hypothesis that all cross-linguistic differences in phonology result from differences in constraint ranking. This hypothesis implies that there are no differences between languages in any other aspects of phonology, including the nature of the possible inputs. Positing inventories of input segments constitutes a relatively minor modification of this approach since the inventory is derived by a ranking of universal constraints and all languages draw on the same space of sounds in constructing an inventory, so it is still the case that phonologies differ only in the ranking of constraints.

1.2. The relationship between phonetics and phonology

Phonetic realization is usually conceived of as following phonology – indeed the name implies that it involves the mapping of phonological representations onto their phonetic form. However, there is now a variety of evidence that the distribution of contrasts is sensitive to details of realization, many of which are language-specific. For example, the viability of stop place contrasts in a given context depends on whether stops are audibly released in that context (Jun 2002), and the viability of contour tones on a given syllable type depends on the patterns of vowel and coda duration in that language (Zhang 2004). So in the model proposed here, the Realization component derives phonetically detailed representations including coarticulatory effects, gestural overlap, etc so that this information is available in ESC to affect the distribution of contrasts. The proposed organization of phonetics and phonology is modular in the sense that there is limited interaction between Realization and ESC: the ESC makes use of the outputs of Realization but cannot affect those outputs.

6. Limits on positional enhancement

Additional evidence for the hypothesized restrictions on the role of distinctiveness constraints comes from limits on positional enhancement. The phenomenon of enhancement is one of the basic pieces of evidence for distinctiveness constraints, as discussed in section 2. The cases of enhancement discussed there, enhancement of backness contrasts by rounding and enhancement of stop voicing contrasts by prenasalization, are analyzed in the same way in the inventory-based model: these enhancements are derived in the Inventory component. In fact, the analyses outlined in section 3 essentially involve deriving inventories only, so they are placed on a firmer footing by giving inventories a theoretical status. The interesting prediction of the inventory-based model is that enhancement effects should only arise via the Inventory. Enhancement is motivated by MINDIST constraints, and these apply in the Inventory. Although MINDIST constraints also apply in ESC, violations can only be avoided by neutralization, not enhancement. This is because candidates in ESC differ only in the number of basic contrasts that they neutralize. The realizations of the contrasting forms are supplied by the Realization component, and therefore cannot be affected by MINDIST constraints. Consequently the model predicts that there should be no positional enhancement as a counterpart to positional neutralization. The evidence bears this out.

The free interaction model predicts that there should be corresponding typologies of positional neutralization and positional enhancement. As discussed above, positional neutralization is motivated by distinctiveness constraints: contrasts are neutralized where they cannot satisfy a critical MINDIST constraint. But neutralization should be only one way to evade a potential violation of a MINDIST constraint, the alternative should be to make the contrast more distinct so it satisfies the constraint – that is, to enhance the contrast. So we would expect that environments where a contrast is more prone to neutralization should also be environments where the same kind of contrast is prone to enhancement. In fact the attested patterns of positional enhancement are rather limited.

Push chains of the kind discussed in section 5 could be regarded as a rather extreme form of positional enhancement in which the whole vowel inventory is rearranged in order to satisfy MINDIST constraints with as many contrasts as possible. But even more modest forms of positional enhancement turn out to be rather limited and best analyzed as consequences of faithfulness to perceptual targets rather than MINDIST constraints. Since the perceptual targets in the Inventory specify dispersed contrasts, faithfulness to these targets can produce the appearance of enhancement, but the attested patterns are not sensitive to the possibility of neutralization in particular contexts. Consequently the environments of ‘enhancement’ do not parallel the environments of positional neutralization. We consider two cases of apparent enhancement: common patterns of vowel epenthesis and the realization of laryngeal contrasts in Native American languages of the Northwest.

Vowel epenthesis provides a good example of a phenomenon that can be perceptually motivated in a broad sense, but is not precisely targetted towards salvaging contrasts. It has been argued that vowel epenthesis can be motivated by the perceptual requirements of consonants (Côté 2000, Wright 1996). For example, the common pattern in which medial clusters of three consonants are broken up vowel epenthesis, as in Yawelmani Yokuts (20) (Newman 1944, Kisseberth 1970), ensures that every consonant is adjacent to a vowel. Similar patterns are observed in Cairene Arabic (Broselow 1976) and Lenakel (Lynch 1978). As Côté observes, many significant cues to consonant contrasts can only be realized with an adjacent vowel, e.g. formant transitions provide cues to place of articulation and the change in intensity provides cues to consonant manner and presence.

(20)		/paʔt+mi/	→	[paʔitmi]	‘having fought’
	cf.	/paʔt+al/	→	[paʔtal]	‘might fight’
		/lihm+mi/	→	[lihimmi]	‘having run’
	cf.	/lihm+al/	→	[lihmal]	‘might run’

However it is not necessary to appeal to MINDIST constraints to account for this pattern of epenthesis. We will see that it is better accounted for in terms of faithfulness to consonant transitions specified in the inventory. If vowel epenthesis were directly motivated by MINDIST and MAXIMIZE CONTRASTS constraints then we would also expect patterns in which vowels are epenthesized precisely where they are needed to prevent neutralization of contrasts. For example, many languages neutralize obstruent voicing contrasts in word-final position. Steriade (1997) argues that voicing contrasts are prone to neutralization in this position because an important cue to voicing, Voice Onset Time (VOT), is unavailable there. In final position, there is no following voiced sound, and hence no VOT. However voicing contrasts in this position could be made more distinct by epenthesizing a following vowel, allowing for the realization of VOT cues, but this pattern of ‘positional enhancement’ is unattested. The only attested repair for the violation of MINDIST constraints in final position is neutralization of the contrast, devoicing all final obstruents (Lombardi 2001, Steriade 2001).

Another well-established pattern of positional neutralization that lacks parallel patterns of positional enhancement involves place assimilation in consonant clusters. Jun (1995, 2004), building on earlier work by Mohanon (1993), shows that patterns of major place assimilation in intervocalic consonant clusters exhibit a number of implicational universals. For example, if stops assimilate in place to a following consonant then so do nasals, but not vice versa. So there are languages in which both stops and nasals assimilate in place to a following stop, such as Korean (21), and there are languages in which nasals assimilate in place to a following stop, but oral stops do not assimilate, e.g. Malayalam (22), but there are no languages in which stops undergo assimilation whle nasals do not.

(21) Malayalam (Mohanon and Mohanon 1984)

Nasals undergo assimilation:

/peɳ-kuʈʈi/	peŋkuʈʈi	‘girl (female child)’	cf. peɳɳə	‘female’
/miin-tʃan̪t̪a/	miin̠tʃan̪t̪a	‘fish market’	cf. miin	‘fish’

Stops do not undergo assimilation:

ut̪karʂam	‘progress’
sapt̪am	‘eight’

(22) Korean (Jun 1995)

/mit-ko/	[mikko]	‘believe and’
/ip-ko/	[ikko]	‘wear and’
/cinan-pam/	[cinampam]	‘last night’
/nam-kɨk/	[naŋkɨk]	‘the South Pole’

Jun argues that this asymmetry arises because place contrasts between nasals that do not precede vowels are less distinct than place contrasts among non-prevocalic stops and thus are more prone to neutralization. It is well established that nasalization has a detrimental effect on the distinctiveness of vocalic elements because it reduces formant intensities and introduces additional formants, rendering the spectra of nasalized vowels less distinct than oral vowels (e.g. Wright 1980?) and it is plausible that nasalization should have the same effect on the distinctiveness of formant transitions. In the terms proposed here, this implies that non-prevocalic nasal place contrasts violate higher-ranked MINDIST constraints than equivalent stop contrasts, and so stop place contrasts can be preserved even where nasal place contrasts are ruled out as insufficiently distinct.

Again, we do not find any patterns of positional enhancement to parallel these generalizations about positional neutralization. For example, epenthesis could be used to rescue nasal place contrasts that would otherwise be neutralized. This would result in a pattern of epenthesis where vowels are inserted into nasal-stop clusters, but not into stop-stop clusters (23).

(23) /anba/ → [anəba] but: /atpa/ → [atpa]

This pattern is unattested. In fact, where there is a choice of positions for epenthesis, epenthesis into nasal-stop clusters seems to be strongly disfavored. For example, -CCC- clusters can be eliminated by inserting a vowel after the first consonant or after the second. The languages cited above consistently adopt one option or the other independent of the manner of consonants involved (cf. Itô 1989), but in Chaha the location of epenthesis depends on the relative sonority of consonants in the cluster (Rose 2002). Epenthesis preferentially applies to insert a vowel between consonants that show a rising sonority profile, such as stop-liquid sequences. Nasal-stop clusters involve a significant fall in sonority whereas stop-stop clusters involve a sonority plateau, so stop-stop clusters should be more likely to be split by epenthesis. A similar dispreference for epenthesis into nasal-stop clusters emerges from a survey of the sites of vowel epenthesis in loanword adaptation (Fleischhacker 2002).

The generaliztion that MINDIST and MAXIMIZE CONTRASTS do not motivate positional enhancement follows from the organization of the grammar: these constraints do not apply in the Realization component, they only apply in the Evaluation of Surface Contrasts where the only ‘repair’ available to avoid violations of MINDIST constraints is to neutralize the offending contrast. Epenthesis arises in the Realization component, so the apparent contextual enhancement of epenthesis in –CCC- clusters is motivated by Correspondence constraints. Before formulating this analysis it is necessary to consider some issues concerning the representation of consonant contrasts and the nature of correspondence constraints in Realization.

The target specifications of consonants are somewhat more complex than vowel targets because cues to consonantal contrasts are generally distributed in time and many of them are realized on adjacent segments. For example, place contrasts among stops are generally cued by formant transitions at the closure and release of the stop and by the quality of the burst at the release of the stop. The Inventory of contrasts is intended to be the maximal set of contrasts, so it is assumed that all these types of cues are available in the specification of consonants. So a voiceless stop may have targets for voicelessness during the closure and a long VOT following the stop release.

Not all of the targets for a segment are necessarily realized precisely, and some of them may not be realized at all. For example a transition is only realized where a consonant is adjacent to a vowel and the first stop in a sequence of two stops may lack an audible release burst if it is overlapped with the following stop as in English. Correspondence constraints in Realization require that the targets of input segments be realized in the output, so failure to realize a specified transition violates a Correspondence constraint.

We will represent transitions between consonants and vowels as ‘sub-segments’ distinct from the consonant constriction and from the vowel proper. These segments bear features specifying the quality of the transition, including formant transitions and VOT. Stops can contain another sub-segment representing the release burst, with features specifying the burst quality, so an intervocalic stop consists of four subsegments: closure transitions, closure, burst and release transitions (24, cf. Flemming 2002:23ff.)[3].

(24)

$$\underset{\textit{closure transition}}{\begin{bmatrix}\text{F1} & \mathit{1}\\ \text{F2} & \mathit{4}\\ \text{(etc)} & \\ & \end{bmatrix}}\underset{\textit{closure}}{\begin{bmatrix}\text{Loudness} & \mathit{0}\\ & \\ & \\ & \end{bmatrix}}\underset{\textit{burst}}{\begin{bmatrix}\text{NF} & \mathit{6}\\ \text{NL} & \mathit{2}\\ \text{(etc)} & \\ & \end{bmatrix}}\underset{\textit{release transition}}{\begin{bmatrix}\text{F1} & \mathit{1}\\ \text{F2} & \mathit{4}\\ \text{VOT} & \mathit{1}\\ \text{(etc)} & \end{bmatrix}}$$

Epenthesis into –CCC- clusters can then be motivated by the need to realize at least one transition segment for each consonant. This is formulated in terms of the Correspondence constraint MAX(trans) which is violated if neither of the specified transitions segments of a consonant are realized. This constraint effectively requires each consonant to be adjacent to a vowel since a transition segment can only be realized between a consonant and a vowel. This is similar to Côté’s (2000) constraint C⇔V ‘A consonant is adjacent to a vowel’, but is implemented in terms of faithfulness instead of markedness. Given the standard assumption that epenthesis of a vowel is penalized by the constraint DEPV, input CCC clusters will be broken up by epenthesis if MAX(trans) >> DEPV.

It is also necessary to posit a constraint MAX(release trans) that specifically requires the realization of release transitions to account for languages in which epenthesis eliminates all consonant clusters (e.g. Fijian). Release transitions are singled out because of their greater perceptual salience and importance (Fujimura et al 1978, Redford and Diehl 1999, Wright 2001) (cf. Côté's constraint C→V, requiring consonants to be followed by vowels).

In general universal restrictions on epenthesis then follow from fixed rankings of Correspondence constraints representing universal differences in the relative importance of cues (cf. Steriade 2001). In this case we can derive the observation that epenthesis is not directly motivated by preservation of voicing or place contrasts without stipulating fixed rankings if we assume that transition specifications for formant frequencies and VOT are protected by IDENT

[3] NF stands for ‘Noise Frequency’ and NL for ‘Noise Loudness’. See Flemming 2002:20-24.

constraints that require corresponding transition sub-segments to have identical specifications for particular features, e.g. IDENT(VOT), IDENT(F2) (cf. McCarthy and Prince 1995). Then epenthesis cannot be motivated by faithfulness to F2 transitions since failure to realize a transition only violates MAX(trans), not IDENT(F2). So ranking a specific IDENT constraint such as IDENT(F2) above DEPV will not result in epenthesis unless MAX(trans) also ranks above DEPV, otherwise it is preferable not to realize the transition, vacuously staisfying IDENT(F2). If MAX(trans) does outrank DEPV, then epenthesis is motivated wherever it is required to ensure that every consonant has at least one realized transition, not just where it is required to prevent place assimilation.

We illustrate this line of analysis with respect to final devoicing. It is assumed that voiced and voiceless stops have the targets shown in (25) for voicing and VOT. /T/ represents a stop that lacks voicing targets, as will be explained below. VOT specifications cannot be realized in word-final position, as indicated in (26), since there is no onset of voicing following the stop.

(25)

targets:	d	t	T
VOT	0	1	-
voice	1	0	-

(26)

realizations:	dV	tV	d#	t#
VOT	0	1	n/a	n/a
voice	1	0	1	0

The realization of stops in final position is shown in (X) and (Y). A final voiced stop is realized as a voiced stop without VOT. The VOT target could be realized via epenthesis of a following vowel (candidate b), but this violates the higher-ranked constraint DEPV. Voiceless stops are realized without epenthesis for the same reason.

The evaluation of surface contrasts for an input word-final voiced stop is shown in (X). The voicing contrast is insufficiently distinct because there is no VOT difference, violating MINDIST = VOT:1, so the contrast has to be neutralized. We see here the effect of an assumption introduced above that segments lack a target for a dimension if there is no contrast on that dimension. That is, the targets for segments in the inputs to realization are derived from the inventory specifications, but where there are neutralizations, some targets may be omitted. If a segment does not minimally contrast with any sound that differs on a given dimension, then there is no target on that dimension since targets only specify cues to contrasts (section 5). In candidate (b), there is no voicing contrast, so the stop lacks targets for [voice] and [VOT]. This unspecified stop is realized as voiceless due to a low-ranking effort constraint against voiced obstruents, *[+voice, -son], as shown in (29).

(27)	/ad/	DEPV	MAX(rel trans)	IDENT[VOT]	IDENT(voice)
a.	☞ ad		*		
b.	adə	*!			
c.	at		*		*

(28)

/at/	DEPV	MAX(rel trans)	IDENT[VOT]	IDENT(voice)
a. ☞ at		*		
b. atə	*!			
c. ad		*		*!

(29)

/aT/	DEPV	MAX(rel trans)	IDENT[VOT]	*[+voice, -son]
a. ☞ at		*		
b. atə	*!			
c. ad		*		*

(30) Evaluation of surface contrasts:

$/ad_1, at_2/$	MINDIST = VOT:1	*MERGE
a. $/ad_1\ at_2/$ ad at	*!	
b. $/aT_{1,2}/$ ☞ at		*

This analysis illustrates the fact that there is no positional enhancement in the inventory model because MINDIST and MAXIMIZE CONTRASTS constraints do not apply in Realization, so MAXIMIZE CONTRASTS cannot favour epenthesis here. In ESC, the only remedy for an indistinct contrast is neutralization.

The non-existence of epenthesis as a repair for final voicing contrasts is just one instance of what Steriade (2001) dubs the 'too many solutions' problem. If it is assumed that final devoicing is motivated by a constraint against word-final voiced obstruents, then we would expect a wider variety of repairs to eliminate final voiced stops, e.g. vowel epenthesis, consonant deletion, nasalization of the final stop, etc, but the only attested repair is devoicing. So far we have only addressed the non-existence of epenthesis as a repair for final voicing, but the distinction between Realization and ESC also helps us to understand the absence of some of the other unattested repairs. Final devoicing occurs because the contrast between final voiced and voiceless obstruents is insufficiently distinct, but this is determined in the ESC not in the Realization component. In Realization, the only concern is to remain faithful to the underlying specifications. As a result many of the unattested modifications of final voiced stops are excluded because they involve gratuitous violations of faithfulness unmotivated by any markedness constraint. For example, nasalization (/tab/ → [tam]), or final approximantization (/tab/ → [taw]) would constitute unmotivated faithfulness violations (or rather, they can only be motivated by *[+voice, -son], but this is not specific to final position). The only specification relevant to voicing that is at risk in Realization is VOT, so the only viable alternative candidate to be excluded is final epenthesis, as discussed above. In ESC, inadequately distinct contrasts cannot be repaired by modifying them – there is no positional enhancement – the only repair is neutralization. More generally, separating Realization from ESC predicts that the repairs to problems of insufficient distinctiveness will often be limited because repairs arise in Realization but distinctiveness constraints apply in ESC, so segmental modifications do not arise in direct response to distinctiveness constraints.

The analysis of the absence of epenthesis to prevent place assimilation is parallel. Following Jun (2004), place assimilation is analyzed as a side-effect of place neutralization in pre-consonantal position where crucial release cues are unavailable. So again neutralization is motivated by MINDIST constraints in ESC but epenthesis is only possible in Realization. In Realization the only problem is the difficulty of realizing release-based place cues such as formant transitions. Here the crucial assumption is that the faithfulness constraints that favor the realization of transitions are relatively general. That is, we posit constraints of the form MAX(trans) and MAX(release trans) that apply equally to all consonants, but no constraint that specifically requires the realization of release transitions of nasals. Consequently, a ranking such as MAX(release trans) >> DEP V that yields epenthesis after nasals will also derive epenthesis in all consonant clusters (cf. Fijian, Schütz 1985), so it is not possible to derive the unattested pattern where epenthesis breaks up nasal-stop clusters but not stop-stop clusters.

The proposed distinction between faithfulness to perceptual targets in Realization and MINDIST constraints in ESC makes it possible to capture the observation that processes like epenthesis can be perceptually motivated in the sense that they facilitate the realization of cues to contrasts, while also accounting for the fact that epenthesis is not specifically targeted to salvage contrasts that are at risk of neutralization. Epenthesis is perceptually motivated in the sense that it can be motivated by Correspondence constraints that require the realization of underlying perceptual targets, but it is not deployed as a precisely targeted positional enhancement because it is independent of MINDIST and *MERGE constraints.

7. Apparent positional enhancement: Hupa laryngeal contrasts

The phonology of Hupa laryngeal contrasts provides a somewhat different example of how apparen positional enhancement can arise in the proposed model. Contrasts derived in the Inventory are generally maximally enhanced since they are not subject to context-sensitive markedness constraints. However Realization is subject to context-sensitive effort constraints, metrical constraints, etc, so not all of the targets specified by the Inventory can be realized in every context. This is the basis for positional neutralization, as discussed above, but in some cases in can also give rise to the appearance of contextual enhancement because a contrast is realized via different cues in different contexts. However, these patterns arise from simple faithfulness to underlying specifications, not from MINDIST constraints.

A particularly striking case of this kind involves laryngeal contrasts in Hupa (Golla 1970, Gordon 2001). Hupa contrasts aspirated, voiceless unaspirated and ejective stops and affricates. We will focus on the realization of ejectives, but similar patterns of allophonic variation are observed with aspirated stops also. The basic pattern of interest here is summarized in (31). In prevocalic and final stops we find ejectives, but in pre-consonantal position the ejective's release is obscured by overlap with the following consonant. However, following a long vowel, the ejectives remain distinct from plain voiceless stops by virtue of the presence of creaky voicing at the end of the preceding vowel. In pre-consonantal position following a short vowel, the ejective is neutralized to a plain voiceless stop. These patterns are complicated by morphological factors analyzed in detail by Gordon (2001), but we will restrict our attention here to the phonological factors just outlined.

(31)	Patterns:		Examples:	
	after long V	after short V	after long V	after short V

pre-vocalic	Vːk’V	Vk’V	tʃ’eːkʲ’ɪɫ	tʰɪkʲ’ɪɫ
phrase final	Vːk’	Vk’	tʃeːkʲ’	tʰɪkʲ’
pre-consonantal	V̰ːk̚C	Vk̚C	tʃ’ḛːkʲ̚tʰe	tʰɪkʲ̚tʰeː

The pattern of variation in the realization of ejectives looks like positional enhancement in the sense that creaky voicing on the preceding vowel seems to be deployed to maintain the contrast where the ejective burst is unavailable. However I propose that creaky voice and ejective burst are both specified as cues to the contrast between ejective and plain stops in the Inventory, but the realization of these targets varies according to context as a result of interactions between Correspondence constraints in Realization.

We will not attempt to develop a general account of the perceptual dimensions relevant to ejective distinctions. All that is necessary for present purposes is to have features that differentiate ejective and plain stop bursts and modal and creaky voice quality. Ejectives are generally characterized by a high intensity burst compared to other types of stops due to high oral pressure (Kingston 1985), so we will adopt a ‘Burst Intensity’ (BI) dimension. For voice quality we adopt a binary feature [creak], although this could perhaps be replaced by a spectral tilt dimension (Ladefoged, Maddieson and Jackson 1988) (32).

(32)

targets:	closure transitions			burst	
	voiceless	ejective		voiceless	ejective
creak	0	1	BI	2	3

The proposed analysis closely follows Gordon (2001), although the formalization of the constraints is a little different. Gordon argues that two factors limit the realization of the burst and creak specifications of ejectives. First, stops are unreleased in pre-consonantal position. That is, the constrictions of adjacent consonants are overlapped so the release burst of a stop is obscured in pre-consonantal position. Second, vowels resist non-modal voicing. Gordon suggests that this is because the reduced intensity of non-modal vowels makes vowel quality less perceptible (p.19). Specifically, vowels that are creaky throughout their duration are unacceptable. Consequently creak can only be realized on a preceding long vowel since these have sufficient duration to accommodate an interval of creak and still be modally voiced for about half their duration. Creak cannot be realized on a preceding short vowel since it would be rendered fully creaky. However, vowels are ideally modally voiced throughout so long vowels also resist creaky voice. We will see that this has the result that creak is only realized where the ejective is unreleased.

Overlap between adjacent consonants is required by the constraint OVERLAP (CF. Cho 1997, Gordon 2001). The essential consequence of this constraint is that the bursts of pre-consonantal stops are unrealized. The resistance of vowels to creaky voice should be due to faithfulness to vowel quality targets, or possibly to formant intensity targets. To circumvent the question of how the effects of non-modal voicing on vowel quality should be represented, we will use faithfulness to modal voicing as surrogates for the proper correspondence constraints. We need to distinguish loss of modal voicing from part of the duration of a vowel from complete loss of modal voicing. This achieved via the two Correspondence constraints in (33). The formulation of these constraints represents an ad hoc attempt to formalize the intuition that it is a greater violation of faithfulness to change an input specification throughout the duration of a vowel than to change the specification for only part of a vowel. It is crucial that failing to realize [creak *0*] (modal voice) on a short vowel is a greater violation of faithfulness than failing to realize [creak *0*] on part of a long vowel, a generalization which is difficult to capture in terms of standard IDENT constraints. It is further assumed that an inviolable constraint prevents short vowels from bearing two specifications for [creak].

(33) IDENTV(creak): A [creak] specification on an input segment must be realized on the corresponding output segment.
IDENT-WHOLEV(creak): A [creak] specification on an input vowel must be realized on the entire duration of the corresponding output segment.

The interesting feature of the data is that we never see all of the ejective cues realized together (except as a result of interactions with Output-Output Correspondence constraints – Gordon 2001:56f.). The environment where we might expect this to be possible is where an ejective appears between vowels, the first of which is long, e.g. [tʃ’eːkʲ’ɪɬ] ‘be peppery (progressive)’, but in this context the ejective burst is realized, but there is no creaky voice on the preceding vowel. I suggest that this results from a kind of scalar evaluation of the overall degree of faithfulness to the perceptual specifications of an input segment, along the lines of Kirchner’s (1996) conception of distantial faithfulness. Essentially, failure to realize the creak specification is a lesser violation of faithfulness than failure to realize both the creak and release burst specifications. This can be formalized in terms of locally conjoined faithfulness constraints (Smolesnky 1996, Kirchner 1996). The two basic faithfulness constraints are IDENTC(creak), which requires preservation of the [creak] specifications of a consonant, and MAX(burst), which requires the realization of the ejective release burst. The conjoined constraint IDENTC(creak)& MAX(burst) is violated if neither cue is realized for a given consonant, i.e. if both basic constraints are violated with respect to a single consonant.

As illustrated in (34), ranking IDENT-WHOLEV(creak) above IDENTC(creak) blocks the realization of creak on a preceding vowel where the ejective burst can be realized. That is, realizing modal voice quality throughout a vowel is preferable to realizing the creak cue from an ejective consonant, other things being equal. The same applies to pre-pausal ejectives since the ejective burst can be realized in that context also ([tʃeːkʲ’], not *[tʃḛːkʲ’]).

(34)

	/Vːk’V/	OVERLAP	IDENTV (creak)	IDENTC(creak) & MAX(burst)	IDENT-WHOLE V(creak)	IDENTC (creak)	MAX (burst)
a.	☞ Vːk’V					*	
b.	V̰ːk’V				*!		

As shown in (35), other things are not equal in preconsonantal position because the high-ranking OVERLAP constraint prevents the realization of the ejective burst in this position (eliminating candidate a). If the preceding long vowel is fully modal (candidate c), then neither creak nor burst cues to the ejective are realized, in violation of the conjoined constraint. This ranks above IDENT-WHOLEV(creak) so it is preferable to realize creak on the latter half of the long vowel (candidate b).

(35)

/Vːk’C/	OVERLAP	IDENTV (creak)	IDENTC(creak) & MAX(burst)	IDENT-WHOLE V(creak)	IDENTC (creak)	MAX (burst)
a. Vːk’C	*!					
b. ☞ V̰ːk˺C				*		*
c. Vːk˺C			*!		*	*

Where the preceding vowel is short, realizing the creak cue to the ejective would make the whole vowel creakiy, violating top-ranked IDENTV(creak), so this is never possible. Failure to realize creak is unproblematic in pre-vocalic or pre-pausal position because the ejective burst can still be realized (36), but in pre-consonantal position following a short vowel neither release burst nor creak can be realized (37), so the realization is indistinguishable from a plain voiced stop.

(36)

/Vːk’C/	OVERLAP	IDENTV (creak)	IDENTC(creak) & MAX(burst)	IDENT-WHOLE V(creak)	IDENTC (creak)	MAX (burst)
a. ☞ Vk					*	
b. V̰k		*!				

(37)

/Vk’C/	OVERLAP	IDENTV (creak)	IDENTC(creak) & MAX(burst)	IDENT-WHOLE V(creak)	IDENTC (creak)	MAX (burst)
a. Vk’C	*!					
b. V̰k˺C		*!		*	*	*
c. ☞ Vk˺C			*			*

Support for this line of analysis comes from a comparable pattern of ejective realization in Takelma (Sapir 1912)[4]. This language contrasts voiceless unaspirated, aspirated and ejective stops. Word-finally and before consonants, all stops acquire an aspirated release, but ejectives remain distinct from the other types of stops by virtue of a preceding ‘glottal catch’ which Sapir transcribes as a glottal stop preceding the oral stop (p.36). This is reminiscent of Hupa in that ejectives are distinguished by preceding glottalization when the ejective release burst is lost, although the mechanism of loss is rather different. But in Takelma the preceding glottalization is a general feature of the realization of all post-vocalic ejectives, including intervocalic ejectives that are described as having strong release bursts. Sapir states that in the production of ejectives (‘fortes’ in his terminology), ‘the glottis is closed just before or simultaneously with the moment of consonant contact’, resulting in a preceding ‘glottal catch’ (pp.33f.). So loss of the burst is not a precondition for exploiting glottalization of the preceding vowel as a cue. In terms of the analysis proposed here, Takelma and Hupa share similar targets for the realization of ejectives, but Takelma differs from Hupa in the extent to which vowels resist glottalization.

So far we have only analyzed the Realization of ejectives across a full range of contexts. The outputs of Realization are still subject to Evaluation of Surface Contrasts. The cases of

[4] Thanks to Donca Steriade for bringing these data to my attention.

neutralization between ejectives and voiceless stops (after a short vowel, preceding a consonant) are relatively trivial in that the realizations of ejectives and voiceless stops in these contexts are identical so clearly violate all MINDIST constraints. However, the fact that surface contrasts based on release burst alone ([tʃeːkʲ’] vs. hypothetical [tʃeːkʲ]) or on glottalization of the preceding vowel alone ([tʃ’ḛːkʲ̚tʰe] vs. hypothetical [tʃ’eːkʲ̚tʰe]) are both acceptable introduces a slight complication. If we assume that MAXIMIZE CONTRASTS and *MERGE occupy the same position in the hierarchy of MINDIST constraints, then any acceptable surface contrast should be an acceptable contrast in the Inventory and vice versa, so we would expect preglottalization and ejection to be independent contrasts, maximizing the number of contrasts. The fact that these are not independent contrasts suggests that *MERGE is ranked somewhat higher than MAXIMIZE CONTRASTS so there is some leeway to reduce the distinctiveness of contrasts from the canonical realizations specified by the inventory before they have to be neutralized.

The proposed analysis does not really contradict the characterization of the Hupa pattern as positional enhancement – it is the case that creaky voice is realized because the ejective burst is not, but this pattern is derived from faithfulness to perceptual targets, not from distinctiveness constraints operating in context. The only kinds of positional enhancement that can be derived in this way involve the realization in different contexts of different subsets of the features specified for a segment in the Inventory. Since the set of features constitute the specifications of a segment, in general they should cooccur in surface realizations in some language, as is the case with creak and ejective burst in Takelma.

8. Conclusions

In this paper it has been argued that there is a preference for languages to maintain a consistent inventory of contrasting sounds in all contexts in addition to the preference to maximize the perceptual distance between contrasting sounds. Both preferences can be understood as having their basis in speech perception: maximally distinct contrasts minimize the likelihood of confusion by listeners, and maintaining a consistent inventory of contrasts reduces the need for context-dependent adjustments in perceptual criteria for catgeorization of speech sounds.

These ideas are implemented in a model according to which phonology is divided into three components, Inventory, Realization and Evaluation of Surface Contrasts. Distinctiveness constraints apply in the Inventory to derive the set of basic contrasts in interaction with MAXIMIZE CONTRASTS and segment-internal effort constraints. The distinctiveness constraints favour maximally distinct contrasts, resulting in dispersion and enhancement effects such as the covariation of backness and rounding in vowels. In Realization the goal is to faithfully realize the inventory of contrasts in all contexts, implementing the preference to maintain a consistent inventory of contrasts. However this goal conflicts with contextual markedness constraints, such as effort constraints and metrical constraints. Contrasts must be adequately distinct in context to be easily recoverable, so distinctiveness constraints also apply to the surface forms of words (ESC). Contrasts that would be insufficiently distinct are neutralized, as in final devoicing, but more radical reorganization of contrasts is not possible in response to context-specific difficulties.

Empirically, this model is motivated by the observation that context-free enhancement effects and positional neutralization effects are well attested, but apparent cases of positional enhancement are of very limited kinds and are better construed as cue preservation rather than as results of maximizing the distinctiveness of contrasts in context.

This inventory-based model of contrast is more tractable than the model proposed in Flemming (2004) according to which distinctiveness constraints and MAXIMIZE CONTRASTS interact with contextual markedness constraints to derive the set of possible words. As discussed above, that model implies that the wellformedness of a candidate word depends on its distinctness from neighbouring words, but their wellformedness depends in turn on distinctness from their neighbours, and so on, so it is difficult to evaluate individual words. In the inventory-based model it is still necessary to evaluate a word with reference to a set of minimally contrasting neighbours, but the potential neighbours are fixed by the inventory, and the evaluation of surface contrasts only has to determine whether to retain or neutralize a contrast, there is no possibility of reorganizing contrasts in order maximize the number or distinctness of contrasts. So in general it should only be necessary to consider near neighbors of a target underlying form in order to derive its surface form.

References

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

Benua, Laura (1997). *Transderivational Identity*. Ph.D. dissertation, University of Massachusetts, Amherst.

Broselow, Ellen I. (1976). *The Phonology of Egyptian Arabic*. PhD dissertation, University of Massachusetts, Amherst.

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.

Colarusso, John (1988). *The Northwest Caucasian Languages: A Phonological Survey*. Garland Publishing, New York.

Côté, Marie-Hélène (2000). *Consonant Cluster Phonotactics: A Perceptual Approach*. Ph.D. dissertation, MIT [distributed as *MIT Working Papers in Linguistics and Philosophy*]

Crosswhite, Katherine (1999). Intra-paradigmatic homophony avoidance in two dialects of Slavic. *UCLA Working Papers in Linguistics* 1, 48-67.

Fleischhacker, Heidi (2002). Cluster-dependent epenthesis asymmetries. Adam Albright and Taehong Cho (eds.) *UCLA Papers in Phonology* 5, 71-116.

Flemming, Edward (2002). *Auditory Representations in Phonology*. Garland Press, New York

Flemming, Edward (2004). Contrast and perceptual distinctiveness. Bruce Hayes, Robert Kirchner, and Donca Steriade (eds.) *The Phonetic Bases of Markedness*. Cambridge University Press, Cambridge. 232-276.

Flemming, Edward (2005). Targetless schwa revisited. Poster presented at the 149th Meeting of the Acoustical Society of America, Vancouver.

Fujimura, Osamu, M. Macchi, and L.A. Streeter (1978). Perception of stop consonants with conflicting transitional cues: A cross-linguistic study. *Language and Speech* 21, 337-346.

Gordon, M. (2001). Laryngeal timing and correspondence in Hupa. *UCLA Working Papers in Phonology* 7.

Itô, Junko (1989). A prosodic theory of epenthesis. *Natural Language and Linguistic Theory* 7, 217-259.

Jun, Jongho (1995). *Perceptual and Articulatory Factors in Place Assimilation: An Optimality-Theoretic Approach*. PhD dissertation, UCLA.
Jun, Jongho (2002). Positional faithfulness, sympathy and inferred input. Ms, Yeungnam University, Daegu, Korea.
Jun, Jongho (2004). Place assimilation. Bruce Hayes, Robert Kirchner, and Donca Steriade (eds.) *The Phonetic Bases of Markedness*. Cambridge University Press, Cambridge. 58-86.
Kao, Diane L. (1971). *Structure of the Syllable in Cantonese*. Mouton, the Hague.
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.
Kisseberth, Charles (1970). On the functional unity of phonological rules. *Linguistic Inquiry* 1, 291-306.
Kuipers, Aert H. (1960). *Phoneme and Morpheme in Kabardian (Eastern Adyghe)*. Janua Linguarum, series minor, no. 8. Mouton, The Hague.
Liljencrants, Johan, and Björn Lindblom (1972). Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language* 48, 839-62.
Lindblom, Björn (1990). Phonetic content in phonology. *PERILUS* 11, 101-118.
Lindblom, Björn, and Olle Engstrand (1989). In what sense is speech quantal? *Journal of Phonetics* 17, 107-121.
Lombardi, Linda (2001). Why Place and Voice are different: constraint-specific alternations and Optimality Theory. Linda Lombardi (ed.) *Segmental Phonology in Optimality Theory: Constraints and Representations*. Cambridge University Press, Cambridge.
Lynch, John (1978). *A Grammar of Lenakel*. Pacific Linguistics, Series B, no. 55. Asutralian National University, Canberra.
Maddieson, Ian (1984). *Patterns of Sounds*. Cambridge University Press, Cambridge.
Martinet, Andre (1952). Function, structure, and sound change.*Word* 8, 1-32.
McCarthy, John J. and Alan Prince (1995). Faithfulness and Reduplicative Identity. Jill Beckman, Suzanne Urbanczyk, Laura Walsh (eds.) UMOP-18: *Papers in Optimality Theory*.
Mohanon, K.P. (1993). Fields of attraction in phonology. John Goldsmith (ed.) *The Last Phonological Rule*: *Reflections on Constraints and Derivations*. Chicago University Press, Chicago. 61-116.
Newman, Stanley (1944). *The Yokuts Language of California*. The Viking Fund Publications in Anthropology 2, The Viking Fund, New York.
Ní Chiosáin, Maire, and Jaye Padgett (2001). Markedness, segment realization, and locality in spreading. Linda Lombardi (ed.) *Segmental Phonology in Optimality Theory*, Cambridge University Press, Cambridge, 118-156.
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. (2004) *Optimality Theory: Constraint Interaction in Generative Grammar*. Cambridge, MA: MIT Press.
Redford, M.A. & Diehl, R.L. (1999). The relative perceptual distinctiveness of initial and final consonants in CVC syllables. *Journal of the Acoustical Society of America*, 106, 1555-1565.
Rose, Sharon (2000). Epenthesis positioning and syllable contact in Chaha. *Phonology* 17, 397-425.
Rosner, B.S. and J.B. Pickering (1994). *Vowel Perception and Production*. OUP, Oxford.

Sapir, Edward (1912). The Takelma Language of Southwestern Oregon. Franz Boas (ed.) *Handbook of American Languages, Part 2*, 1-296. Smithsonian Institution, Washington D.C.
Schütz, Albert (1985). *The Fijian Language*. Honolulu: University of Hawaii Press.
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.
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.
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 (1999). Alternatives to syllable-based accounts of consonantal phonotactics. Osamu Fujimura, Brian D. Joseph, and Bohumil Palek (eds.) *Proceedings of LP '98: Item Order in Language and Speech*. Karolinum Press, Prague, 205-245.
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.
Steriade, Donca (2001). The phonology of perceptibility effects: the P-map and its consequences for constraint organization. To appear in Kristin Hanson and Sharon Inkelas (eds) *The nature of the word: essays in honor of Paul Kiparsky*. MIT Press, Cambridge.
Walker, D.C. (1984). *The Pronunciation of Canadian French.* University of Ottawa Press, Ottawa.
Wright, Richard (1996). *Consonant clusters and cue preservation in Tsou*. PhD dissertation, UCLA.
Wright, Richard (2001). Perceptual cues in contrast maintenance. Elizabeth Hume and Keith Johnson (eds.) *The Role of Speech Perception in Phonology*. Academic Press, New York, 251-277.
Yip, Moira (1988). The obligatory contour principle and phonological rules: a loss of identity. *Linguistic Inquiry* 19, 65-100.
Zhang, Jie (2004). The role of contrast-specific and language-specific phonetics in contour tone distribution. Bruce Hayes, Robert Kirchner, and Donca Steriade (eds.) *The Phonetic Bases of Markedness*. Cambridge University Press, Cambridge. 157-190.