10/20/08

Asymmetries between assimilation and epenthesis

Edward Flemming
Department of Linguistics & Philosophy, MIT

1. Introduction

This paper provides evidence that scales of perceptual distance between sounds play a central role in both markedness and faithfulness constraints. Perceptual distance governs the markedness of contrasts: contrasts between sounds are more marked the smaller the distance between the sounds (Flemming 2002, 2004). Correspondence between input and output forms is evaluated in terms of perceptual distance: greater distances between corresponding sounds incur greater violations of faithfulness (Steriade 2001a,b). Using the notation $\Delta(a, b)$ for the perceptual distance between sounds *a* and *b*, if $\Delta(a, b) < \Delta(c, d)$ then a contrast between *a* and *b* is more marked than a contrast between *c* and *d*. On the other hand a change from *a* to *b* between input and output incurs a lesser violation of faithfulness constraints than a change from *c* to *d*.

These two classes of constraints provide the basis for an account of some puzzling generalizations about where vowel epenthesis does and does not apply. The main problem addressed here is that vowel epenthesis should, in principle, be able to serve as a very general repair for dispreferred consonant clusters since inserting a vowel into a cluster necessarily eliminates that cluster, but in fact vowel epenthesis does not generally serve this function. For example, there is evidence that certain heterorganic consonant clusters are marked because they are often eliminated through place assimilation. This pattern has been analyzed in terms of a markedness constraint, AGREE(place), requiring adjacent consonants to agree in place features. Violating clusters such as [-ng-], [-mt-] can be brought into line with this constraint via assimilation of the first consonant to the second, [-ng-] → [-ŋg-], [-mt-] → [-nt-], but in principle the AGREE constraint could also be satisfied by epenthesizing a vowel between the heterorganic consonants ([-ng-] → [-nəg-]).

The same reasoning applies to any markedness constraint that penalizes a class of C_1C_2 clusters: it should be possible to satisfy the markedness constraint by altering one of the consonants, or by epenthesizing a vowel between the consonants, eliminating the cluster. So on this basis we should expect to find that markedness constraints on consonant clusters give rise to parallel typologies of processes involving featural changes such as assimilation, and vowel epenthesis. That is, the same clusters that are singled out by place assimilation processes should also be targeted by vowel epenthesis. For example, nasals are particularly prone to assimilation in the sense that they can undergo place assimilation in contexts where consonants of other manners, e.g. oral stops, do not undergo assimilation. If vowel epenthesis is an alternative to place assimilation as a repair for heterorganic clusters, then vowel epenthesis should also specifically target nasal-C clusters, but this is not the case: vowel epenthesis never specifically targets nasal-C clusters (cf. Hayes & Steriade 2004:20).

The typology of voicing assimilation in obstruents also lacks a parallel pattern of vowel epenthesis: it is common for obstruents to assimilate in voicing to a following obstruent, eliminating mixed-voicing obstruent clusters, e.g. [-tg-] → [-dg-], but these clusters are never specifically targeted by vowel epenthesis (Myers 2002, Steriade 2001).

These restrictions on vowel epenthesis can be accounted for in terms of a framework according to which both markedness and faithfulness constraints refer to the same scales of perceptual distance. The outline of the analysis is as follows: The marked clusters under discussion here are marked because they enter into contrasts that are insufficiently distinct – that is, they violate high-ranked distinctiveness constraints. For example, a heterorganic cluster like [-ng-] is marked because it is relatively indistinct from [-ŋg-] (cf. Jun 2004). The relevant distinctiveness constraint can, in principle, be satisfied by assimilation (-ng- → -ŋg-), neutralizing the problematic contrast, or by epenthesis (-ng- → -nəg-), which would allow more cues to the place of the nasal to be realized, yielding a sufficiently distinct contrast. The fact that neutralization is always preferred over epenthesis in these cases is explained in terms of the preference imposed by correspondence constraints to minimize the perceptual differences between input and output. We will argue that assimilation is preferred over epenthesis because it is a smaller perceptual change, i.e. Δ(ng, ŋg) < Δ(ng, nəg).

Consideration of the distance scales for place contrasts and contrasts between C_1C_2 and C_1VC_2 sequences supports the plausibility if this claim. The place difference, Δ(ng, ŋg), is relatively small – this is the basis for the markedness of the [ng] cluster. On the other hand, we will see evidence that Δ(ng, nəg) is at the high end of the scale of C_1C_2 - C_1VC_2 distances. $\Delta(C_1C_2, C_1VC_2)$ is smallest where sonority rises from C_1 to C_2, and greatest where sonority falls from C_1 to C_2 (Fleischhacker 2005, Steriade 2006). Sonority falls in nasal-stop clusters like [-ng-], so Δ(ng, nəg) is substantial.

This line of analysis can be generalized to account for the preference for neutralization over epenthesis as a repair for a wider range of marked C_1C_2 clusters, including mixed voicing obstruent clusters, such as [-tg-]. Processes that neutralize contrasts in C_1 position of C_1C_2 clusters are driven by distinctiveness constraints – that is, the C_1C_2 clusters are marked because the contrast between C_1 and some other consonant C_C is insufficiently distinct in the context preceding C_2, so $\Delta(C_1C_2, C_CC_2)$ falls below the threshold specified by a distinctiveness constraint. The empirical observation is that we do not find vowel epenthesis applying as an alternative to neutralization in the same contexts, i.e. the unfaithful mapping $C_1C_2 \rightarrow C_CC_2$ is always preferred over $C_1C_2 \rightarrow C_1VC_2$ although both candidate outputs would satisfy the distinctiveness constraint. This is explained in terms of the hypothesis that faithfulness constraints prefer a minimal perceptual distance between input and output: neutralization is preferred over epenthesis because $\Delta(C_1C_2, C_CC_2) < \Delta(C_1C_2, C_1VC_2)$.

$\Delta(C_1C_2, C_CC_2)$ is necessarily relatively small since this is the source of the markedness problem, so neutralizing the contrast ($C_1C_2 \rightarrow C_CC_2$) is always a modest violation of faithfulness. For major place and voice contrasts, where $\Delta(C_1C_2, C_CC_2)$ is small, $\Delta(C_1C_2, C_1VC_2)$ is liable to be large. This is because place and voice contrasts in C_1 position are least distinct where C_2 is an obstruent (Jun 2004, Steriade 1997), so the cluster cannot contain any significant sonority rise. Consequently epenthesis is a significant violation of faithfulness.

In each of the case studies analyzed here, major place assimilation and obstruent voicing assimilation, more direct evidence that $\Delta(C_1C_2, C_CC_2) < \Delta(C_1C_2, C_1VC_2)$ is provided by a study of half rhymes in limericks. Half rhymes in poetry are rhymes like *man-ham* where there are mismatches between the rhyming segments (in this case [n]-[m]). The frequency of a given type of segmental mismatch in poetic rhymes has been found to be a good index of the perceptual distance between the corresponding segments – mismatches between more similar segments are more frequent (Zwicky 1976, Steriade 2003, Kawahara 2007a). So frequency of half rhymes can be used to compare the relative perceptual similarity of diverse phonetic differences such as place and V-∅ contrasts. We will see that nasal place and obstruent voicing mismatches are much more frequent than C_1C_2-C_1VC_2 mismatches.

The lack of parallelism between the typologies of assimilation and epenthesis is demonstrated in more detail in the next section. Analyses of the individual patterns are developed in sections 3-4. Section 5 presents evidence for the differences in perceptual distance presupposed by these analyses from a variety of sources including a survey of half rhymes in English limericks. Then in section 6 we examine a prediction that the analyses developed here make about the kinds of consonant clusters that should be subject to vowel epenthesis. Specifically, we expect vowel epenthesis to break up consonant clusters where $\Delta(C_1C_2, C_1VC_2)$ is small, because then the contrast between presence and absence of a vowel is marked due to being perceptually indistinct, and epenthesis of a vowel is a small violation of faithfulness. As outlined above, $\Delta(C_1C_2, C_1VC_2)$ is smallest where the rise in sonority from C_1 to C_2 is largest, so we expect to find epenthesis processes that only target rising sonority clusters. We will see that this pattern is attested in languages like Winnebago. Section 7 discusses the broader typology of epenthesis into consonant clusters, and section 8 concludes by discussing broader implications of the proposed analyses.

2. Apparent mismatches between assimilation and epenthesis

The basic puzzle addressed here is the observation that vowel epenthesis is not employed to break up clusters that are identified as marked according to the typology of assimilation processes. Vowel epenthesis does break up consonant clusters, but only a limited variety of properties of consonant clusters affect the occurrence of vowel epenthesis.

A widely attested pattern of epenthesis depends only on cluster length: epenthesis applies to all medial clusters of three or more consonants (-CCC- → - CVCC- or -CCVC-) and to initial and final clusters of two consonants, regardless of the nature of the consonants involved (e.g. Yawelmani Yokuts, Newman 1944, and Lenakel, Lynch 1978). There are also epenthesis processes that break up clusters with particular sonority profiles: epenthesis can be restricted to apply only adjacent to sonorants (cf. Hall 2003), and may be further restricted to apply only preceding a sonorant in a rising-sonority cluster. The first pattern is exemplified by some Salishan languages where schwa is epenthesized before all sonorants that are not preceded by vowels, e.g. /mn/→[mən] but /mt/→[mt] (Flemming, Ladefoged & Thomason 2008). The second pattern is exemplified by Winnebago (Miner 1979), where epenthesis breaks up obstruent-sonorant clusters, e.g. /pr/→[par], but leaves obstruent-obstruent clusters intact, e.g. /ps/→[ps]. This paper provides an account of some surprising gaps in this typology of epenthesis into consonant clusters: many segmental properties that do contribute to the markedness of consonant clusters do not condition epenthesis. The system of markedness and faithfulness constraints that provides the basis for this account also derives epenthesis into rising sonority clusters, as in Winnebago (section 6.1). The other patterns of epenthesis are discussed briefly in section 7.

The evidence that vowel epenthesis does not serve to break up marked clusters comes from the observation that implicational universals governing patterns of place assimilation are not paralleled by implicational universals concerning vowel epenthesis in clusters. In intervocalic consonant clusters, - VC_1C_2V-, C_1 may assimilate in major place of articulation to the following consonant, C_2. Surveys by Mohanan (1993) and Jun (1995) establish an implicational universal concerning place assimilation: if stops in C_1 are targets of assimilation then so are nasals, but not vice versa. So there are languages like Malayalam in which nasals assimilate in place to a following stop (1a) but oral stops do not undergo place assimilation (1b) (Mohanan & Mohanan 1984), and there are languages like Korean in which both stops and nasals assimilate in place to a following stop (2) (Jun 1995), but there are no languages in which stops undergo place assimilation where nasals do not.

(1) Malayalam

a. Nasals undergo assimilation:

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

b. Stops do not undergo assimilation:

u̪tkarʂam ‘progress’
sap̪tam ‘eight’

(2) Korean – stops and nasals undergo assimilation:

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

This implicational universal suggests that heterorganic clusters of a nasal followed by a consonant violate some constraint that is not violated by heterorganic stop-consonant clusters. For illustrative purposes we will entertain an analysis according to which place assimilation is motivated by AGREE(place) constraints, which require that adjacent consonants must agree in major place features. To account for the implicational universal concerning targets of assimilation, we would have to posit manner-specific versions of this constraint, AGREE(place)-NC, applying to nasal-consonant clusters, and AGREE(place)-TC which applies to stop-consonant clusters. In addition we have to assume that AGREE(place)-NC universally ranks above AGREE(place)-TC, capturing the generalization that nasals are more subject to place assimilation than oral stops.

However the AGREE(place) constraints could also be satisfied by vowel epenthesis. That is, if a vowel is inserted into a heterorganic cluster then the consonants are no longer adjacent and thus satisfy AGREE(place). So given the fixed ranking AGREE(place)-NC >> AGREE(place)-TC we should find a typology of epenthesis environments that parallels the typology of place assimilation. That is epenthesis into stop-consonant clusters should imply epenthesis into nasal-consonant clusters, but not vice versa. For example, the ranking in (3) derives a pattern in which vowel epenthesis breaks up heterorganic nasal-stop clusters but not heterorganic stop-stop clusters, e.g. /anpa/ → [anəpa] but /atpa/ → [atpa]. This pattern is unattested (cf. Hayes & Steriade 2004:20)[1].

(3) AGREE(place)-NC, IDENT(place) >> DEPV >> AGREE(place)-TC

A similar issue arises with respect to obstruent voicing assimilation. In intervocalic obstruent clusters, it is common for the first obstruent to assimilate in voicing to the second, as in German, Russian, Lithuanian, Hungarian etc (e.g. Lombardi 1999, Steriade 1997). This process eliminates mixed voicing clusters in favor of clusters with uniform voicing, which implies that mixed voicing clusters are more marked. Again this is often formalized in terms of a constraint, AGREE(voice), requiring adjacent obstruents to have the same specification for [voice] (e.g. Lombardi 1999). This constraint could be satisfied by epenthesizing a vowel to break up mixed voicing clusters, e.g. /ab-ka/→ [abəka] but /ap-ka/ → [apka], but this pattern is not attested (Myers 2002, Steriade 2001).

So vowel epenthesis is not employed as a repair for clusters that are identified as marked based on the typology of place and voicing assimilation processes although vowel epenthesis could in principle eliminate the marked clusters.

As outlined above, we will propose that distinctiveness constraints account for the markedness of heterorganic and mixed-voicing obstruent clusters: both place and voicing contrasts are relatively indistinct in the pre-obstruent context (Steriade 1997, Jun 2004). In principle the relevant distinctiveness constraints could be satisfied by neutralizing the offending contrasts through assimilation, or by a epenthesizing a vowel into the cluster, which would allow more distinct realization of the contrasts. The fact that assimilation is the only attested repair is explained in terms of the preference for perceptually minimal modification imposed by the correspondence constraints: in these contexts, neutralizing an indistinct contrast is a smaller perceptual change than epenthesizing a vowel into the cluster, so neutralization is always preferred as the more faithful repair.

This line of analysis is developed with respect to voicing assimilation in the next section, and place assimilation in section 4.

3. Voicing assimilation

The analysis of voicing assimilation largely follows Steriade (1997, 2001a), but is formulated in rather different terms (see Flemming 2002:41f. for a related formulation). The markedness of mixed-voicing obstruent clusters, such as [agpa], is a consequence of the perceptual indistinctness of voicing contrasts in pre-obstruent position – i.e. Δ(agpa-akpa) is small compared to Δ(ga-ka). The preference to repair these indistinct contrasts through neutralization (agpa → akpa) rather than epenthesis (agpa → agəpa) is attributed

[1] Wilson (2001) proposes a variant of OT in which this kind of prediction can be avoided by making markedness constraints 'targeted', which in effect restricts the range of repairs that is considered for violations of a given markedness constraint. See McCarthy (2003) for a critique of this approach.

to the relative magnitudes of the perceptual distances involved: Δ(agpa, akpa) < Δ(agpa, agəpa), so neutralization is the more faithful repair.

This analysis depends on two perceptual distance scales, one concerning the distance between voiced and voiceless obstruents in various contexts, and the other concerning the distance between C_1C_2 and C_1VC_2 sequences.

3.1 Distinctiveness of obstruent voicing contrasts

Obstruent voicing contrasts are cued by differences on a variety of perceptual dimensions (e.g. Wright 2004), but the crucial dimension for present purposes relates to Voice Onset Time (VOT), the duration of the interval between consonant release and onset of voicing. Voiced stops have short VOT while voiceless stops have longer VOT. This is a salient cue to voicing contrasts, but it is not available in all contexts, and where it is unavailable, voicing contrasts are less distinct and thus prone to neutralization (Steriade 1997).

In particular VOT cues can only be realized properly where a voiced sonorant follows the obstruent, so they are unavailable in pre-obstruent contexts. If no voiced sound follows then there is simply no voice onset, but clear VOT cues also cannot be realized if the following sound is a voiced obstruent. The intensity of voicing during an obstruent is low (Stevens 1998:471, 479) so onset of voicing during an obstruent is not perceptually salient, and the release of the first obstruent is liable to be obscured if it is released into the following obstruent, so neither of the landmarks that usually make VOT such a salient cue are realized clearly in this context. Furthermore, it is not simply the lag between stop release and onset of voicing that cues voicelessness, it is also the frication and aspiration noise that fills that interval (Repp 1979), and these cues cannot be realized during an obstruent since aspiration requires an open vocal tract. In the absence of VOT, other cues remain available, such as presence of voicing during the closure of voiced obstruents, but this is a relatively weak cue due to the low intensity of voicing during obstruents.

Consequently voicing contrasts in pre-obstruent contexts violate higher-ranked distinctiveness constraints than voicing contrasts in pre-sonorant contexts, and thus can be singled out for neutralization. To formalize this analysis, we posit two perceptual dimensions relevant to obstruent voicing: VOT and Voice, where VOT refers to the lag between consonant release and voice onset and Voice refers to strength of periodicity during the consonant constriction. Pre-sonorant voiced stops have short VOT ([VOT *0*]) while voiceless stops have longer VOT ([VOT *1*])[2]. Voice takes a value of 0 in voiceless stops and 1 in voiced stops.

Note that the minimum value of VOT is 0, so we do not follow Lisker & Abramson (1964) in treating closure voicing as negative VOT. In the present context it is important to distinguish differences in closure voicing from differences in positive VOT since the later are much more salient. In any case closure voicing cannot always be treated as negative VOT, since it is not unusual for voicing to die out before stop release (cf. Keating, Linker & Huffman 1983).

[2] The VOT of voiceless stops can vary, most notably between aspirated and unaspirated stops, so further levels of VOT need to be differentiated in a full account of laryngeal contrasts, but all that matters for present purposes is a distinction between longer and shorter VOT.

Following Flemming (2002, 2004), the preference for perceptually distinct contrasts is implemented in terms of a ranked set of constraints requiring a specified minimum perceptual distance between contrasting sounds (4). Distances are specified along perceptual dimensions such as formant frequencies, VOT, etc. For example, the constraint MINDIST = F2:2 in (4) requires that contrasting sounds differ by at least 2 units on the second formant frequency (F2) dimension.

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

The crucial distinctiveness constraints for the analysis of voicing assimilation is MINDIST = VOT:1, which requires that contrasting consonants differ in VOT. Neutralization of voicing before obstruents is derived if MINDIST=VOT:1 outranks the faithfulness constraint IDENT(voice), as illustrated in (6-7).

To evaluate the distinctiveness of contrasts, it is necessary to evaluate sets of minimally distinct inputs together (Flemming 2002, 2004, Itô & Mester 2007, Lubowicz 2003, Ní Chiosáin & Padgett 2008). In this case, the tableaux (6) and (7) show the evaluation of obstruent voicing contrasts in the context preceding an obstruent. The candidates in each tableau involve faithful maintenance of the input contrast (a), neutralization through devoicing of the first obstruent (b), and neutralization through voicing of the first obstruent (c).

A contrast between voiced and voiceless obstruents preceding an obstruent, as in [apka - abka] (6, candidate *a*), violates the distinctiveness constraint MINDIST=VOT:1 since pre-obstruent consonants do not differ in VOT. This violation can be eliminated by neutralizing the contrast at the cost of changing the closure voicing of one of the stops, violating IDENT(voice). Since IDENT(voice) is lower-ranked, neutralization is preferred (candidates *b*, *c*).

Neutralization results in assimilation to the following obstruent because this is the least effort realization of the neutralized consonant. Following Flemming (2002:41f.) this is formalized in terms of two effort constraints (5). The first is a general constraint against voiced obstruents, *[+voice, -son], reflecting the difficulty of sustaining vocal fold vibration during an obstruent, due to rising oral pressure (Ohala 1983, Westbury & Keating 1986). The second is a more specific constraint, *TD, which is violated where a voiced obstruent follows a voiceless sound. This context is singled out because it is particularly difficult to initiate voicing during an obstruent, as opposed to maintaining vocal fold vibration into an obstruent, because a larger pressure drop across the glottis is required to initiate voicing than is required to sustain voicing (Westbury & Keating 1986, Titze et al 1995). Accordingly *TD universally outranks *[+voice, -son].

(5) a. *[+voice, -son] - no voiced obstruents.
b. *TD: *[-voice][+voice, -son] – a voiced obstruent must not follow a voiceless sound.

Given these two effort constraints, the preferred realization of a neutralized stop is voiced before a voiced obstruent to satisfy *TD (7) and voiceless before a voiceless obstruent to satisfy *[+voice, -son], given that *TD is irrelevant (6). *[+voice, -son] also accounts for the fact that obstruents are realized as voiceless when voicing is neutralized in final position.

(6)

apka, abka	MINDIST = VOT:1	IDENT (voice)	*TD	*[+voice, -son]
a. apka, abka	*!			*
b. ☞ apka		*		
c. abka		*		*!

(7)

apga, abga	MINDIST = VOT:1	IDENT (voice)	*TD	*[+voice, -son]
a. apga, abga	*!		*	***
b. apga		*	*!	*
c. ☞ abga		*		**

3.2 Why epenthesis is not an alternative to voicing neutralization

This analysis does not yet solve the problem of why epenthesis is never used to eliminate the marked mixed-voicing clusters. The problematic voicing contrasts in (6) and (7) could in principle be brought into conformity with the distinctiveness constraint, MINDIST = VOT:1, by epenthesizing vowels into the clusters, e.g. /abka/ → [abəka], since the vowel would allow for the realization of VOT cues. This repair is not attested - the only attested repair is assimilation, neutralizing the insufficiently distinct contrast. We adopt here Steriade's (2001a) account of this generalization in terms of the P-map hypothesis: Assimilation is the preferred repair because the perceptual difference between voiced and voiceless obstruents is small in the context preceding another obstruent, whereas vowel epenthesis between obstruents is perceptually salient, as shown schematically in (8). As a result, assimilation best satisfies the correspondence constraints.

This analysis requires an account of perceptual distance between C_1C_2 and C_1VC_2 sequences. This is developed in the next section, then we turn to evidence for the specific claim about relative distance schematized in (8).

(8) Δ(abka, apka) < Δ(abka, abəka)

3.2.1 Vowel-∅ similarity

Perceptual distance between C_1C_2 and C_1VC_2 sequences has been investigated extensively by Fleischhacker (2001, 2005). She provides evidence for the ranking of consonant cluster types shown in (9), where a cluster C_1C_2 at the left of the scale is more similar to a sequence consisting of the same consonants separated by a vowel, C_1VC_2.

(9) TL > SL > SN > ST
(L = liquid, N = nasal, S = sibilant, T = stop).

The evidence that most directly concerns vowel epenthesis comes from a study of the acceptability of vowel epenthesis as a repair for illicit onset clusters in loanword adaptation. She found that the susceptibility of initial clusters to epenthesis follows the hierarchy shown in (9), where epenthesis is more acceptable in clusters that are further to the left. But Fleischhacker shows that vowel epenthesis into C_1C_2 clusters is just one of several phenomena that place a C_1C_2 cluster in correspondence with a C_1V sequence, including partial copying of onsets in reduplication, alliteration and puns. In each case correspondence between C_1C_2V and C_1V is more acceptable with clusters that are further to the left in (9), providing substantial additional support for the validity of the hierarchy. Zuraw (2007) found that the same hierarchy governs the acceptability of infixing a /VC/ morpheme into consonant clusters in Tagalog.

Fleischhacker analyzes this hierarchy in terms of perceptual similarity between the corresponding C_1C_2 and C_1V sequences, i.e. a stop preceding a liquid, TL, is more similar to a stop preceding a vowel, TV, than a sibilant preceding a stop, ST, is to a sibilant preceding a vowel. An experimental study of similarity ratings of C_1C_2V- C_1VC_2V found the ordering TL, SL > SN > ST, which is consistent with the scale in (9) although it fails to distinguish the first two cluster types (Fleischhacker 2005:116).

Building on Fleischhacker's findings, Steriade (2006) proposes that the perceptual similarity between a consonant cluster C_1C_2 and the same cluster broken up by a vowel C_1VC_2, depends on the magnitude of the sonority rise in the cluster. According to Steriade's analysis, this is because vowel epenthesis introduces a sonority rise following C_1, but this is a smaller change if there is a similar sonority rise following C_1 in C_1C_2. In other words, the magnitude of the difference in sonority rise following C_1 is a key factor in the perceptual similarity of C_1C_2 and C_1VC_2. For example, epenthesis into a [kl] cluster, kl → kil, replaces a large sonority rise (kl) by a slightly larger rise (ki), and this constitutes a relatively small change. On the other hand epenthesis into an [st] cluster, st → sit, is a substantial change because it replaces a sonority fall (st) with a large sonority rise (si).

Building on these insights, we propose a measure of the perceptual distance $\Delta(C_1C_2, C_1VC_2)$ based on the difference in sonority rise following C_1 in the two sequences. This distance measure then provides the basis for sets of distinctiveness and faithfulness constraints. The measure is based on a sonority scale with six levels (10). We assume, following Parker (2008), that the notion of sonority that is relevant here corresponds to the perceived loudness of a sound, so it is fundamentally a perceptual quantity rather than a more abstract property of speech sounds (cf. Clements 1990:287ff.).

(10) Sonority scale

6	Vowel
5	Glide
4	Liquid
3	Nasal
2	Fricative
1	Stop

The sonority rise of a consonant is the difference in sonority between that consonant and the following segment, so for example the cluster [pr] has a sonority rise of 3 because sonority rises from 1 in the stop to 4 in the liquid. MINDIST and IDENT constraints evaluate the difference in sonority rise between C_1C_2 and C_1V. Difference is measured in terms of the ratio of the corresponding sonority rises – i.e. the sonority rise in C_1C_2 divided by the sonority rise in C_1V. A ratio is employed rather than a simple difference because subtracting the sonority rise in C_1C_2 from the sonority rise in C_1V simply gives the sonority of C_2, so such a measure would not be sensitive to the magnitude of the sonority rise in C_1C_2, contrary to the observed generalization. It is not unusual for perceptual distances to approximate ratios of the quantities being compared – this is essentially the generalization made by Weber's law in psychophysics (e.g. Gescheider 1997:2ff.).

To obtain a distance measure where larger differences in sonority rise correspond to larger numbers, the ratio is subtracted from 1. Perfect correspondence yields a ratio of 1, since the sonority rise in unchanged, so after subtraction from 1, perfect correspondence yields a Sonority Rise distance of 0 and differences in sonority rise yield positive distances. For example, the initial cluster of [**pr**a] has a rise in sonority of 3 (= 4-1), while the CV sequence that results from epenthesizing a vowel into that cluster, [**pə**ra], has a sonority rise of 5 (= 6-1), so the ratio of the sonority rises is 3/5=0.6, resulting in a distance of 1-0.6 = 0.4. The table in (11) shows sonority rise distances between a selection of C_1C_2 and C_1V sequences.

(11)

C_1C_2	Rise	C_1V	Rise	Distance
pr	3	pər	5	0.4
ps	1	pəs	5	0.8
pt	0	pət	5	1
np	-2	nəp	3	1.7

The correspondence constraints that refer to Sonority Rise distance are of the form IDENT(SONRISE)<*n*, requiring that the difference in Sonority Rise between corresponding input and output consonants is less than *n*.

3.2.2 A fixed ranking of correspondence constraints

The hypothesized universal, Δ(abka, apka) < Δ(abka, abəka), translates into a fixed ranking of the correspondence constraints that penalize these changes (12), since larger perceptual changes violate higher-ranked faithfulness constraints. The crucial value of sonority rise distance is 0.8 since epenthesis into an obstruent-obstruent cluster involves a distance of at least 0.8 between input and output (11).

(12) IDENT(SONRISE)<0.8 >> IDENT(voice)

Given this fixed ranking, voicing assimilation is always preferred over epenthesis into an obstruent cluster where the two are alternative means of satisfying a markedness constraint (13).

(13)

	apga, abga	IDENT(Son Rise)<0.8	MINDIST = VOT:1	IDENT (voice)	*TD	*[+voice, -son]
a.	apga, abga		*!		*	***
b.	apəga, abga	*!				***
c.	☞ abga			*		**

The fixed ranking (12) is plausible given that the difference between presence and absence of a vowel is relatively large between obstruents since an obstruent cluster involves at best a very small sonority rise (from stop to fricative). On the other hand, the small perceptual distance between voiced and voiceless obstruents preceding an obstruent is the source of the markedness violation incurred by mixed voicing clusters in the first place. More direct evidence for this ranking would ideally come from studies of the relative confusability of the two types of contrasts, but it is rare for perceptual studies to compare such disparate contrasts. However supporting evidence can be gleaned from a variety of phenomena that diagnose perceptual similarity.

For the comparison between epenthesis and a change in pre-obstruent voicing, Steriade (2001a) cites the work of Magen (1998) who assessed the perceptual salience of pronunciation errors in English produced by native speakers of Spanish. Magen asked native English-speaking listeners to evaluate how close recorded sentences were to native English, a task that presumably involves judging the perceptual distance of the utterance from the subjects' pronunciation norms. The stimuli involved utterances including a variety of errors characteristic of Spanish speakers of English, and versions of those utterances edited to eliminate or reduce the error, so the perceptual magnitude of the modifications can be assessed by comparing the ratings of corresponding edited and unedited stimuli. Three errors are relevant here: (i) schwa epenthesis in word-initial position (e.g. [ə]*speak*) and in final clusters (e.g. *clo*[zəd] for *clo*[zd]), which was corrected by deleting the schwa, (ii) devoicing of [z] (e.g. *rea*[s]*ons* for *rea*[z]*ons*), which was mitigated by shortening the fricative and lengthening the preceding vowel, and (iii) substitution of voiceless unaspirated stops for aspirated stops (e.g. [p]*ut* for [p^h]ut), which was corrected by addition of aspiration noise.

Epenthesis into final clusters is comparable to epenthesis into medial obstruent clusters, the change at issue here, since the final consonant was an obstruent in all of Magen's stimuli. This was among the most salient errors, together with initial epenthesis, whereas editing of the voicing-related errors did not have any significant effect on evaluations. While the voicing-related modifications did not involve comparisons between fully voiced and voiceless consonants, they are still relevant to our concerns because we are interested in pre-obstruent voicing contrasts which are only distinguished by a subset of the full cues to voicing, as discussed above. In particular, the [s]-[z] stimuli manipulated preceding vowel duration, which is a significant cue to voicing in post-vocalic obstruents (Raphael 1972). So Magen's results support the hypothesis that the perceptual difference between presence and absence of a vowel is greater than a difference in voicing.

Another source of evidence concerning the relative perceptual magnitude of phonetic differences comes from the study of half-rhymes in songs and poetry (Zwicky 1976, Steriade 2003, Kawahara 2007a). This is a valuable source because it makes it possible to compare diverse differences on an equal footing. Section 5 presents a study of half rhymes in English limericks that provides support for the fixed rankings posited to account for restrictions on epenthesis, including (12), above.

3.3 Distinctiveness constraints or AGREE(voice)?

We propose that both markedness and faithfulness constraints refer to the same scales of perceptual distance, but it should be noted that it is the faithfulness constraints that are essential to deriving the result that epenthesis does not serve as an alternative to voicing neutralization. That is, the assumption that epenthesis is a greater violation of faithfulness than a change in post-vocalic obstruent voicing is sufficient to account for the absence of epenthesis targeting mixed-voicing clusters even if we do not assume that assimilation is motivated by distinctiveness constraints. For example, an analysis based on AGREE(voice) would derive the non-occurrence of epenthesis into mixed-voicing obstruent clusters if epenthesis is universally a greater violation of faithfulness than a change in voicing. However the analysis in terms of distinctiveness constraints still has advantages over an AGREE-based analysis when we consider a broader range of facts about neutralization of obstruent voicing contrasts, and about positional neutralization in general.

Steriade (1997) shows that obstruent voicing assimilation is part of a broader implicational hierarchy relating positions in which obstruent voicing is neutralized. For example, neutralization of voicing contrasts in final position implies neutralization before obstruents, but not vice versa (Lombardi 1991, Steriade 1997, Wetzels & Mascaró 2001), so there are languages like Russian in which voicing contrasts are neutralized finally (14b) and before obstruents (14a), and there are languages like Hungarian in which voicing is neutralized before obstruents (15a) but the contrast is maintained word-finally (15b) (Petrova et al 2006, Siptár & Törkenczy 2000), but there are no languages in which voicing is neutralized finally without also being neutralized before obstruents. Steriade provides evidence that these implicational universals are related to the distinctiveness of voicing contrasts in the relevant contexts: voicing contrasts are neutralized first where they would be less perceptually distinct. So neutralization in word-final position implies neutralization preceding obstruents because obstruent voicing contrasts are more distinct in final position, mainly because cues from the quality of the release burst are available in that context.

An analysis of voicing neutralization in terms of distinctiveness constraints constitutes a reformulation of Steriade's (1997) analysis: less distinct contrasts violate higher ranked distinctiveness constraints, deriving the generalization that neutralization in better cued contexts implies neutralization in contexts where the contrast would be less distinct[3]. An analysis of voicing assimilation in terms of a constraint like AGREE(voice) fails to capture this implicational universal because AGREE(voice) cannot motivate final devoicing, so the implicational universal would have to be accounted for in terms of a fixed ranking of AGREE(voice) above an apparently unrelated constraint against voiced obstruents in word-final position.

(14) Russian voicing neutralization

	nom. sg.	gen. sg.	
a.	le[d]-ok	le[t-k]a	'ice'

[3] Steriade (1997) formulates her analysis in terms of a fixed ranking of context-specific *[αvoice] constraints, banning obstruent voicing contrasts from the specified context. These are essentially distinctiveness constraints since their ranking is determined by the distinctiveness of voicing contrasts in the specified context, but ban the occurrence of [voice] specifications rather than banning contrasts per se.

pro[s’]-it’ pro[z’-b]a ‘request’

b. vra[k] vra[g]-a ‘enemy’
bra[t] bra[t]-a ‘brother’

(15) Hungarian voicing neutralization

a. haːz ‘house’ haːs-toːl ‘from the house’
kɛrt ‘garden’ kɛrd-bɛ ‘in the house’

b. rɔb ‘prisoner’
kɔlɔp ‘hat’

Analyzing voicing assimilation in terms of distinctiveness constraints also accounts for the generalization that obstruents assimilate in voicing to following obstruents, but not to following sonorants[4]. Obstruent voicing contrasts are perceptually distinct before sonorants since VOT cues can be realized in that context, so this environment is not singled out for neutralization. AGREE-based analyses generally stipulate this property in the formulation of the AGREE(voice) constraint, specifying that it applies only to sequences of obstruents (e.g. Beckman 1998:26, Lombardi 1999).

4. Major place assimilation

This section addresses the generalization that major place assimilation can target nasals, leaving stops unaffected, but there is no parallel pattern of vowel epenthesis where epenthesis breaks up only nasal-obstruent clusters, leaving stop-obstruent clusters intact. The analysis is very similar in structure to the analysis of obstruent voicing assimilation in the preceding section: we will argue that assimilation is preferred over epenthesis in this case because changing place of articulation in these contexts is a smaller perceptual modification than epenthesizing a vowel, and is thus a lesser violation of faithfulness. Different nasal places of articulation are relatively similar in pre-consonantal position, so a change in place is a modest violation of faithfulness, but nasal-obstruent clusters are falling sonority clusters, so inserting a vowel into the cluster is a substantial perceptual change. The fact that nasal place contrasts are marked in the context preceding an obstruent and the fact that place neutralization is the preferred repair for this marked configuration are linked by the hypothesis that markedness and faithfulness constraints both refer to the same measures of perceptual distance between sounds.

As discussed by Jun (1995, 2004), major place contrasts (labial vs. coronal vs. dorsal) are more distinct preceding vowels than preceding consonants (Redford & Diehl 1999, Wright 2001). In both pre-vocalic and post-vocalic contexts place of articulation is cued by formant transitions, but release transitions carry more perceptual weight than

[4] More precisely, obstruents can assimilate to the voicing of sonorants, but this assimilation is never neutralizing. So intervocalic voicing of non-contrastive voiceless stops is attested (e.g. in Tümpisa Shoshone, Dayley 1989), as is voicing of word-final obstruents before voiced sonorants in some languages where obstruent voicing is independently neutralized in word-final position, e.g. Cracow Polish (Rubach 1996) and Slovak (Rubach 1993). These patterns are of interest, but are orthogonal to our primary concern here which is the environments in which obstruent voicing contrasts are neutralized.

transitions into the consonant closure (Fujimura et al 1978). In addition, prevocalic stop contrasts are distinguished by burst quality whereas overlap between consonants can result in the burst of a preconsonantal stop being obscured, so stop place contrasts benefit from additional cues in pre-vocalic contexts. For nasals it has been shown that a short interval around the release of a nasal carries significant cues to place (Kurowski & Blumstein 1993), whereas the interval around the closure does not carry comparable cues (Repp & Svastikula 1988).

An additional factor that plausibly affects the distinctiveness of pre-consonantal place contrasts is articulatory overlap with the following consonant. That is, the movements towards the second consonant of a cluster may begin before the first consonant constriction is formed, so the second consonant can affect the formant transitions into the first (Byrd 1992, Surprenant & Goldstein 1998). These effects are assimilatory in nature: the resulting formant transitions are shifted towards those characteristic of the second consonant (Dilley & Pitt 2007, Gow 2003), so overlap can result in less distinct contrasts between heterorganic and homorganic clusters.

Accordingly, pre-consonantal place contrasts violate higher ranking distinctiveness constraints than pre-vocalic place contrasts, and so are more prone to neutralization. If pre-consonantal place contrasts are neutralized, this results in regressive place assimilation because this is the least effort realization of the neutralized consonant – that is, a homorganic cluster like [mp] involves a single closure gesture, whereas a heterorganic cluster like [np] involves two (Jun 2004:72).

We are interested here in the Malayalam pattern where nasals assimilate in place to following stops while oral stops do not undergo assimilation. Jun (2004) argues that nasals are more prone to assimilation than oral stops because nasal place contrasts are less distinct than oral place contrasts in pre-consonantal contexts, that is Δ(anba, amba) < Δ(adba, abba). This is attributed to the deleterious effects of anticipatory nasalization on the distinctiveness of formant transitions, effects that have been demonstrated for vowel formants by Wright (1986). Unlike an oral stop closure, the nasal murmur can provide cues to place, but these cues are easily masked in normal speaking conditions (Malécot 1956, Johnson 2003:160) and so do not compensate for the reduction in the distinctness of the formant transitions.

In formalizing this analysis, we will formulate constraints on place contrasts in terms of the types of dimensions along which the contrasts differ rather than specific values on individual dimensions. The precise nature of the differences between places of articulation depends on vowel context (e.g. Kewley-Port 1982, Repp & Lin 1989), so a full treatment of place contrasts requires a relatively complex set of MINDIST constraints, specifying distances based on combined formant transitions and release properties (Flemming 2002). For the analysis of the Malayalam pattern, the crucial distinction is that a difference in nasalized closure transitions is less distinct than a difference in oral closure transitions. A difference in release transitions is in turn more distinct than a difference based on closure transitions. Accordingly we have the rankings of distinctiveness and correspondence constraints shown in (16) and (17).

(16) MINDIST = {nasal closure transitions} >>
MINDIST = {oral closure transitions} >>
MINDIST = {release transitions}

(17) IDENT{release transitions} >>
IDENT{oral closure transitions} >>
IDENT{nasal closure transitions} >>

Nasal place assimilation is then derived by the ranking in (18). A preconsonantal nasal place contrast violates MINDIST = {nas clos transitions) (candidate a). It is preferable to neutralize the contrast since the change in nasal place only violates IDENT{nas clos transitions}, which is lower ranked. The constraint *GESTURE favors neutralization to the homorganic cluster (candidate c) because this cluster involves only a single constriction gesture.

(18)

	anba, amba	MINDIST = {nas clos trans}	IDENT {oral clos trans}	MINDIST = {oral clos trans}	IDENT {nas clos trans}	MINDIST = {rel trans}	*GESTURE
a.	anba, amba			*!		*	***
b.	anba				*		**!
c.	☞ amba				*		*

Stop place contrasts are not neutralized in the preconsonantal context (19) because changing stop place (candidates b, c) would violate IDENT{oral clos transitions} which outranks the only distinctiveness constraint that is violated by the contrast, MINDIST = {release transitions} (candidate a). (Note that violation of IDENT{oral clos transitions} implies violation of IDENT{nas clos transitions}, since these descriptions are intended to index positions on a scale of perceptual distance, and a change in oral closure transitions is a greater perceptual modification than a comparable change in nasalized closure transitions).

(19).

	adba, abba	MINDIST = {nas clos trans}	IDENT {oral clos trans}	MINDIST = {oral clos trans}	IDENT {nas clos trans}	MINDIST = {rel trans}	*GESTURE
a.	☞adba, abba					*	***
b.	adba		*!		*		**
c.	abba		*!		*		*

Vowel epenthesis into nasal-obstruent clusters would also satisfy MINDIST={oral clos transitions} since it would allow for the realization of release formant transitions, which are more distinct than closure transitions, but as discussed above, epenthesis never specifically targets nasal-obstruent clusters. This option is ruled out because vowel epenthesis into a nasal-obstruent cluster would be a substantial violation of faithfulness, while a change in nasal place in this context constitutes a relative modest violation. Nasal-obstruent clusters have a falling sonority profile, i.e. a negative sonority rise, so epenthesis of a vowel involves a large change in sonority rise from a negative to a large positive value. The distance between –ms- and –məs- is 1.33 ([ms] has a sonority rise of -1 (i.e. sonority falls), [mə] has a sonority rise of 3, so the sonority rise distance between them is 1-(-1/3) = 1.33). On other hand, a change in the place of articulation of a pre-obstruent nasal is a comparatively small perceptual change – indeed this is the source of the markedness problem that drives place assimilation in the first place. So the contexts where place contrasts are less distinct involve clusters where V- Ø contrasts are perceptually distinct. Accordingly we hypothesize that Δ(anba, amba) < Δ(anba, anəba). This translates into the fixed ranking of correspondence constraints in (20).

(20) IDENT(SonRise) < 1.3 >> IDENT{nasal clos transitions}

Given this ranking, assimilation is always preferred over epenthesis (21). We will see below that this ranking is supported by evidence from a study of half rhymes.

(21)

	anba, amba	IDENT(SON RISE)<1.3	MINDIST = {nas clos trans}	IDENT {oral clos trans}	MINDIST = {oral clos trans}	IDENT {nas clos trans}
a.	anba, amba				*!	
b.	anəba, aməba	*!				
c.	☞ amba					*

4.1 Distinctiveness constraints or AGREE(place)?

As with the analysis of voicing assimilation, much of the work of explaining the absence of epenthesis in heterorganic nasal-obstruent clusters is done by the fixed ranking of Correspondence constraints rather than the distinctiveness constraints, so again we can ask what evidence there is that place assimilation is motivated by distinctiveness constraints rather than a constraint against heterorganic consonant clusters, AGREE(place).

As with obstruent voicing contrasts, there is an implicational hierarchy relating contexts of major place neutralization that follows from an analysis based on distinctiveness constraints but does not follow from an analysis based on AGREE(place). Neutralization of place contrasts in word-final position implies neutralization in pre-obstruent contexts, but not vice versa (de Lacy 2002:365). So there are languages like Diola Fogny (Sapir 1952) and Latin (Devine & Stephens 1977) where nasals must be homorganic to a following consonant, but nasal place contrasts are permitted in word-final position, and there are language like Spanish (Harris 1983) and Japanese (Vance 1987) where nasals must be homorganic to following obstruents, and nasal place contrasts are also neutralized in final position, to either [ŋ] (Japanese, Spanish dialects) or [n] (other Spanish dialects), but there are no languages which neutralize nasal place contrasts in word-final position while maintaining them in pre-obstruent contexts.

This implicational universal follows from an analysis where neutralization in both contexts is motivated by distinctiveness constraints. Place contrasts are more distinct in final position than in pre-obstruent contexts because they are not subject to the effects of gestural overlap with a following consonant, and release cues may be available, so pre-obstruent place contrasts violate higher-ranked distinctiveness constraints than final place contrasts. As a result pre-obstruent nasal place contrasts will always be neutralized if final place contrasts are neutralized. On the other hand, an analysis based on AGREE(place) cannot account for the implicational universal without further stipulations because AGREE(place) cannot motivate final place neutralization since it only applies to clusters. So final place neutralization would have to be motivated by other constraints, perhaps place markedness constraints such as *DORSAL, *LABIAL, which means that the implication between final and pre-obstruent neutralization would have to be derived by stipulating that AGREE(place) ranks above the place markedness constraints[5].

5. A survey of half-rhymes in limericks

The fact that epenthesis is not used as a repair for clusters that are eliminated by assimilation processes is analyzed here in terms of a pair of fixed rankings of correspondence constraints (22) that are the consequence of universal differences in perceptual distance between and among CC and CVC sequences, schematized in (23).

(22) a. IDENT(SONRISE) < 0.8 >> IDENT(voice)
b. IDENT(SonRise) < 1.3 >> IDENT{nasal clos transitions}

(23) a. Δ(abka, apka) < Δ(abka, abəka)
b. Δ(anba, amba) < Δ(anba, anəba)

There is good evidence that voicing contrasts are relatively indistinct preceding obstruents compared to their distinctiveness in pre-sonorant contexts, i.e. Δ(abka, apka) < Δ(ba, pa), and, as a result, voicing contrasts are commonly neutralized preceding obstruents. On the other hand, we have seen that C_1C_2 - C_1VC_2 contrasts are more distinct where C_1C_2 has a flat or falling sonority profile than when it has a rising sonority profile, e.g. Δ(abla, abəla) < Δ(abka, abəka), and we will see in section 6 that contrasts like [abla]-[abəla] are subject to neutralization.

---

[5] de Lacy (2002:363ff.) tries to derive the implication by effectively allowing place markedness constraints to be satisfied by assimilation, so a place markedness constraint that is ranked high enough to force final neutralization will also force neutralization in medial clusters. However, this approach does not actually derive the required generalization since it allows for an unattested pattern where place of articulation contrasts are neutralized to coronal in final position, but some place contrasts are still permitted in medial codas.

In de Lacy's system, neutralization to coronal in final position is derived by the ranking in (i), where *{KP} is a constraint against [labial] and [dorsal] consonants (de Lacy actually uses IDENT{KPT} rather than IDENT(place), and *{KPT}{KPT} rather than AGREE(place) for reasons that do not concern us here).

(i) onset-IDENT(place) >> *{KP} >> IDENT(place), AGREE(place)

The same ranking bans heterorganic labials and velars from medial codas, e.g. *[mt],*[ŋp] etc, but labial and velar codas are permitted when they are homorganic with the following onset, e.g [mp, ŋk], because they can share their place specification with the onset, and are then protected by higher-ranked onset-IDENT(place). So this ranking maps a cluster like /ŋp/ onto [mp] or [np]. Assimilation ([mp]) is favored by AGREE(place) even if that constraint is low-ranked since both outcomes are equally unfaithful.

However this ranking does not derive neutralization of all place contrasts in medial codas because heterorganic coronals are permitted to surface since they do not violate *{KP}. That is the contrasts [np] vs. [mp] and [nk] vs. [ŋk] are permitted by this ranking. So it derives an unattested pattern where there are no place contrasts in final position but place contrasts are preserved in pre-obstruent contexts. The generalization that de Lacy's system derives is that neutralization in final position implies neutralization of some place contrasts in medial codas, but this is not sufficient: Neutralization of all place contrasts in final position implies neutralization of all place contrasts in pre-obstruent contexts.

So it is clear that Δ(abka, apka) is on the low end of the scale of distinctiveness of voicing contrasts, and Δ(abka, abəka) is on the high end of the scale of distinctiveness of C_1C_2 - C_1VC_2 contrasts, so the hypothesized universal in (23a) is plausible. But ideally we would like some means to directly compare the distinctiveness of these two kinds of contrasts. As mentioned above, perceptual studies generally do not compare such diverse contrasts, although we saw that there is some evidence supporting Δ(abka, apka) < Δ(abka, abəka) from a study of the perceptual salience of signal modifications that include vowel deletion and change to voicing cues (Magen 1998). Here we turn to the study of half-rhymes in poetry and song as a general source of evidence concerning perceptual similarity that is applicable to a wide variety of contrasts. In this sections we will see that evidence of this kind drawn from previous studies, together with a new survey of half rhymes in English limericks, supports the rankings in (22).

Perfect rhymes involve a match between the sound sequences from the stressed vowels to the ends of the rhyming words. In half rhymes (a.k.a imperfect rhymes, semi rhymes) there is some mismatch between the rhyming sequences, e.g. *roo**m**-moo**n***. It is clear that poets and lyricists who employ half rhymes do not abandon the requirement of identity between rhyming sequences, rather they relax it to a requirement of similarity since the mismatches typically involve correspondences between similar sounds, and rhymes that involve greater similarity are more common. Moreover, the relevant notion of similarity appears to be the same as the confusability-related notion that we are concerned with here (Zwicky 1976, Steriade 2003, Kawahara 2007a). Accordingly the frequency of a given mismatching correspondence in half rhymes can be used as an index of the perceptual similarity of the corresponding sounds (Steriade 2003).

Previous studies of half rhymes in English, Japanese and Romanian have found that voicing and place mismatches are common (e.g. *died-light*, *room-moon*), while V-Ø correspondences are not reported at all (Zwicky 1976, Steriade 2003, Kawahara 2007:131f.) indicating that presence versus absence of a vowel is a much larger perceptual difference than a difference in place or obstruent voicing[6]. This conclusion is supported here by a study of half rhymes in English limericks. In Limericks V-Ø mismatches are sometimes observed in half rhymes so we can be certain that these rhymes are not excluded independently of the similarity requirement for a satisfactory rhyme (e.g. by metrical requirements), but V-Ø mismatches are rare compared to voicing and place mismatches.

5.1 The half rhyme corpus

The corpus for the study was collected from the 1739 limericks compiled in Legman (1969), and consists of 154 half rhymes involving consonant mismatches or differences in syllable count. No attempt was made to identify half rhymes based on vowel mismatches because similarity between vowels is not at issue here, and because dialectal variation in vowel systems often makes it difficult to determine whether a rhyme actually

[6] Zwicky (1976) mentions the existence of apocopated rhymes where one of the rhyming words has a n additional final unstressed syllable not present in the other, e.g. *end-offend**ed***. However, this pattern of rhyming seems to involve a departure from the canonical definition of rhyme domain rather than imperfect matching between the segments in the rhyming domains. See Kawahara (2007:117) for a similar phenomenon in Japanese hip-hop rhymes.

involves a mismatch. If a rhyme would be perfect given some established pronunciation of the words involved, it was not counted as a half rhyme. For example, apparent correspondence between presence and absence of [h] in an unstressed syllable, e.g. *drove her-over*, were not counted as half rhymes since [h] can be deleted in this context. The exclusion that is most relevant to the contrasts of concern here concerns apparent mismatches in which the velar nasal [ŋ] of the suffix *–ing* corresponds to an alveolar [n]. These were not counted as half rhymes since many dialects allow the suffix to be pronounced [-ɪn].

To a give an overview of the range of half rhymes, counts are given for broad categories of mismatches in (24). Manner differences include nasality mismatches, e.g. *Cor**b**ie-bore **m**e*, as well as mismatches in constriction degree, e.g. *dou**b**ly-lo**v**ely*. In addition mismatches such as [b-v] and [θ-t] are counted as mismatching only in manner although there is some difference in place, as long as the articulator matches. Corresponding segments are counted as having the same place of articulation as long as they share the same articulator. 14 rhymes contain more than one mismatch. In these cases, the mismatches are counted separately (cf. Zwicky 1976:680). For example, *mistress-kissed his* contains a C-Ø mismatch (-str- vs. –st-) and a voicing mismatch (-s vs. –z) and these are both included in the counts in (24), so the sum of the frequencies of the mismatch types is greater than the number of rhymes. The rhyme scheme of a limerick is AABBA, so the 1st, 2nd and 5th lines rhyme with each other. Mismatches among these three rhymes are only counted once. For example, if the rhyming words are *clean-cream-queen*, this is counted as a single place mismatch, it is not treated as two half rhymes *clean-cream* and *cream-queen*.

(24)

Type	Frequency	Description	Examples
voice	48	mismatching segments differ only in voicing	*ja**zz**-ga**s**, plea**s**ure-pre**ss**ure*
place	40	mismatching segments differ only in place	*fi**n**e-ti**m**e, spe**c**ies-fe**c**es*
manner	14	mismatching segments differ only in manner	*dou**b**ly-lo**v**ely, Cor**b**ie-bore **m**e*
C-Ø	42	presence vs. absence of a consonant	*gown-foun**d**, clo**th**es-woes*
syllable count	16	rhymes differ in number of syllables	*kidney-didn’ he, Pitlochry-mockery*
other	11	e.g. mismatching segments differ on multiple dimensions	*Spitzbergen-virgin, whether-better*

From the table in (24), it is apparent that the most common mismatches are voice, place and C- Ø mismatches, which are also common mismatches in Zwicky’s (1976) study of half rhymes in rock lyrics. The frequencies of three types of mismatch, voice, place and syllable count, bear on the generalizations about perceptual distance that we are interested in, Δ(abka, apka) < Δ(abka, abəka) and Δ(anba, amba) < Δ(anba, anəba), so we will examine these types in more detail.

5.1.1 Voice mismatches vs. T_1T_2-T_1VT_2 mismatches

Mixed-voicing obstruent clusters do not generally occur within words in English, but we can obtain a measure of the distinctiveness of voicing contrasts that lack VOT cues from examination of voicing mismatches in word-final, post-vocalic contexts, e.g. *relate it-belate**d***. As discussed above (3.1), word-final voicing contrasts are generally more distinct than pre-obstruent voicing contrasts, i.e. Δ(ab, ap) > Δ(abka, apka), so if Δ(abka, abəka) is in turn greater than Δ(ab, ap), we can be confident that it is also greater than Δ(abka, apka). Consequently, we will compare the frequency of word-final voicing mismatches to the frequency of C_1C_2 - C_1VC_2 mismatches where C_1C_2 is an obstruent-obstruent cluster.

Of the 48 voicing mismatches, 29 involve word-final consonants (e.g. *jazz-gas*), while the remaining 19 involve consonants preceding vowels or sonorant consonants (e.g. *pleasure-pressure*). This asymmetry is expected, given the evidence that voicing contrasts are more distinct preceding sonorants.

There are ten C_1C_2 - C_1VC_2 mismatches. The remaining syllable count mismatches involve four V.V-V correspondences, e.g. *liable-bible*, and two other cases, where VC(C) corresponds to Ø *tremendous-pendulous* and *Herman-permanent* (the latter could be an apocopated rhyme (Zwicky 1976:676)). 9 of the 10 cases of C_1C_2 - C_1VC_2 correspondence involve rising sonority clusters (25). Just one involves a falling sonority cluster, *Axmi**ns**ter-si**nis**ter*[7].

(25) *Pitlo**chr**y-moc**ker**y* (in two limericks)
*Minnea**pol**is-ha**pl**ess*
***Limer**ick-sli**m R**ick*
*colo**str**um-fos**ter** 'em*
***Leg**m**an-e**ggin**' 'em*
*fi**nall**y-divi**nel**y*
*did**dle** her-Hi**tl**er*
*ki**dn**ey-di**dn**' he*

So there are no examples of C_1C_2 - C_1VC_2 mismatches where C_1 and C_2 are obstruents, and only one where sonority is not rising in the consonant cluster. On the other hand there are 29 voicing mismatches among final obstruents. These counts, support the claim that vowel epenthesis into an obstruent cluster is a greater perceptual modification than a change in voicing in pre-obstruent context, Δ(abka, apka) < Δ(abka, abəka). In fact pre-sonorant voicing mismatches are also more frequent than C_1C_2 - C_1VC_2 mismatches, suggesting that voicing contrasts are in general less distinct than V-Ø contrasts. The results are also consistent with the claim that a V-Ø difference is less distinct in rising sonority contexts than in falling sonority contexts, as discussed above, although the small numbers involved (9 vs. 1) make this comparison less reliable.

[7] This rhyme is peculiar in another respect: the initial syllable of *Axminster*is unmatched in *sinister*, so [æksmɪ] appears to be in correspondence with the [ɪ] of *sinister*, assuming that stress is initial in both words. This suggests that some other stress pattern is intended. There may be a secondary stress on *–min-*, of *Áxmìnster*, in which case the rhyme domain may run from this secondary stress: Axm(*inster)-s(inister)*.

5.1.2 Nasal place mismatches vs. NT-NVT mismatches

To evaluate the hypothesis that Δ(anba, amba) < Δ(anba, anəba) we need to compare the frequency of place mismatches among post-vocalic nasals to the frequency of C_1C_2 - C_1VC_2 mismatches just described. We cannot count place mismatches among nasals preceding obstruents, since English does not allow heterorganic nasal-obstruent clusters in most contexts, but as with voicing contrasts, word-final nasal place contrasts should be more distinct, i.e. Δ(an, am) > Δ(anba, amba). So if Δ(an, am) < Δ(anba, anəba), we can conclude by transitivity that Δ(anba, amba) < Δ(anba, anəba).

Of the 40 place mismatches, 30 involve stops and/or nasals. The remainder involve fricatives (8) or glides (2). Among the stop and nasal place mismatches, 23 are in word-final position or precede an obstruent, e.g. *fi**n**e-ti**m**e, gu**mp**tion-fu**nc**tion*, while only 7 precede vowels or approximants, e.g. *co**v**ers-mo**th**ers*, *Ni**x**on-benedi**ct**ion*. This difference is expected given that place contrasts are more distinct in the pre-vocalic context. Of the post-vocalic place mismatches, 22 involve nasals (e.g. *damn-man*), or nasal-stop clusters (e.g. *trunk-cunt*), and 1 involves oral stops (*hog-god*). We group together nasals and nasal-stop clusters because both suffer from the effects of nasalization on their formant transitions, although nasal-stop clusters might be expected to benefit from cues in the release of the stop. The higher frequency of nasal place mismatches supports Jun's argument that nasal place contrasts are less distinct than stop place contrasts in this context. Recall also that the count of nasal mismatches is conservative because it excludes many apparent [ŋ]-[n] mismatches involving the suffix *–ing*. Zwicky (1976) found a similar asymmetry between nasal place and stop place mismatches.

So post-vocalic nasal place mismatches are much more frequent than C_1C_2 - C_1VC_2 mismatches where C_1C_2 has a falling sonority profile, supporting the hypothesis that the latter contrasts are more distinct, i.e. Δ(anba, amba) < Δ(anba, anəba).

6. Neutralization of V- Ø contrasts

The analyses developed so far have employed a set of correspondence constraints, IDENT(SonRise)<*n*, that refer to the perceptual distance between C_1C_2 and C_1VC_2 sequences. Given the hypothesis that markedness and faithfulness constraints refer to the same scales of perceptual distance, we also expect there to be markedness constraints on the distinctiveness of C_1C_2 and C_1VC_2 contrasts based on sonority rise distance. In this section we examine further predictions that follow from positing these sets of constraints.

Markedness and faithfulness constraints based on Sonority Rise distance predict that, on the one hand, contrasts between C_1C_2 and C_1VC_2 should be more marked where they are less distinct, i.e. where the sonority rise in C_1C_2 is larger, and that epenthesis should also be more acceptable in these contexts. Both of these considerations lead to the prediction that there should be epenthesis processes that only target rising sonority clusters. This prediction is confirmed: there are patterns of vowel epenthesis that target obstruent-sonorant clusters but leave obstruent-obstruent clusters intact (/pra/ → [para], /pʃo/ → [pʃo]), e.g. Winnebago.

A further prediction of the system of constraints proposed here is that the same scale of C_1C_2 and C_1VC_2 similarity should govern both vowel epenthesis and vowel deletion processes. I.e. both vowel deletion and vowel epenthesis place C_1C_2 and C_1VC_2

sequences in correspondence, only the direction of the mapping differs: $C_1C_2 \rightarrow C_1VC_2$ or $C_1VC_2 \rightarrow C_1C_2$. Accordingly, the same scale of perceptual distance should regulate both types of process, so vowel deletion should incur a smaller violation of faithfulness in the same contexts as vowel epenthesis. Patterns of schwa deletion in English and Dutch show that these processes do follow the same similarity scale as vowel epenthesis into consonant clusters.

6.1 TR epenthesis

Epenthesis processes that target rising sonority clusters ('TR epenthesis') have been documented by Steriade (1990) and Hall (2003). The most striking example is Dorsey's Law in Winnebago (Miner 1979, 1993). In this process, a copy of the following vowel is inserted into all obstruent-sonorant sequences, word-initially and word-medially (26), while obstruent-obstruent clusters are allowed to surface (27) (Miner 1993:113, 117).

(26)	/pras/	parás	'flat'
	/knak/	kãnã́k	'put something not having length'
	/hakwe/	hakewé	'six'
	/hipres/	hiperés	'know'
	/sni/	sĩ́nĩ	'cold'
	/ʃ-ruxruk/	ʃuruxúruk	'you earn'

(27)	psĩ́ipsĩ́tʃ	'awkward'
	ratʃgã́	'to drink'
	xdʒaːnã́ne	'yesterday'
	ksáatʃ	'stiff'
	kdʒée	'revenge'
	skáa	'white'
	xgããsák	'energetic'
	pʃoopʃótʃ	'fine'

A similar pattern of optional epenthesis is observed in Late Latin, although restricted to obstruent+liquid clusters (Steriade 1990), e.g. *Aleksan**dr**i-Aleksan**diri**, s**cr**iptum-s**ciri**ptum.*

These patterns initially appear somewhat surprising since obstruent-sonorant clusters are generally regarded as relatively unmarked clusters. It has been argued that rising sonority clusters represent a marked syllable contact when divided by a syllable boundary (Venneman, 1988; Davis, 1998; Gouskova, 2002), but this cannot account for TR epenthesis since it applies to initial clusters as well as medial clusters. However, the impression that rising sonority clusters are relatively unmarked comes from the fact that most consonantal contrasts, such as place and voice contrasts, are relatively well cued in these clusters (better than in obstruent-obstruent clusters), but V-Ø contrasts are less distinct, and therefore more marked, in these clusters, i.e. $\Delta(\mathrm{TR}, \mathrm{TVR}) < \Delta(\mathrm{TT}, \mathrm{TVT})$.

TR epenthesis can be analyzed in terms of a set of MINDIST constraints that refer to Sonority Rise distance, as in (28).

(28) MINDIST = SonRise:0.7 >> MINDIST = SonRise:0.8 >> MINDIST = SonRise:0.9…

The table in (29) shows the sonority rises for a variety of Winnebago consonant clusters (left column) and minimally distinct C_1VC_2 sequences (center column), together with the Sonority Rise distance (rightmost column). It can be seen that Winnebago requires a minimum difference in sonority rise of 0.8 for an adequate contrast between C_1C_2 and C_1V. Any smaller differences are neutralized by epenthesizing a vowel to break up the consonant cluster. We will now see how this threshold is established by the ranking of MINDIST constraints with respect to Correspondence constraints.

(29)

	C_1C_2	Rise	C_1V	Rise	Distance
	pra	3	**pa**ra	5	0.4
epenthesis	**sr**a	2	**sa**ra	4	0.5
	sni	1	**si**ni	4	0.75
no	**ps**i	1	**pi**si	5	0.8
epenthesis	**kt**ʃe	0	**ke**tʃe	5	1

A contrast like [pra]-[para] violates MINDIST=SonRise:0.8. This constraint violation can be avoided by neutralizing the contrast, for example by vowel epenthesis, [pra] → [para]. But this deviation from the input form [pra] incurs violations of the faithfulness constraints, IDENT(SONRISE)<*n*, which require that the difference in Sonority Rise between corresponding input and output consonants should be less than *n*.

The tableaux in (31)-(33) illustrate how the ranking of IDENT and MINDIST constraints in (30) derives neutralization of V- Ø contrasts in the context between an obstruent and a sonorant (31, 32) but not between obstruents (33). The tableaux show the evaluation of C_1C_2V and C_1VC_2V contrasts. The candidates in each tableau involve faithful maintenance of the input contrast (a), neutralization by mapping both inputs onto a C_1C_2V output by vowel deletion (b), and neutralization by mapping both inputs onto a C_1VC_2V output by vowel epenthesis (c). The tableaus in (31) and (32) only illustrate the derivation of V- Ø neutralization – that is candidates (b) and (c) are tied with respect to the constraints shown because the violations of IDENT(SonRise) are the same for epenthesis and deletion. The choice between epenthesis and deletion is analyzed below.

(30) MINDIST = SonRise:0.8 >> IDENT(SonRise)<0.7 >> MINDIST = SonRise:0.9

It can be seen that the critical rankings in (30) establish the minimum Sonority Rise distance of 0.8 for a V- Ø contrast by comparing contrasts that lie on each side of this threshold: [sni] vs. [sini], with a distance of 0.75 (32), and [psi] vs. [pisi], with a distance of 0.8 (33). A [sni]-[sini] contrast violates MINDIST = SonRise:0.8 (candidate *a*) and the highest ranked correspondence constraint that is violated by neutralization is IDENT(SonRise)<0.7, which is lower-ranked, so neutralization of the V- Ø contrast is preferred (32, candidates *b*, *c*). On the other hand, a difference of 0.8, as in [psi]-[pisi], only violates lower-ranked MINDIST = SonRise:0.9 (33, candidate *a*) and cannot be neutralized without violating higher ranked IDENT(SonRise)<0.8 (and IDENT(SonRise)<0.7) (candidates *b*, *c*), so the contrast is permitted. The sonority rise distance of 0.75 between [sni] and [sini] is the largest that is possible in an obstruent-sonorant cluster (29) and the V-Ø contrast is neutralized due to insufficient distinctiveness in this context, so the same applies to all other obstruent-sonorant clusters, as illustrated for [pra] vs. [para] in (31).

(31)	pra, para	IDENT(Son Rise)<0.8	MINDIST = SonRise:0.8	IDENT(Son Rise)<0.7	MINDIST = SonRise:0.9
a.	pra, para		*!		*
b.	(☞) pra				
c.	☞ para				

(32)	sni, sini	IDENT(Son Rise)<0.8	MINDIST = SonRise:0.8	IDENT(Son Rise)<0.7	MINDIST = SonRise:0.9
a.	sni, sini		*!		*
b.	(☞) sni			*	
c.	☞ sini			*	

(33)	psi, pisi	IDENT(Son Rise)<0.8	MINDIST = SonRise:0.8	IDENT(Son Rise)<0.7	MINDIST = SonRise:0.9
a.	☞ psi, pisi				*
b.	psi	*!		*	
c.	pisi	*!		*	

The preference for epenthesis over deletion as the outcome of neutralization derives from other constraints. One factor that favors epenthesis is that this improves the distinctiveness of C_1 place contrasts by allowing for the realization of formant transitions into the vowel. In formalizing this analysis, we will formulate constraints on place contrasts in terms of the types of dimensions along which the contrasts differ rather than specific values on individual dimensions, as in section 4. The only essential claim here is that the addition of release formant transitions results in more distinct contrasts, so we will employ the simplified constraints, MINDIST={burst}, requiring that contrasting sounds differ in burst properties, MINDIST={burst & rel. transitions}, requiring that contrasting sounds differ in both burst and release formant transitions, ranked as in (34). This ranking enforces preferences for intervocalic stops over post-vocalic stops, [ipere] > [ipre], and for prevocalic stops over preconsonantal stops in word-initial position, [para] > [pra], thus favoring epenthesis into both initial and medial clusters. The MINDIST constraints relevant to fricative place contrasts are similar but with ‘noise spectrum’ in place of ‘burst’.

(34) MINDIST={burst} >>
MINDIST={burst & rel. transitions}

The tableau in (35) illustrates how these constraints select epenthesis over vowel deletion in the obstruent_sonorant context. We have to consider both place and V- Ø contrasts here, resulting in a total of four inputs. Inputs and outputs are subscripted to clarify the input-output mappings in each candidate set. As seen in (31), constraints on Sonority Rise differences make the V- Ø contrast in pairs like [pra-para] and [kra-kara] unacceptable, so the tableau in (35) only shows the evaluation of two alternative strategies for neutralizing this contrast: deletion (candidate *a*) and epenthesis (candidate *b*). Epenthesis is preferred because it results in a more distinct contrast between [p] and [k], i.e. it satisfies MINDIST={burst & rel. transitions}.

Note that a difference in burst alone is sufficient for contrast in this language since IDENT(burst) ranks above MINDIST={burst & rel. transitions} – i.e. a contrast based on burst properties alone is preferred over place neutralization because the latter would incur a violation of IDENT(burst). This is necessary to allow for place contrasts in initial obstruent clusters, e.g. [ps-] vs. [ks-], which are attested in Winnebago.

(35)

	pra_1, $para_2$, kra_3, $kara_4$	MINDIST = {burst}	IDENT (burst)	MINDIST = {burst & rel.trans}
a.	$pra_{1,2}$, $kra_{3,4}$			*!
b.	☞ $para_{1,2}$, $kara_{3,4}$			

Combining the two parts of the analysis raises the issue of comparison of perceptual distances on very difference dimensions. That is, if we try to evaluate the combination of place and V- Ø contrasts in (35) with respect to the combined constraints from the tableaus in (31) and (35), it seems that place contrasts like [pra-kra] or [para-kara] violate all the MINDIST constraints on Sonority Rise differences since these pairs do not differ along this dimension. Similarly the V- Ø contrasts, [pra-para] and [kra-kara] appear to violate MINDIST={burst & rel. transitions}. This is not the correct result: a contrast is acceptable if it is sufficiently distinct along some dimension(s), it doesn't have to be distinct along all dimensions.

So MINDIST constraints should ideally refer to differences on a general measure of perceptual distance that can be satisfied by any acoustic differences that exceed the required perceptual distance (cf. Flemming 2002:30ff.). In other words, MINDIST = SonRise:0.8 should be replaced by a constraint of the form MINDIST = *d* which is satisfied by a difference in Sonority Rise of 0.8, but is also satisfied by equivalent differences in burst spectrum and/or formant transitions. A contrast like [pra-kra] might actually satisfy this reformulated MINDIST constraint. Given our limited understanding of perceptual distances among speech signals, we cannot be at all certain what the relevant equivalences are between place cues and sonority rise distances, but a variety of assumptions yield the desired result here. (36) illustrates one possibility, specifying the alternative place and sonority rise distances as disjuncts in each MINDIST constraint. Violation marks are labeled to indicate which contrast incurs each violation.

(36)

	pra_1, $para_2$, kra_3, $kara_4$	IDENT (SonRise)<0.8	MINDIST = SonRise:0.8 or {burst}	IDENT (SonRise)<0.7	MINDIST = SonRise:0.9 or {burst & rel.trans}
a.	pra_1, $para_2$, kra_3, $kara_4$		*(kr-kar)! *(pr-par)		*(pr-par) *(kr-kar) *(pr-kr)
b.	$pra_{1,2}$, $kra_{3,4}$				*(pr-kr)!
c.	☞ $para_{1,2}$, $kara_{3,4}$				

A final issue in the analysis of Winnebago concerns the quality of the epenthetic vowel: why does it have the same quality as the vowel following the sonorant? Given the hypothesis that Input-Output Correspondence constraints favor perceptually minimal modification of the input, these constraints should favor the epenthetic vowel quality V that makes the output C_1VC_2 most similar to the input C_1VC_2. As argued by Steriade (2001a), this should generally favor a short vowel that is assimilated to its context, a quality that is usually transcribed as schwa [ə]. It is conceivable that the transcription of the epenthetic vowel as a copy of the following vowel is simply an approximation representing a schwa vowel that is affected by coarticulation with the following vowel, but this account is problematic where C_2 is a glide, because in these cases, we would expect a vowel quality close to that of the glide to constitute the perceptually minimal vowel insertion, i.e. we would expect [u] before [w] and [i] before [j] rather than a copy of the following vowel. Thus we tentatively hypothesize that the quality of the epenthetic vowel is the result of a form of vowel harmony (cf. Kawahara 2007b), perhaps motivated by AGREE(F) constraints, requiring identity between consecutive vowels (Baković 2000). However the main concern here is the distribution of epenthetic vowels rather than their quality, so we will not develop this aspect of the analysis further.

6.1.1 Could TR epenthesis be motivated by distinctiveness constraints on place contrasts?

The analysis in (36) raises the possibility that epenthesis is motivated by constraints on the distinctiveness of consonant place contrasts rather than constraints on the distinctiveness of V- Ø contrasts. That is, we have seen that the constraint MINDIST={burst & rel. transitions} favors epenthesis in all C_1C_2 consonant clusters to allow for the realization of burst and release transition cues to place in C_1. In the analysis above, it is assumed that MINDIST={burst & rel. transitions} is a low-ranked constraint that simply breaks the tie between epenthesis and deletion as the means to neutralize insufficiently distinct V- Ø contrasts, but if this MINDIST constraint were ranked between IDENT(SonRise)<0.8 and IDENT(SonRise)<0.7 it would be able to force epenthesis into CC clusters as long as the resulting change in SonRise is less than 0.8 – i.e. in obstruent-sonorant clusters, but not in obstruent-obstruent clusters.

This line of analysis still depends on the perceptually-based IDENT(SonRise)<*n* constraints to derive TR epenthesis, but calls into question the need for distinctiveness constraints based on the same scale, MINDIST=SonRise:*d*. However, there are reasons to favor the analysis according to which epenthesis is motivated by MINDIST=SonRise:0.8, based on the range of variation in patterns of TR epenthesis observed so far. Varying the relative ranking of MINDIST=SonRise:*d* constraints and IDENT(SonRise)<*n* constraints derives variation in the threshold of Sonority Rise distance below which C_1C_2 - C_1VC_2 contrasts must be neutralized. As we have seen, the ranking in (30) derives a threshold distance of 0.8, so epenthesis applies to all obstruent-sonorant clusters. The ranking in (37) establishes a threshold of 0.6, which yields epenthesis into obstruent-liquid clusters only, as in Late Latin.

(37) MINDIST = SonRise:0.6 >> IDENT(SonRise)<0.5 >> MINDIST = SonRise:0.7

Based on the small sample of attested TR epenthesis processes, languages do differ in how low they set the distance threshold below which epenthesis applies (Steriade 1990, Hall 2003). On the other hand, any plausible development of the analysis according to which TR epenthesis is motivated by the preference for more distinct place contrasts will predict a much richer typology of patterns of epenthesis, including patterns that are not attested. This is because the distinctiveness of major place contrasts in C_1 position of a C_1C_2 cluster vary significantly depending on the manner of the consonants involved. For example Δ(pr, kr) is plausibly greater than Δ(pn, kn) because some cues to the place contrast can be realized on a liquid like [r], e.g. in its formant structure (Flemming & Jones 2006), but not on a following nasal. Accordingly the need for vowel epenthesis to enhance place distinctiveness is greater in stop+nasal clusters than in stop+liquid clusters – i.e. a [pn] vs. [kn] contrast violates higher-ranked MINDIST constraints than [pr] vs. [kr]. This leads to the prediction that epenthesis could target stop-nasal clusters only – stop-liquid clusters are not targeted because they are already sufficiently distinct (i.e. the increase in distinctiveness is not worth the violation of faithfulness). This pattern is not attested, suggesting that constraints like MINDIST = {burst & rel. trans} never rank high enough to force epenthesis on their own.

6.2 V- Ø similarity in vowel deletion

The framework developed here also predicts that the same scale of C_1C_2 and C_1VC_2 similarity should govern both vowel epenthesis and vowel deletion processes. I.e. both vowel deletion and vowel epenthesis place C_1C_2 and C_1VC_2 sequences in correspondence, only the direction of the mapping differs: $C_1C_2 \rightarrow C_1VC_2$ or $C_1VC_2 \rightarrow C_1C_2$. Accordingly, the same scale of perceptual distance should regulate both types of process, so vowel deletion should incur a smaller violation of faithfulness in the same contexts as vowel epenthesis. Patterns of schwa deletion in English and Dutch show that these processes do obey the same similarity scale as vowel epenthesis into consonant clusters.

Dutch schwa deletion is an optional process that deletes schwa between an obstruent and a liquid (38a) (Booij 1995:128ff.) – i.e. where the consonant cluster that results from deletion has a large sonority rise. Deletion is not possible with smaller sonority rises (38b) or with level or falling sonority (38c)[8]. So deletion is possible where the resulting C_1C_2 cluster is sufficiently similar to the C_1VC_2 input, specifically the maximum

[8] Booij (1995) does not cite any examples of this type, but his rule implies that deletion is unacceptable in the words in (38c). Pronunciations are from CELEX (Baayen, Piepenbrock & Gulikers 1995).

acceptable sonority rise distance between them is 0.5 (39). The unacceptability of deletion in cases like [rɑmələn] (38b) might be attributed to phonotactic constraints since [-ml-] does not generally occur morpheme-internally, except in loans (*Kremlin, umlaut*), but the this line of analysis will not account for the other cases where schwa cannot be deleted because the clusters that would result are well-formed, e.g. -kn- [akne] 'acne', -ln- [kɛlnər] 'waiter', -ns- [ɪnsɛkt] 'insect', -rn- [stornɪs] 'disorder'.

(38)	a.	supələ	~ suplə	'smooth'
		ɔfərən	~ ɔfrən	'to sacrifice'
		bɪbərən	~ bɪbrən	'to tremble'
	b.	tekənən	~ *teknən	'to draw'
		rɑmələn	~ *rɑmlən	'to rattle'
	c.	vølənən	~ *vølnən	'foals'
		vɔnəsən	~ *vɔnsən	'verdicts'
		xəzworənə	~ *xəzwornə	'juror'

(39)

	C_1C_2	Rise	C_1V	Rise	Distance
	plə	3	**pələ**	5	0.4
deletion OK	**frə**	2	**fərə**	4	0.5
no deletion	**knə**	2	**kənə**	5	0.6
	mlə	1	**mələ**	3	0.67

Schwa deletion only applies where the following vowel is also unstressed, so deletion is not possible in bəˈloven/*blovən 'to promise' (Booij 1995:130, van Oostendorp 2000:202), so it serves to eliminate a stress lapse. Accordingly we analyze schwa deletion as being motivated by *LAPSE, forbidding adjacent unstressed syllables, and limited by IDENT(SonRise)<*n* constraints. As shown in (40), *LAPSE is variably ranked with respect to IDENT(SonRise)<0.4, so deletion is possible in stop_liquid contexts. Deletion is never acceptable where it results in a change in Sonority Rise of 0.6 or greater, as in the nasal_liquid context exemplified in (41), because IDENT(SonRise)<0.6 ranks above *LAPSE.

(40)

	supələ	IDENT (SonRise) <0.6	*LAPSE	IDENT (SonRise) <0.4
a.	☞ súpələ		*	
b.	☞ súplə			*

(41)

	rɑmələn	IDENT (SonRise) <0.6	*LAPSE	IDENT (SonRise) <0.4
a.	☞ rámələn		*	
b.	rámlən	*		*

Optional schwa deletion in English follows a similar pattern. Hooper (1978) describes a dialect in which a basic constraint on schwa deletion is that it can only apply where the resulting cluster has rising sonority and a sonorant in second position (42). Schwa vowels can be reduced to the point of apparent deletion in a wider range of contexts in casual speech, but the rising sonority contexts were those where full deletion was judged to be acceptable even in less casual speech.

(42)	a.	ˈɛvəɹi	~ ˈɛvɹ̩i	‘every’
		ˈmɛməɹi	~ ˈmɛmɹi	‘memory’
		ˈpɝsənəl	~ pɝsnəl	‘personal’
		ˈfæməli	~ ˈfæmli	‘family’
	b.	ˈpɪkətɪŋ	~*ˈpɪktɪŋ	‘picketing’
		kəˈpæsəti	~*kəˈpæsti	‘capacity’

The metrical context for deletion is similar to Dutch as well: schwa only deletes before another unstressed syllable, e.g. [ˈmɛməˌɹaɪz]/*[ˈmɛmˌɹaɪz], and schwa is only deleted in word-initial pre-stress syllables (e.g. /bəloʊ/ → [bloʊ] ‘below’) in ‘very rapid and casual styles’ (Hooper 1978:198). This restriction on pre-stress deletion is confirmed by Patterson, LoCasto & Connine’s (2003) corpus study of schwa deletion: they found a much lower frequency of pre-stress schwa deletion compared to post-stress deletion. So the process can also be motivated by *LAPSE, differing only in that this constraint can rank above IDENT(SonRise)<0.75, permitting deletion in a wider range of environments than in Dutch.

(43)

	C_1C_2	Rise	C_1V	Rise	Distance
	pl	3	**pə**l	5	0.4
	fr	2	**fə**r	4	0.5
deletion OK	**kn**	2	**kə**n	5	0.6
	ml	1	**mə**l	3	0.67
	sn	1	**sə**n	4	0.75
no deletion	**ps**	1	**pə**s	5	0.8
	sk	-1	**sə**k	4	1.25

Previous analyses have accounted for the sonority conditions on schwa deletion by positing a requirement that the cluster that results from deletion must be a syllable onset (Booij 1995, Hooper 1978). The idea is to derive the requirement that the cluster resulting from schwa deletion must have rising sonority from the requirement that onset clusters must have rising sonority. This approach is problematic because the resulting clusters are not necessarily acceptable onsets in these particular languages, e.g. Dutch does not allow word-initial [tl], but [tl] clusters can be derived by schwa deletion: [kitələn]/[kitlən] ‘to tickle’. In English, [ml] is not an acceptable onset, but results from schwa deletion in ‘family’ (42). In addition, deletion can be blocked even where the result would be an acceptable onset: English allows s-stop onsets, but schwa deletion does not apply between [s] and a stop (42b). So both authors have to propose that constraints on possible onsets are modified specifically for the output of schwa deletion (Booij 1995:129, Hooper 1978:193). The analysis proposed here avoids this problem because it makes no reference to syllabification, the sonority-related constraints on deletion are derived purely from universals of C_1C_2 - C_1VC_2 similarity.

7. The typology of epenthesis into consonant clusters

We have seen that vowel epenthesis does not serve as an alternative to assimilation in repairing marked heterorganic clusters or mixed-voicing clusters. These generalizations raise the question of what kinds of clusters are targeted by vowel epenthesis. In the previous section we have seen that epenthesis can single out obstruent-sonorant clusters, leaving obstruent-obstruent clusters intact. This pattern arises because the difference between V and Ø is smaller in the obstruent-sonorant context. In this section we review additional patterns of vowel epenthesis into consonant clusters in order to situate the occurring and non-occurring patterns of epenthesis discussed above in relation to the broader typology of epenthesis. The goals are to show that other patterns of epenthesis into consonant clusters can be analyzed in terms of additional constraints that are consistent with system of constraints proposed so far, and to show how the unattested patterns of epenthesis differ from these additional epenthesis processes.

In one widely attested pattern, epenthesis applies to all consonant clusters that exceed a certain length, regardless of the properties of the consonants that make up the cluster. This pattern is exemplified by Yawelmani Yokuts (Newman 1944, Kisseberth 1970) where epenthesis breaks up medial clusters of any three consonants (44a) and final clusters of two consonants (44b), ensuring that every consonant is adjacent to a vowel. Similar patterns are observed in Cairene Arabic (Broselow 1976), and Lenakel (Lynch 1978).

(44) Yawelmani Yokuts (Newman 1944)

a.	/paʔt-mi/ →	paʔitmi	'having fought'	cf.	paʔt-al	'might fight'
	/lihm-mi/ →	lihimmi	'having run'	cf.	lihm-al	'might run'
	/ʔilk-mi/ →	ʔilikmi	'having sung'	cf.	ʔilk-al	'might sing'
b.	/paʔt/ →	paʔit	'fighting'			
	/lihm/ →	lihim	'running'			
	/ʔilk/ →	ʔilik	'singing'			

This pattern is quite different from TR epenthesis since it is insensitive to the sonority contour of consonant clusters. It also cannot be motivated by constraints on the distinctiveness of consonant contrasts since neutralization is generally more faithful than epenthesis as a repair for insufficiently distinct consonant contrasts, so it is clear that additional constraints are required to account for this pattern.

Following Côté (2000:158), we propose that the relevant constraint is C↔V, which requires every consonant to be adjacent to a vowel. This constraint motivates epenthesis into medial $–C_1C_2C_3$- and final $–C_1C_2$ clusters, since in each case C_2 is not adjacent to a vowel. The constraint C↔V can be understood as expressing a preference that the presence of each consonant should be marked by the acoustic discontinuities that are generated at the transition from a vowel to a consonant or vice versa (cf. Stevens 1998:245f.). If this constraint is sufficiently highly ranked with respect to faithfulness constraints, then epenthesis applies in all consonant clusters of sufficient size regardless of their sonority profile.

If C↔V ranks in the hierarchy of IDENT(SonRise) constraints, we expect to find patterns where the application of epenthesis is influenced by the sonority profile of consonant clusters. This is the case in Lebanese Arabic where epenthesis generally breaks up final CC clusters, but applies only variably in falling sonority clusters (Haddad 1984), i.e. where epenthesis is a greater violation of faithfulness. Similarly in Bedouin dialects of the Northern Sinai Littoral, a vowel is generally epenthesized after the first consonant of a $–C_1C_2C_3$- cluster, but this process is variable where sonority falls from C_1 to C_2 (de Jong 2000:125f.).

Another widespread type of epenthesis into consonant clusters applies only adjacent to sonorant consonants. This can be exemplified from a number of Salishan languages, including Montana Salish (Flemming, Ladefoged & Thomason 2008), where a schwa vowel is epenthesized between a sonorant and a preceding consonant in almost all contexts, e.g. [tʃ’ɬtsts**ə**ˀ**m**u] ‘bead, beads’. Similar patterns are found in Kalispel (Vogt 1940), Spokane (Bates & Carlson 1992), Shuswap (Kuipers 1974) and Thompson (Thompson & Thompson 1992). Although this pattern of epenthesis is sensitive to sonority, it is unlike TR epenthesis in that it applies even if the preceding consonant is more sonorous, as in [wáˀlə**wə**ˀ**l**] 'long-billed curlew', where the second schwa is epenthesized between a glide and a liquid. Hall (2003) notes other epenthesis processes that apply adjacent to high sonority consonants. For example, Dutch has an optional process of schwa epenthesis that applies to word-final liquid-C clusters, e.g. [mɛlk]~[mɛlək] 'milk', [ɑrm]~[ɑrəm] 'arm' (Booij 1995:127).

These processes suggest that in addition to sonority rise distance, there is another dimension of perceptual distance between V and Ø which is related to the absolute sonority of adjacent consonants. It is plausible that epenthesis is a lesser violation of faithfulness adjacent to a sound that is closer in sonority to a vowel. These epenthesis processes can be regarded as a form of 'fission' in which a sonorant consonant splits into two sonorant sounds, a vowel and a consonant, whereas an epenthetic vowel between obstruents lacks any sonorant correspondent in the input.

Neither of these attested patterns of epenthesis resembles the unattested patterns that are the focus of this paper: epenthesis that specifically targets nasal-obstruent clusters, and epenthesis that breaks up mixed-voicing obstruent clusters. These unattested processes cannot be characterized as epenthesis into clusters that exceed a certain length since they target specific types of clusters. Epenthesis into mixed-voicing obstruent clusters obviously would not constitute epenthesis adjacent to a sonorant. A process that epenthesizes vowels adjacent to sonorants could apply to nasal-C clusters and would not apply to obstruent-obstruent clusters, but if it applies adjacent to nasals it would also be expected to apply adjacent to more sonorant consonants, e.g. liquids. So it is not possible to derive a pattern that only targets nasal-obstruent clusters on this basis either.

Finally we turn to a pattern of epenthesis which appears to be problematic for the hypothesis that perceptually minimal repairs are preferred, since an illegal cluster is broken up by epenthesis when deletion of a consonant would be expected to be a smaller perceptual change. This pattern is exemplified by alternation in the forms of the English plural /-z/ and past tense /-d/ suffixes and by a similar alternation in Lithuanian verb prefixes (Baković 2005). We will see that these patterns are morphologically restricted in crucial ways that can explain why the minimal repair is eschewed in these cases.

The English /-z/ suffixes are realized as [-əz] after stridents [s, z, ʃ, ʒ, tʃ, dʒ], e.g. *dog*[z] but *mesh*[əz]. Epenthesis of schwa avoids an ill-formed cluster of two stridents, e.g. **me*[ʃz]. It has been suggested that this cluster is ill-formed because it violates an OCP constraint (e.g. Yip 1988), or, according to Baković (2006), because English has processes of voicing and anteriority assimilation that make a geminate strident the expected outcome if epenthesis did not apply, e.g. *me*[ʃʃ], and geminates are not permitted in this context in English. Either way, we have seen that epenthesis between two obstruents is a perceptually significant change, so we might expect that deletion of the second fricative should be preferred as the more faithful repair, i.e. /mɛʃ+z/→[mɛʃ].

In fact deletion is the attested repair for sequences of stridents in all other contexts in English, e.g. *dissimilar* /dɪs+sɪmɪlɚ/ → [dɪsɪmɪlɚ], *[dɪsəsɪmɪlɚ]. The avoidance of deletion in the particular case of the /-z/ suffix is plausibly motivated by the fact that this would completely delete the suffix, resulting in homophony between singular and plural. In other words, epenthesis is motivated by a higher-ranked constraint requiring that inflected forms of the same stem should be distinct although it is less faithful than deletion (Crosswhite 1999, Trón & Rebrus 2005, Löfstedt 2008).

The same line of analysis applies to the past tense suffix, /-d/. This is realized as [-əd] after other coronal stops [t, d], e.g. *walk*[t] but *wait*[əd], avoiding an ill-formed sequence of coronal stops. Again, deletion would be more faithful than vowel epenthesis but is blocked by morphological distinctiveness constraints.

A similar pattern of epenthesis is observed in Lithuanian (Ambrazas 1997, Bakovic 2005): the verb prefix /ap-/ is realized as [api-] before bilabials [b, p] (*ap^{j}i-b^{j}ek^{j}t^{j}i* 'to run around', cf. *ap-kal^{j}b^{j}et^{j}i* 'to slander', while /at-/ is realized as [ati] before [t, d] (*at^{j}i-d^{j}et^{j}i* 'to delay, put off, postpone', cf. *at-ras^{j}t^{j}i* 'to find, discover'). Lithuanian has regressive voicing assimilation in obstruent clusters, so a geminate would be expected where /at-/ is prefixed to [t] or [d]. As in English, Lithuanian disallows geminates, so Bakovic attributes epenthesis to a constraint against geminates. Again, the question arises why epenthesis is preferred over degemination here.

Degemination is the regular repair for geminates in Lithuanian (Kenstowicz 1972:21). Assimilation and degemination both apply to other prefixes such as /iʃ-/, /uʒ-/, e.g. /iʃ+ʃoːkti/→[iʃoːk^{j}t^{j}i] 'jump out', /uʒ+sienis/ → [us^{j}ienis] 'abroad' (Mathiassen 1996:26), so the V- Ø alternations observed with /ap-/ and /at-/ are exceptional. One possible line of analysis would be to treat this as a case of phonological selection of listed allomorphs (cf. Kager 1996, Bonet, Lloret & Mascaró 2007). That is, each prefix has two listed allomorphs, /ap-/, /api-/, /at-/, /ati-/, and they compete to appear in a word that is morphologically specified for the relevant prefix. The disyllabic form is selected where it permits avoidance of consonant deletion through degemination, otherwise the shorter form is preferred.

8. Conclusions: Limitations on positional enhancement

This paper has provided evidence for the proposal that both markedness and faithfulness constraints refer to the same scales of perceptual distance. Distinctiveness constraints are markedness constraints that penalize contrasts according to the perceptual distance between the contrasting sounds: contrasts are more marked if the perceptual distance between the sounds is smaller. Faithfulness constraints evaluate correspondence between input and output forms in terms of perceptual distance: greater distances between corresponding sounds incur greater violations of faithfulness. This framework provides the basis for an account of generalizations about when vowel epenthesis does and does not break up consonant clusters. We have seen that vowel epenthesis can target rising sonority clusters, leaving obstruent-obstruent clusters intact, as in Winnebago. This pattern follows from the generalization that the perceptual distance between C_1C_2 - C_1VC_2 sequences is smaller where the sonority rise from C_1 to C_2 is larger. As a result, contrasts between obstruent-sonorant and obstruent-V-sonorant sequences are marked because they are relatively indistinct, and, for the same reason, neutralizing these contrasts is a modest violation of faithfulness. On the other hand, vowel epenthesis does not serve as an alternative to place and voicing assimilation in consonant clusters, although inserting a vowel could eliminate the same clusters that undergo these assimilation processes. This is because epenthesis in these contexts constitutes a larger perceptual change than assimilation. We observe in this concluding section that these non-occurring patterns of epenthesis can be viewed as instances of a broader class of positional enhancement processes, and that the analysis predicts limitations on this class of processes.

Positional enhancement processes are expected counterparts to positional neutralization processes, given the line of analysis pursued here. That is, according to the Dispersion Theory of Contrast (Flemming 2002, 2004), positional neutralization is analyzed as elimination of contrasts in contexts where they would be insufficiently distinct, so positional neutralization is driven by distinctiveness constraints, as exemplified in the analyses of place and voice neutralization above. But in principle there are two ways that a contrast that violates a MINDIST constraint could be changed to satisfy that constraint: one is to eliminate the contrast, i.e. neutralization, and the other is to modify the contrast to make it more distinct, i.e. enhancement. As discussed above, epenthesis into heterorganic nasal-C clusters and mixed-voicing obstruent clusters could serve to improve the distinctiveness of insufficiently distinct contrasts, and thus would be examples of positional enhancement.

However we have seen no evidence that vowel epenthesis functions as a positional enhancement of place or obstruent voicing contrasts. According to the analysis proposed here, this is because correspondence constraints impose a preference for perceptually minimal modification of the input and, in the cases examined, vowel epenthesis is a greater violation of faithfulness than neutralization of place or voice contrasts. The broader prediction of this analysis is that positional enhancement of any kind will only be preferred over positional neutralization if it is the more faithful option. This is a restrictive requirement because contrasts that are at risk of neutralization are necessarily relatively indistinct – i.e. they violate high-ranking distinctiveness constraints – so neutralization of these contrasts always involves a relatively small change, and the threshold for faithfulness of a viable enhancement is thus set relatively low. The hypothesis that markedness and faithfulness constraints refer to the same scales of perceptual distance establishes this connection between the nature of the markedness problem and preferred repairs for that problem: If the perceptual distance $\Delta(a, b)$ between two sounds is small then contrasts between *a* and *b* are marked, and neutralizing a contrast between *a* and *b* is a modest violation of faithfulness.

Specifically, we predict that vowel epenthesis should be very limited in its occurrence as a positional enhancement in spite of the fact that vowel epenthesis could in principle enhance the distinctiveness of the wide range of consonant contrasts that are better realized adjacent to, and preferably preceding, a vowel (Wright 2004, Côté 2000). This is because vowel epenthesis is a significant perceptual change, particularly in the contexts where consonant contrasts are less distinct. That is, many patterns of positional neutralization of consonant contrasts arise where the consonant contrast relies on external or transitional cues for its distinctiveness, as with the obstruent voicing contrasts and stop and nasal place contrasts discussed above (sections 3 and 4). Most external cues, such as stop bursts, formant transitions and rate of change of intensity are best realized with an open vocal tract, i.e. with an adjacent vowel, and are lost as the constriction of the surrounding sounds narrows, so the context adjacent to an obstruent in general offers the fewest external cues (Wright 2004). Consequently place and voice contrasts are commonly neutralized in pre-obstruent contexts. But we have seen that the perceptual change incurred by epenthesizing a vowel is greatest in clusters where the second member is an obstruent – i.e. where there is no significant sonority rise in the cluster. Contexts where vowel epenthesis would be a lesser violation of faithfulness, e.g. obstruent-sonorant clusters, are not usually the contexts where consonantal contrasts are particularly indistinct – e.g. voicing contrasts can be realized by VOT differences in this context, so vowel epenthesis is not necessary as an enhancement.

This prediction is supported by preliminary examination of additional patterns of positional neutralization. For example, Mohanan (1993) and Jun (2004) provide evidence for a universal of place neutralization: if non-coronal consonants assimilate in place to a following consonant, then coronal consonants do also, but not vice versa. Jun (2004) argues that this reflects the lesser distinctiveness of place contrasts involving coronals in pre-obstruent contexts compared to place contrasts among non-coronals. Vowel epenthesis does not serve as a positional enhancement for these coronal place contrasts: there are no epenthesis processes that specifically target coronal-obstruent clusters. A second example comes from the distribution of contrasts between apical alveolar and retroflex consonants. Steriade (1995, 2001b) shows that these contrasts are frequently neutralized where there is no preceding vowel, i.e. in word-initial and post-consonantal contexts, because F3 and F4 closure transitions are crucial cues to this contrast, but cannot be realized in the absence of a preceding vowel. These problematic contrasts could be enhanced by epenthesis of a preceding vowel but epenthesis preceding only apical consonants is not attested.

Positional enhancement is predicted to be more feasible in contexts where the distinctiveness of a cue is reduced, rather than a set of cues being completely lost. Here we consider a case of this kind involving post-nasal voicing contrasts. The distinctiveness of closure voicing differences between voiced and voiceless obstruents tends to be reduced where the obstruents follow a nasal, which can result in neutralization of voicing contrasts. However, in this case we also find positional enhancement: the voicing contrast can be made sufficiently distinct by aspirating the voiceless stops. This enhancement is viable because it involves a relatively small perceptual change to the input, namely an increase in VOT.

Post-nasal voicing is a cross-linguistically common process whereby voiceless stops are voiced after nasals (Pater 199, Hayes 1999), and it can result in the neutralization of voicing contrasts, e.g. in Kikuyu, Ki-Nande and Bukusu (Hyman 2001, Flemming 2001:14). Hayes (1999) argues that the tendency to voice stops after nasals results from the relative articulatory difficulty of devoicing a stop in this context. Devoicing of a stop is usually facilitated by the rise in oral pressure that occurs following the formation of the closure because rising oral pressure reduces the pressure drop across the glottis below the threshold required for vocal fold vibration (Ohala 1983, Westbury & Keating 1984). But following a nasal, velum lowering is liable to persist into the stop, and the lowered velum allows air to flow out through the nose, slowing the rise in oral pressure. So, unless additional measures are taken, part of the closure of a voiceless stop following a nasal will be voiced, making it perceptually similar to a voiced stop. The lower distinctiveness of the voicing contrast can motivate neutralization.

A number of Bantu languages show a different pattern of laryngeal modification following nasals: voiceless stops are aspirated in this context, although they are unaspirated elsewhere, as in Kongo (45) (Hyman 2003:50, Meinhof 1932:159[9]).

(45)	plural, class 10	sing., class 11	
	ŋ-kʰuni	lu-kuni	'fire wood'
	n-tʰombo	lu-tombo	'shoot'

As Pater (1999) observes, aspiration of post-nasal voiceless stops is plausibly analyzed as an enhancement, in that it serves to maintain the voicing contrast in a context where stops are commonly subject to voicing. Aspiration can enhance this contrast because it is associated with wider glottal opening than in unaspirated stops (Dixit 1987) and sufficient glottal opening prevents vocal fold vibration even if oral pressure is not raised (Stevens 1998:81). The long VOT characteristic of aspirated stops is probably a necessary concomitant of wide glottal opening since the time it takes to open the glottis wide and then return it to a width at which voicing can resume is plausibly longer than the duration of stop closure (Stevens 1998:42). However, this increased VOT could constitute an enhancement in its own right, increasing the difference from the short VOT of the contrasting voiced stop, offsetting any reduction in the difference in the duration of closure voicing. Post-nasal aspiration is an acceptable alternative to post-nasal voicing neutralization because the resulting increase in VOT of voiceless stops is a modest perceptual change.

References:

Ambrazas, V. (ed.),1997. Lithuanian Grammar. Baltos Lankos, Vilnius.

Baayen, R.Harald, Richard Piepenbrock & Leon Gulikers (1995). *The CELEX lexical database (Release 2)*. Linguistic Data Consortium, University of Pennsylvania.

Baković, Eric (2000). *Harmony, Dominance and Control*. PhD dissertation, Rutgers University, New Jersey.

Baković, Eric (2005). Antigemination, assimilation and the determination of identity. *Phonology* 22, 279-315.

[9] Meinhof does not mark aspiration in his transcriptions, but post-nasal aspiration is described on p.158.

Baković, Eric (2006). Partial identity avoidance as cooperative interaction. Michael Temkin Martínez, Asier Alcázar & Roberto Mayoral Hernández (eds.) *Proceedings of the thirty-third Western Conference on Linguistics*, 13-25. California State University, Fresno.

Bates, Dawn, & Barry F. Carlson (1992) Simple syllables in Spokane Salish. *Linguistic Inquiry*. 23, 653-59.

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

Bonet, Eulàlia, Maria-Rosa Lloret & Joan Mascaró (2007). Allomorph selection and lexical preferences: Two case studies. *Lingua* 117, 903-927.

Booij, Geert (1995). *The Phonology of Dutch*, Oxford University Press.

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

Byrd, Dani (1992). Perception of assimilation in consonant clusters : a gestural model. *Phonetica* 49. 1–24.

Clements, G. N. (1990). The role of the sonority cycle in core syllabification. John Kingston & Mary E. Beckman (eds.) *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. 283-333. CUP, Cambridge.

Côté, Marie-Hélène (2000). *Consonant Cluster Phonotactics: A Perceptual Approach*. Ph.D. dissertation, MIT.

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

Davis, Stuart (1998). Syllable contact in Optimality Theory. *Journal of Korean Linguistics* 23, 181-211.

Dayley, Jon P. (1989). *Tümpisa (Panamint) Shoshone grammar*. University of California Press, Berkeley.

de Jong, Rudolf (2000). *A Grammar of the Bedouin Dialects of the Northern Sinai Littoral*. Brill, Leiden.

de Lacy, Paul V. (2002). *The Formal Expression of Markedness*. PhD dissertation, University of Massachusetts, Amherst.

Devine, Andrew M., & Laurence D. Stephens (1977). *Two Studies in Latin Phonology*. Anma Libri, Saratoga, CA.

Dilley, L. and Pitt, M. (2007). A study of regressive place assimilation in spontaneous speech and its implications for spoken word recognition. *Journal of the Acoustical Society of America, 122(4)*, 2340-2353.

Dixit, Prakash (1987). Mechanisms for voicing and aspiration. *UCLA Working Papers. in Phonetics* 67. 49–102.

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

Fleischhacker, Heidi (2005). *SImilarity in Phonology: Evidence from Reduplication and Loan Adaptation*. PhD dissertation, University of California, Los Angeles.

Flemming, Edward (2001). Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* 18, 7-44.

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

Flemming, Edward (2004). Contrast and perceptual distinctiveness. Hayes, B., Kirchner, R. & Steriade, D. (eds.) *Phonetically-Based Phonology*, 232—276. Cambridge: Cambridge University Press.

Flemming, Edward, & Patrick J. Jones (2006). Cues to place in stop-liquid clusters. Poster presented at 152nd Meeting of the Acoustical Society of America, Honolulu, Hawaii.

Flemming, Edward, Peter N. Ladefoged, & Sarah G. Thomason (2008). Phonetic Structures of Montana Salish. *Journal of Phonetics* 36, 465-491.

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.

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.

Gescheider, George A. (1997). *Psychophysics: The Fundamentals*, 3rd Edition. Lawrence Erlbaum Associates, Mahwah, NJ.

Gouskova, M., 2002. Relational hierarchies in OT. In: Mikkelsen, L., Potts, C. (Eds.), Proceedings of the West Coast Conference in Linguistics 21, pp. 113-126.

Gow, D. (2003) Feature parsing: Feature cue mapping in spoken word recognition. *Perception & Psychophysics*, 65(4), 575-590.

Haddad, Ghassan F. (1984). Epenthesis and sonority in Lebanese Arabic. *Studies in the Linguistics Sciences* 14, 57-88.

Hall, Nancy E. (2003). *Gestures and Segments: Vowel Intrusion as Overlap*. PhD dissertation, University of Massachusetts, Amherst.

Hayes, Bruce, & Donca Steriade (2004). Introduction: the phonetic bases of phonological markedness. Bruce Hayes, Robert Kirchner, & Donca Steriade (eds.) *Phonetically Based Phonology*. Cambridge University Press, Cambridge.

Harris, James (1983) *Syllable structure and stress in Spanish: A nonlinear analysis*. MIT Press, Cambridge, MA.

Henderson, Janette B., & Bruno H. Repp (1982). Is a stop consonant released when followed by another stop consonant? *Phonetica* 39, 71-82.

Hooper, Joan B. 1978. Constraints on schwa deletion in American English. In Jacek Fisiak (ed.). *Recent developments in historical phonology*. The Hague: Mouton. 183-207.

Hooper, Joan B. 1978. Constraints on schwa deletion in American English. In Jacek Fisiak (ed.). Recent developments in historical phonology. The Hague: Mouton. 183-207.

Hyman, Larry (2003). Segmental phonology. Derek Nurse & Gérard Philippson (eds.) *The Bantu Languages*. Routledge, New York.

Itô, Junko, & Armin Mester (2007). Systemic markedness and faithfulness. *CLS 39: Papers from the 39th Regional Meeting of the Chicago Linguistic Society 2003*, 665-689.

Johnson, Keith (2003). *Acoustic and Auditory Phonetics*. 2nd Edition. Blackwell, Oxford.

Jun, Jongho (1995). *Perceptual and Articulatory Factors in Place Assimilation: An Optimality-Theoretic Approach*. PhD dissertation, UCLA.

Jun, Jongho (2004). Place assimilation. Bruce Hayes, Robert Kirchner, and Donca Steriade (eds.) *The Phonetic Bases of Markedness*. Cambridge University Press, Cambridge. 58-86.

Kager, René (1996). On affix allomorphy and syllable counting. Ursula Kleinhenz (ed.), Interfaces in Phonology, 155-171. Akademie Verlag, Berlin.

Kawahara, Shigeto (2007a). Half rhymes in Japanese rap lyrics and knowledge of similarity. *Journal of East Asian Linguistics* 16, 113-144.

Kawahara, Shigeto (2007b). Copying and spreading in phonological theory: Evidence from echo epenthesis. L. Bateman, M. O'Keefe, E. Reilly & A. Werle (eds.) *UMOP 32: Papers in Optimality Theory III,* 111-144. GLSA, Amherst, MA.

Keating, Patricia, Wendy Linker & Marie Huffman (1983). Patterns in allophone distribution for voiced and voiceless stops. *Journal of Phonetics* 11, 277-290.

Kenstowicz, Michael (1972). Lithuanian Phonology. *Studies in the Linguistic Sciences* 2.2, 1-85.

Kewley-Port, Diane (1982). Measurement of formant transitions in naturally produced stop consonant-vowel syllables. *Journal of the Acoustical Society of America* 72, 379–389.

Kisseberth, Charles (1970). On the functional unity of phonological rules. *Linguistic Inquiry* 1, 291-306.

Kuipers, Aert H. (1974) *The Shuswap Language*. The Hague: Mouton.

Kurowski, Kathleen M., & Sheila Blumstein (1993). Acoustic properties for the perception of nasal consonants. Marie K. Huffman & Rena A. Krakow (eds) *Nasals, Nasalization, and the Velum. Phonetics and Phonology, vol. 5*, 147-167. Academic Press, San Diego.

Legman, G. (1969). *The Limerick: 1700 examples, with notes variants and index*. Bell Publishing Co., New York.

Lisker, Leigh, and Arthur S. Abramson (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20, 384-422.

Löfstedt, Ingvar (2008). Perceptually driven allomorphy: The Swedish non-neuter definite article. To appear in *NELS 38: Proceedings of the 38th Meeting of the North East Linguistics Society*,

Lombardi, Linda (1999). Positional faithfulness and voicing assimilation in Optimality Theory. *Natural Language & Linguistic Theory* 17, 267-302.

Lubowicz, Anna (2003). *Contrast Preservation in Phonological Mappings*. PhD dissertation, University of Massachusetts, Amherst.

Lynch, John (1978). *A Grammar of Lenakel*. Pacific Linguistics, Series B, no. 55. Asutralian National University, Canberra.

Malécot, André (1956). Acoustic cues for nasal consonants: An experimental study involving tape-splicing technique. *Language* 32, 274-284.

Magen, Harriet S. (1998). The perception of foreign-accented speech. *Journal of Phonetics,* 26, 381-400.

Mathiassen, Terje (1996). *A Short Grammar of Lithuanian*. Slavica, Columbus, Ohio.

McCarthy, John J. (2003). On targeted constraints and cluster simplification. *Phonology* 19, 273-292.

McDonough, Joyce (2003). *The Navajo Sound System*. Kluwer, Boston, MA.

Meinhof, Carl (1932) *Introduction to the Phonology of the Bantu Languages*. Translated by N.J. v. Warmelo. Dietrich Reimer, Berlin.

Miner, Kenneth (1979). Dorsey's Law in Winnebago-Chiwere and Winnebago accent. *International Journal of American Linguistics* 45, 25-33.

Miner, Kenneth (1993). On some theoretical implications of Winnebago phonology. *Kansas Working Papers in Linguistics* 18, 111-130.

Mohanan, 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.

Mohanan, K.P., and Tara Mohanan (1984). Lexical phonology of the consonant system in Malyalam. *Linguistic Inquiry* 15, 575-602.

Myers, Scott (2002). Gaps in factorial typology: The case of voicing in consonant clusters. Ms, University of Texas, Austin.

Newman, Stanley (1944). *The Yokuts Language of California*. The Viking Fund Publications in Anthropology 2, The Viking Fund, New York.

Ní Chiosáin, Máire & Jaye Padgett (2008). Contrast, comparison sets, and the perceptual space. Steve Parker (ed.) *Phonological Argumentation: Essays on Evidence and Motivation*. Equinox, London.

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.

Parker, S. (2008), Sound level protrusions as physical correlates of sonority, *Journal of Phonetics*, 36, pp. 55-90.

Pater, Joe (1999). Austronesian nasal substitution and other NC effects. René Kager, Harry van der Hulst, & Wim Zonneveld (eds.) *The Prosody-Morphology Interface*. 310-343. CUP, Cambridge.

Patterson, David, Paul C. LoCasto, & Cynthia M. Connine (2003). Corpora analyses of frequency of schwa deletion in conversational American English. *Phonetica* 60, 45-69.

Petrova, Olga, Rosemary Plapp, Catherine Ringen & Szilárd Szentgyörgyi (2006). Voicing and aspiration: evidence from Russian, Hungarian, German, Swedish, and Turkish. *The Linguistic Review* 23, 1-35.

Raphael, L.J. (1972). Preceding vowel duration as a cue to the perception of the voicing characteristics of word-final consonants in American English. *Journal of the Acoustical Society of America*, 51, 296-1303.

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.

Repp, Bruno H. (1979). Relative amplitude of aspiration noise as a voicing cue for syllable-initial stop consonants. *Language and Speech* 22, 173-189.

Repp, Bruno H., & Hwei-Bing Lin (1989). Acoustic properties and perception of stop consonant release transients. *Journal of the Acoustical Society of America* 85, 379-396.

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.

Rubach, Jerzy (1993). *The Lexical Phonology of Slovak*. OUP, New York.

Rubach, Jerzy (1996). Nonsyllabic analysis of voice assimilation in Polish. *Linguistic Inquiry* 27, 69-110.

Sapir, J. David (1965). *A Grammar of Diola-Fogny, a language spoken in the Basse-Casamance region of Senegal*. Cambridge University Press, Ibadan.
Shaw, Patricia A. (2008). Constraints on the sequencing and syllabification of obstruents. Handout from CUNY Conference on the Syllable, CUNY, New York.
Siptár, Péter & Miklós Törkenczy (2000). *The Phonology of Hungarian*. OUP, New York.
Steriade, Donca (1990). Gestures and autosegments: comments on Browman and Goldstein's paper. John Kingston & Mary E. Beckman (eds.) *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. 382-397. CUP, Cambridge.
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 (2001a). 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.
Steriade, Donca (2001b). Directional asymmetries in place assimilation. Elizabeth Hume and Keith Johnson (eds.) *The Role of Speech Perception in Phonology*. Academic Press, New York.
Steriade, Donca (2003). Knowledge of similarity and narrow lexical override. *Proceedings of Berkeley Linguistics Society* 29, 583-598.
Steriade, Donca (2006). Contour correspondence: evidence from cluster interruption. Handout, OCP3, Budapest.
Stevens, Kenneth N. (1998). *Acoustic Phonetics*. MIT Press, Cambridge.
Surprenant, Aimée M. & Louis Goldstein (1998). The perception of speech gestures. *JASA* 104. 518–529.
Thompson, L.C., & Thompson, M.T. (1992) *The Thompson Language*, *University of Montana Occasional Papers in Linguistics* 8.
Titze, Ingo, Sheila S. Schmidt & Michael R. Titze (1995). Phonation threshold pressure in a physical model of the vocal fold mucosa. *Journal of the Acoustical Society of America* 95, 3080-3084.
Trón, Viktor, & Péter Rebrus (2005). Re-presenting the past: Contrast and uniformity in Hungarian past tense suffixation. Christopher Piñón & Péter Siptár (eds.) *Papers from the Düsseldorf Conference: Approaches to Hungarian* 9, Akadémiai Kiadó, Budapest.
Van Oostendorp, Mark (2000). *A Theory of Feature Content and Prosodic Structure*. Walter de Gruyter, Berlin.
Vance, Timothy J. (1987). *An Introduction to Japanese Phonology*. State University of New York Press, Albany, NY.
Venneman, T., 1988. Preference Laws for Syllable Structure and the Explanation of Sound Change: With Special Reference to German, Germanic, Italian, and Latin. Mouton de Gruyter, Berlin.
Vogt, H. (1940) *The Kalispel language*. Det Norsk Videnskaps-Akademi, Oslo.
Westbury, John, and Patricia A. Keating (1986). On the naturalness of stop consonant voicing. *Journal of Linguistics* 22, 145-66.

Wetzels, W. Leo, & Joan Mascaró (2001). The typology of voicing and devoicing. *Language* 77, 207-244.
Wilson, Colin (2001). Consonant cluster neutralization and targeted constraints. *Phonology* 18, 147-197.
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.
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.
Wright, Richard (2004). A review of perceptual cues and robustness. Hayes, B., Kirchner, R. & Steriade, D. (eds.) *Phonetically-Based Phonology*, 34—57. Cambridge: Cambridge University Press.
Yip, Moira (1988). The Obligatory Contour Principle and phonological rules: a loss of identity. *Linguistic Inquiry* 19, 65-100.
Zoll, Cheryl (1994). Subsegmental parsing: floating featuresin Chaha and Yawelmani. *Phonology at Santa Cruz* 3, 111-128.
Zuraw, Kie (2007). The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog. *Language* 83, 277-316.
Zwicky, Arnold (1976). Well, this rock and roll has got to stop: Junior's head is hard as a Rock, *Proceedings of Chicago Linguistics Society* 12, 676-697.
Zsiga, Elizabeth C. (1994). Acoustic evidence for gestural overlap in consonant sequences. *Journal of Phonetics* 22. 121–140.