Directional asymmetries in place assimilation: a perceptual account

Donca Steriade

1. Introduction

The project presented here seeks to explain observed regularities in the direction of place assimilation. The best known among these is the fact that assimilation proceeds regressively in intervocalic clusters composed of alveolars, palato-alveolars, labials or velars. This fact is consistent with a variety of interpretations, some of which are discussed below. However the range of analyses narrows down drastically once we observe that assimilation is consistently *progressive* in clusters composed of retroflexes and alveolars. These two observations are schematically illustrated below. Instances of each type are presented in the body of the paper.

(1) **Regressive** assimilation in VC_1C_2V;

$$C_1,C_2 \in \{(Palato)\text{-}Alveolar, Labial, Velar\}$$

(a)	{Alveolar, Labial}	anpa $\rightarrow$ ampa;	amta $\rightarrow$ anta
(b)	{Velar, Labial}	aŋpa $\rightarrow$ ampa;	amka $\rightarrow$ aŋka
(c)	{Alveolar, Velar}	anka $\rightarrow$ aŋka;	aŋta $\rightarrow$ anta
(d)	{Palatoalveolar, Labial}	anpa $\rightarrow$ ampa;	amtʃa $\rightarrow$ antʃa

(2) **Progressive** assimilation in VC_1C_2V; $C_1, C_2 \in \{Alveolar, Retroflex\}$

$$anṭa \rightarrow anta; \qquad anṭa \rightarrow anṭa$$

Assimilation is a form of contrast neutralization: it reduces, in a C_1C_2 sequence, the number of potential bearers of place features from two (each of C_1 and C_2) to one (the whole C_1C_2 cluster). Neutralization can also take non-assimilatory forms. Thus the [s]-[ʃ] contrast is neutralized pre-consonantally in English, as only [ʃ] occurs before [r] and only [s] occurs before obstruents, nasals and [l]. This is in part a non-assimilatory effect: the fricative in [sp], [sm] and [sk] cannot contrast for anteriority but remains unassimilated to the following C. Similarly, in Chumash (Poser 1982), the [s]-[ʃ] contrast is reduced to [ʃ] before the non-strident coronals [t,l,n]: clusters like [ʃn] have an unassimilated, but place-neutralized first member.

Place assimilation patterns correspond to patterns of non-assimilatory place neutralization. Contrasts between alveolars, labials, velars and palatoalveolars (abbreviated here as *major C-place contrasts*) are typically neutralized in preconsonantal or domain-final position, i.e. **in consonants not followed by vowels**. Thus the pre-consonantal C_1 in the clusters in (1) is not only the typical target of assimilation, as in (1), but also, in non-assimilating cases, the typical target of place neutralization. This is the pattern observed above for English and Chumash fricatives: pre-consonantal strident fricatives neutralize, while prevocalic fricatives continue to contrast for anteriority. An extension of this pattern is the case in which word-final and pre-

consonantal C's are place neutralized: thus the alveolar-labial contrast between A. Greek [n] and [m] is eliminated through assimilation before a C, and in non-assimilatory fashion word finally, where only [n] is permitted.

Non-assimilatory neutralization also affects the contrast between apico-alveolars and retroflexes (referred to here as *the apical - or t/ʈ - contrast*). In this case however it is typically the domain initial and post-consonantal positions that are place-neutralized, that is the apicals **not preceded by a vowel.** Thus C_2 in the clusters in (2) is not only the typical target of apical place assimilation but also the position in which the apical contrast is generally neutralized. An instance of this sort is the distribution of apicals in the Muriṇbaṭa morpheme internal clusters (Hamilton 1996): in this language apicals contrast with retroflexes postvocalically (in V_# and V_C) but the t/ʈ contrast is suspended after C's. Postconsonantal apicals are uniformly realized as alveolars after non-apicals (e.g. [ŋd], [nt], [md]) and as homorganic with a preceding apical C (e.g. [nd], [ɳḍ]). Miriwung (Hamilton 1996) is an extension of this system: the apical contrast is realized in all post-vocalic positions, and neutralized post-consonantally, as in Muriṇbaṭa. In addition, the Miriwung t/ʈ contrast is neutralized word initially: only alveolars surface initially. The observations about contexts of non-assimilatory neutralization are summarized below (cf. Hamilton 1996, Steriade 1999 for more details):

(3)   Contexts of non-assimilatory place neutralization for major place contrasts

      Pre-C: [s]-[ʃ] contrast reduced to [ʃ]/ _[t, l, n] (Chumash)

      Domain-final: [n]-[m] reduced to [n]/ __# (A. Greek)

(4)   Contexts of non-assimilatory place neutralization for major place contrasts

      Post-C: [t]-[ʈ] contrast reduced to [t] in C_ (Muriṇbaṭa)

      Domain initial: [t]-[ʈ] contrast reduced to [t] in C_ ; #_(Miriwung)

We have observed that assimilatory and non-assimilatory neutralization draw the same distinctions between contexts but that these distinctions are contrast specific. Major C-place contrasts are shielded in pre-V position from both types of neutralization[1]. The apical contrast, on the other hand, is shielded, again from both forms of neutralization, in the post-V context, while post-C and initial apicals are potential targets to both processes. The task before us is to provide an explanation for this systematic pattern and a framework for its phonological analysis.

The gist of the argument developed here is that both varieties of neutralization select their targets on the basis of a hierarchy of perceived similarity between the input and the output strings. One key assumption is that the perception of phonological similarity is influenced by auditory factors such as the availability of cues to the relevant contrast: the terms of poorly cued

---

[1] Word final consonants undergo major C-place assimilation in phrasal contexts; here too assimilation is invariably regressive . The picture is more complex in the case of apical assimilation applying at word boundaries: see section 3.

contrasts being more similar than those of a better cued contrast. The comparison between major place and apical contrasts is revealing because we know independently that their perceptual correlates have a different contextual distribution: this perceptual difference matches observed differences in neutralization and assimilation patterns. A more general principle emerges from the comparison between apical and major place assimilation: *assimilation for any feature F targets positions in which the F contrast, if realized, would be less salient.*

Perceptual factors identify not only the direction of assimilation but also the likelihood that it will occur at all: we observe that different CC clusters give rise to considerably different rates of place assimilation, depending again on the salience of place contrasts in each one of the cluster's components. Assimilation is infrequent in cases where each C carries cues that allow a reliable identification of its place category; assimilation is prevalent if one C lacks its primary place correlates. Taken together, the observations about predictability of direction and incidence in assimilation suggest the hypothesis in (5):

(5) Perceptual similarity to input

The likelihood that a lexical representation R will be realized as modified R′ is a function of the perceived similarity between R and R′.

Thus, if [n] in a sequence /anpa/ is confusable with [m], then the lexical representation /anpa/ is similar to an assimilated variant [ampa]: then regressive assimilation is likely. If a progressively assimilated variant [anta] is perceived as more dissimilar to the lexical form /anpa/ then progressive assimilation is correspondingly less likely. Finally, if neither [n] is perceived as similar to [m] nor [p] to [t] in the original /anpa/, then neither form of place assimilation is likely to occur. The reference to *perceived* similarity in (5) is meant to convey the central idea here: that perceptual factors – among them cue distribution - play a critical role in defining degrees of similarity between lexical forms and their conceivable modifications.

Although the focus here is on establishing the link between cue distribution, assimilatory direction and rates of assimilation, the paper touches also on the form taken by speakers' knowledge of similarity and the evidence that this knowledge has consequences for grammatical organization.

2. Major place assimilation

The initial evidence for linking assimilatory direction to perceptibility in regressive place assimilation comes from studies by Fujimura, Macchi and Streeter (1978), Ohala (1990) and Jun (1995). The first two works have established selective attention on the part of speakers to release-related cues to place (CV transitions) to the detriment of transitional place cues clustered at the onset of closure (VC transitions). In stimuli containing an interlude with the duration of a single C and contradictory place cues (VC transitions excised from one C; CV transitions excised from another; no burst) listeners identify the C on the basis of CV transitions. Longer interludes allow the hearer to interpret the contradictory information as evidence of a heterorganic CC cluster. However, as the interlude duration decreases, the interpretation shifts to a single C and this sets the stage for a forced choice between conflicting place cues. In this forced choice, the CV transitions are given preference over VC transitions. It is just the CV transitions alone, without the burst, that can have this effect (Fujimura et al. 1978).

A further detail makes it likely that we are dealing here with a language-independent perceptual bias that can be safely invoked in explaining cross-linguistic patterns. Fujimura and colleagues compared the performance of Japanese and English subjects, based on the observation that Japanese phonotactics make the VC transitions redundant: in careful speech, all Japanese coda C's are homorganic with following onsets. The CV transitions, on the other hand, are indispensable for the identification of word initial place features. Therefore the Japanese phonotactics may train the speakers to ignore VC transitions: this may be the language-specific, phonological origin of the CV dominance effect. But in English there is a much smaller asymmetry between pre and postvocalic C's with respect to the range of place contrasts: place features contrast in stops in all positions. Therefore, as the English subjects displayed exactly the same bias in favor of the CV transitions, their behavior cannot be attributed to the effect of language specific phonotactics.

Fujimura's experiment thus settled, for this case, the issue of phonotactically dependent perception biases and allowed a direct comparison of the relative contribution of CV and VC transitions, independently of information present in bursts. Further, it showed that the CV-bias does not have an articulatory basis: when played backwards, the stimuli were processed in the same way, with the CV transitions (now originating as VC transitions) dominating the percept again. This means that the effect of CV transitions could not have been due to an asymmetry in coarticulation. This result invites one to speculate then that major place assimilation targets C_1 in VC_1C_2V simply because C_1's place cues are less well attended to and hence a place-modified C_1 is a lesser departure from the input than an altered C_2.

The notion that assimilation asymmetries have a perceptual basis is further supported by the observation that manner classes differ in their propensity to assimilate in ways that mirror confusion rates for place features. Kohler (1990) notes that nasals are more likely to assimilate than stops and stops in turn are more likely than fricatives, observations confirmed by Jun's (1995) survey. Thus final [bn] in German assimilates progressively to [bm] (*haben* [habm̩]) but final [pt] does not (*liebt* [liːpt], *[liːtt][2]). Medial [t] assimilates to a following obstruent (*mitbringen* [mipbriŋən]) but [s] does not (*Ausfahrt* *[auffaːrt]). The correspondence between place assimilability and rates of place confusion was later established by Hura, Lindblom, Diehl 1992, who presented listeners with word sequences of the form $XVC_1\#C_2VY$ where C_1 varied between a stop, a nasal, a fricative and C_2 was a heterorganic stop. The resulting misperception rates display the hierarchy nasals > stops > fricatives, with nasals being the most confusable class. Kohler's and Hura et al.'s studies suggest that both the incidence of assimilation and its direction are controlled by perceptibility differences. We return to this point below.

However, if we limit our attention to major place assimilation, three interpretations of the directionality data are possible, as outlined by Fujimura et al. (1978). The first possibility is that the CV transitions are dominant in the perception of major place contrasts, but not necessarily for other contrasts: this is the contrast-specific interpretation of perceptibility differences that I pursue here. A different view has become the standard syllable-based interpretation (cf. Jun 1995, Beckman 1998): assimilation is regressive because the target C_1 is a coda and the trigger C_2 is an onset. Perceptibility is controlled by syllable position because listeners pay more

---

[2] The progressive assimilation in the case of final [bn] does not invalidate observations made earlier, which pertain only to intervocalic clusters.

attention to onsets than to codas. Finally, the third possible interpretation is that the information encoded in C_2 is dominant simply because C_2 is more recent.

The contrast-specific account of assimilation predicts that assimilation will work regressively only for features cued primarily by CV transitions. Progressive assimilation is not ruled out; indeed, it is predicted for any feature cued mainly by VC transitions. Since C_2 in a VC_1C_2V sequence lacks the VC transitions and C_1 possesses them, assimilation for any features cued by VC transitions should target C_2. In contrast, the syllable-based and the recency accounts of regressive assimilation do not differentiate among feature types, or at least do not do so on the basis of cue distribution: there is no reason why it should be major place features, and not others, that spread from onsets to codas or from more to less recent C's . Both these accounts lead one to expect, wrongly, that all forms of local intervocalic assimilation will be regressive, regardless of the feature involved.

3. Apical assimilation

We turn now to the t/ʈ apical contrast. The reason to consider this case is that cues to the t/ʈ distinction lie primarily in the VC transitions, as noted by Ladefoged and Maddieson (1986:12), Dave (1977) for Gujarati, Stevens and Blumstein (1975) for Hindi, and Bhat (1973:235) in a cross-linguistic survey of retroflexion. The VC transitions preceding retroflexes point to distinctively low F3, F4 values relative to those of dentialveolars: typical F3 loci are at 1800 Hz for retroflexes, 2700 Hz for dentialveolars; F4 loci at 2750 for retroflexes and 3500 for dentialveolars (based on Gooniyandi data reported by McGregor 1990 and Gujarati data in Dave 1976). In contrast, the CV transitions of the two classes are similar or indistinguishable[3]. There is a clear articulatory explanation for this asymmetry: during the retroflex closure, the tongue tip slides forward: at release, it reaches a site nearly identical to that of an apico-alveolar (Butcher 1995, Henderson 1997). Therefore the release-related cues, including the CV transitions, are misleading if both apicals are released from the same constriction point.

The VC transitions are not the only acoustic properties distinguishing retroflexes from alveolars. Thus Anderson and Maddieson (1994) show that Tiwi [t] and [ʈ] are distinguished by closure duration (shorter for [ʈ]), VOT values (shorter for [ʈ]), and burst amplitude (lower for [ʈ]). Similarly, Dart (1991: 127) finds that the t/ʈ contrast of Malayalam involves small differences in VOT values (shorter for [ʈ]). Some of these properties may serve as cues for the t/ʈ distinction in initial and post-C position: for languages like Hindi, where the apical contrast is maintained initially and after C's, this is a necessary assumption.

This said, the data reviewed thus far suggests two differences between the perception of the apical contrast and that of major place contrasts. First, it suggests a reversal in the status of CV and VC transitions in the perception of apical classes relative to the role of transitions in the

---

[3] The following is Dave's (1977: 98) description of representative Gujarati data "A comparison between formant transitions in vowels adjacent to dental and retroflex consonants shows that the vowels [a], [ə], [u], [o] before retroflex stops have a very clear negative transition of F3 and F4 which is not found before dental stops […]. The vowels *following* dental and retroflex consonants do not show any consistent differences except that [o] has a lower F4 after retroflex C's and [a] has a significantly lower F3 and higher F2 after retroflex C's." But, although [ta] and [ʈa] CV transitions are distinguishable, those of [ka] and [ʈa] are not (p.118): only the VC transitions distinguish the retroflexes from all other consonant classes, across vocalic contexts.

perception of major place contrasts: VC transitions provide unambiguous information distinguishing among apical classes, in contrast to the CV transitions. By contrast, we have seen earlier that the CV transitions dominate in the perception of major place categories. Second, if transitions represent the main source of place information, then apical contrasts differ from major place contrasts in yet another way: apicals should be more confusable in the absence of VC transitions than major place classes should be in the absence of CV transitions. That is because the apicals' CV transitions are ambiguous, whereas the VC transitions of major class categories are not, in most vocalic contexts. This second point is borne out informally by scores of Australianists who report their inability to distinguish auditorily among initial apicals[4]. By contrast, few – if any - field workers report difficulties in distinguishing postvocalic unreleased [pˀ] from [tˀ] and from [kˀ]. A more rigorous confirmation of this point comes from the results of Anderson's (1997) perceptual confusion experiment with speakers of Western Arrente. Anderson compared rates of identification of the medial C in aCə words and in Cə word fragments extracted from aCə: the identification rates were similar in the two conditions for labials, velars and laminals, but listeners' performance dropped to chance levels in distinguishing t/ʈ in the truncated Cə fragments. More interestingly, the rates of apical confusion reported in Anderson's experiment far exceed the rates of confusion among major place classes ([p], [k], coronals) reported in a symmetrical experiment by Ohala and Ohala (1999), in which listeners had to identify Hindi consonants based on VC stimuli with bursts excised. Ohala and Ohala's aC stimuli – the closest symmetric counterparts to Anderson's Arrernte Cə stimuli - showed that labials, velars and coronals as a class continued to be reliably identified (although loss of release caused confusion between [tʃ] and [t])[5].

Thus if the relative perceptibility of place distinctions determines (a) the sites of non-assimilatory place neutralization; (b) the direction of assimilation; and (c) the incidence of place assimilation, then we predict (a) that apicals will have place distinctions neutralized typically in initial and post-C contexts, where they lack VC transitions; (b) that apical assimilation will be triggered by the better cued C_1 and undergone by the poorly cued C_2 in VC_1C_2V sequences; and (c) that progressive place assimilation will be considerably more common in apical clusters (where C_2 lacks reliable transitional cues) than any form of place assimilation should be in non-apical clusters (where C_1 continues to be identifiable, after at least some V's).

These points are verified below, beginning with the predictions concerning place neutralization in apicals. This process targets mainly contexts lacking VC transitions: initial and post-C positions. The main trends observed in the distribution of apical neutralization are summarized below.

(6)   Patterns of apical neutralization

        (a) **The Law**: if the t/ʈ contrast occurs in a language, it occurs after V.

        (b) **The General Case:**   t/ʈ contrast *only* after V

---

[4] Some of these reports are cited in Hamilton (1996) and Steriade (1995).

[5] In different vocalic contexts, the rates of misidentification went up in the Ohala and Ohala 1999 study for the coronals (after [i]) and velars (after [u]).

(c) **The Initial Deviation** : t/ṭ contrast only after V and in #__

(e.g. Djinang: Waters 1979)

(d) **The I-Deviation** : t/ṭ contrast after central and back V; reduced to [t] after [i]

(e.g. Maṭuṭuni«a: Dench 1995)

The generalization in (6.a) is self-explanatory: the context of optimal perceptibility for apical subtypes is after a vowel. This is also the first context where the contrast surfaces, if it occurs at all. Neutralization in all other contexts is widespread (6.b). The initial deviation (6.c) is a general effect, not specific to this contrast and will not concern us further here (cf. Steriade 2000). The i-deviation (6.d) is more revealing for a perceptually based analysis: in at least some languages, iC transitions do not distinguish clearly t/ṭ (Dave 1977:103) and it/iṭ confusions have been reported (Ohala and Ohala 1999). The cause of this is the conflict between the gesture of tongue body fronting and raising (for [i]) and the curling back of the tongue tip required for [ṭ]: if the conflict is resolved in [i]'s favor, the result is a diminished retroflexion gesture and thus a perceptually reduced t/ṭ contrast[6]. Thus not only the general law regarding the context of apical neutralization but also the details support a link between differential perceptibility and the selection of neutralization targets.

The patterns of apical assimilation support the same view. There are a number of clear generalizations that can be supported regarding the direction of assimilation in this case. Beginning with the best documented and most striking one, place assimilation is progressive in the vast majority of apical clusters and it is 100% progressive in apical clusters that belong to the same word and are of identical stricture level (both stops, or both fricatives, or both nasals, or both liquids).The Appendix summarizes relevant cases. Major points are outlined below:

(7)   Patterns of inter-apical assimilation

(a) **The Law**: all else equal, assimilation is progressive in apical clusters

(b) **Final Deviation**: assimilation may be regressive across the boundary of content words: e.g. Sanskrit (Allen 1962), Punjabi (Malik 1995).

(c) **Nasal Deviation**: assimilation may be regressive in nasal-stop clusters e.g. Sanskrit (Allen 1962); Malayalam (Asher and Kumari 1997)

The data in the Appendix shows that progressive assimilation is encountered with clusters of identical or different stricture, in retroflex+dentialveolar as well as dentialveolar+retroflex sequences. The latter are harder to document, because retroflexes are rare in suffixes, but enough relevant cases occur to insure that what we analyze as progressive assimilation (i.e.[ṭt]→[ṭṭ],

---

[6] Inspection of the palatograms in Dave (1977:38-39) for [aṭa]-[uṭu]-[iṭi] reveals that retroflex contact is initiated at a point considerably further front after [i] than after [a] and [u]. Thus the [aṭa]-[ata] and [uṭu]-[utu] pairs are much better differentiated than [iṭi]-[iti] in terms of constriction site.

[tt̪]→[tt̪]) cannot be reanalyzed as a retroflex-dominance effect ([t̪t]→[t̪t̪], [tt̪]→[t̪t̪]). The two deviations noted in (7) are systematic. The final deviation (7.b) indicates the attested possibility that a word initial C will trigger regressive assimilation. This effect may be invariant and it is likely to reflect contrast-independent factors, as it has counterparts in hiatus resolution (Casali 1996): word initial segments—whether poorly cued or not—are more likely to be invariant. The nasal deviation (7.c) records the occasional occurrence of regressive inter-apical assimilation in nasal+stop clusters. This can be attributed to the fact that F3 is attenuated by nasal zeroes: since F3 is a diagnostic value for the alveolar/retroflex distinction this means that in an apical+apical cluster where C_1 is a nasal and C_2 is a stop, there may be no constant perceptual advantage of C_1 over C_2, as there is in other heterorganic apical clusters. The variable direction of assimilation in N-stop clusters may be tied to this fact.

The overall number of apical assimilations attested is small but the trend is very clear. If there is assimilation in word-internal apical clusters with identical manner the direction is exclusively progressive. The direction is predominantly progressive also in clusters of apicals with different manner: word internally, regressive assimilation is documented only among nasal-stop clusters.

Returning to points made at the outset, I note that the direction of place assimilation is contrast-specific. Apical assimilation targets C_2 while major-place assimilation targets C_1. In both cases the consonant undergoing F-assimilation—for any F—possesses fewer or weaker F cues. This generalization poses an analytical challenge: if the synchronic analysis of individual systems is to reflect cross-linguistic markedness properties, then the analysis of assimilation must succeed in identifying its targets on the basis of *relatively reduced perceptibility*. The question is how, and whether, to build this notion into a synchronic analysis.

5. Likelihood of assimilation

However, before addressing the formal issue, I verify a different prediction of the central hypothesis in (5): assimilation should affect more frequently clusters in which place features are not reliably identified in one component of the sequence, less frequently clusters where both components are reliably cued. This conjecture extends Kohler's (1990) and Hura et al's (1992) ideas to systems in which both apical and major place assimilation can in principle occur.

A source of relevant data is Hamilton's (1996) survey of morpheme internal phonotactics in 115 Australian languages. This represents the largest known group of languages contrasting t/ʈ and thus allows a comparison between rates of apical vs. major place assimilation. Hamilton's survey contains 76 relevant languages: these permit CC clusters and contrast apical types in some contexts. In this group, we observe first that apical clusters are virtually always homorganic, in contrast to clusters consisting of apicals and non-apicals. Second, assimilation between apical in C_1 and non-apical in C_2 is uncommon and implied by all other assimilation types: if any unassimilated CC is allowed, that sequence consists of apical+non-apical. A further heterorganic cluster that occurs frequently in Hamilton's corpus is that consisting of a palatal in C_1 and a non-coronal in C_2. The least common unassimilated sequence are the apical clusters. The count in (8) substantiates these points. The languages are grouped by types of heterorganic CC clusters they permit morpheme internally, with the * sign marking impermissible heterorganic clusters in a given language group. All but one language permit homorganic CC sequences: languages listed as disallowing all heterorganic clusters, permit homorganic NC, LC.

(8) Incidence of heterorganic cluster types in Hamilton's (1996) Australian corpus

Number of languages	Apical+non-apical (e.g. ŋp, np, nk,ŋk)	Palatal+C (e.g. ≠p, ¥k)	Non-apical+C (e.g. mk, ŋp, mc, ṉp)	Apical+apical (e.g. ɳt, nt̠, n, ɳl)
2 (3%)	*	*	*	*
12 (15%)	✓	*	*	*
40 (53%)	✓	✓	*	*
21 (28%)	✓	✓	✓	*
1 (1%)	✓	✓	✓	✓

The first striking fact in this data is the near-absence of heterorganic apical clusters. We ask first whether this fact has an articulatory basis: perhaps the only clusters avoided in this corpus are those in which two successive C's engage the same active articulator. But this cannot be the reason: in apical+laminal clusters (e.g. [nt̠], [lt̠] [ɳc]), the tongue blade must form in quick succession two distinct constrictions. Yet these heterorganic clusters occur in 49 (64%) of the 76 languages, frequently yielding contrasts of the form [nt]-[nc], [ɳt]-[ɳc], [nt]-[nt̠], [ɳt]-[ɳt̠]. Conversely, a retroflex+alveolar (e.g. [ɳt]) should pose no articulatory difficulty, as its trajectory is similar to that involved in a single retroflex: recall that the tongue tip reaches a site close to the alveolar region anyway by the end of the retroflex closure. But in fact the only retroflex+alveolar cluster in the corpus, [ɳt], is reported in only one language, Nyigina, and its status remains debatable according to Hamilton.

The factor determining which clusters assimilate is not ease of articulation but rather the perceived similarity between the assimilated output and the input. The implicational relation between assimilation of the four types of clusters identified in (8) follows from the contextual distribution of cues to place. We have seen that for all intents and purposes there are no cues to the apical contrast in the CV transitions. Then if a language contrasts apico-alveolars and retroflexes, this very fact trains the speakers to attend primarily to the VC transitions of all apicals, as it is these transitions that reliably distinguish apical subtypes. In this way, the VC transitions become a major cue to place—and the unique transitional cue—for all apicals, whether retroflex or not. Consider now an apical+non-apical cluster in intervocalic position, e.g. [t̠k]. The apical [t̠] is identified by its major transitional cue (the VC transitions); and so is the non-apical [k], for which the CV transitions serve this role. The place features of the two C's are then equally salient in this cluster. We predict that place assimilation is least likely to affect this sequence, because modifying either C_1 or C_2 will result in an equally noticeable departure from the perceived properties of the input. For Australian languages, this is indeed the case[7].

---

[7] I lack information about the distribution of cues to lamino-palatals in the languages discussed and thus cannot comment on the frequent occurrence of palatals in the C_1 position of unassimilated clusters and on the virtual exclusion of the other laminal class, the dentals, from this context. It is possible that the preceding V is heavily influenced the palatals as well.

Compare now heterorganic apical+apical clusters (e.g. [nt̪], [ɳʈ]) with heterorganic clusters consisting of two non-apicals (e.g. [mk], [ŋp]). There is a clear difference in Hamilton's corpus between these two types: the apical clusters are virtually always assimilated, while non-apical [ŋp], [mk] clusters surface unassimilated in 22 (29%) of the 76 languages. If we include in the count the palatal+non-apical sequences then 62 languages (81%) possess heterorganic non-apical+non-apical clusters. Why this difference in the rate of assimilation between apicals and non-apicals? The apical clusters contain one member—C_2—that lacks all transitional cues to place distinguishing it from another apical[8]. In contrast, place is identifiable in each consonant of the non-apical clusters on the basis of some contextual cue, CV or VC transitions. If cue distribution contributes to perceived similarity, then the similarity between assimilated and unassimilated apical clusters ([t̪t̪] and [tt], [ʈʈ] and [t̪t̪]) is greater than that between assimilated and unassimilated non-apicals ([kp] and [pp], [pk] and [kk]): hence the much greater likelihood that an assimilatory sound change will be initiated for [t̪t̪], [ʈʈ] clusters. The chances of detecting and repressing incipient assimilation in the two types of clusters discussed - [t̪t̪], [ʈʈ] vs. [kp], [pk] - are thus very different: and this is reflected in the rate of success of such innovations and, ultimately, in their effects on phonotactic typology.

6. Manner effects in apical assimilation

Manner differences have an effect on the direction and incidence of place assimilation. We have seen that place contrasts are more confusable among stops than among sibilants (Hura et al. 1992) and that this correlates with different rates of regressive assimilation. Progressive apical assimilation displays the same asymmetry, as observed in Sanskrit (Allen 1962, Steriade 1995): clusters with identical manner features (two nasals, two stops, or two fricatives) word-internal assimilate in obligatorily and strictly progressive manner. I attribute this to the fact that C_1 possesses the major transitional cue to the apical contrast and C_2 lacks it (cf. 9.a). In this case, C_2 has no compensating perceptual advantage. In the case where C_1 is a fricative and C_2 is a stop or nasal, C_1 possesses both transitional and internal cues (the latter representing the fricative's noise spectrum) while C_2 possesses neither. Here too, assimilation is progressive, as predicted (cf. 9.b). Nasals assimilate regressively to stops and progressively to any preceding apicals (9.b-c): the possibility of regressive assimilation in nasal-stop clusters was discussed above and may be tied to the reduced perceptibility of place distinctions in nasals. (I do not however understand the difference between the Sanskrit and the Australian pattern where all apical clusters, including N+stop, assimilate progressively.) Finally, we consider (9.d), the case where C_1 is a retroflex stop [ʈ] and C_2 is an alveolar sibilant [s]. In this case, [ʈ] could have been identified through VC transitions and [s] through its noise spectrum.[9] In this case no assimilation takes place and Vʈ̩sV surfaces intact. One can speculate that this case is comparable to the heterorganic apical+non-apical clusters of Hamilton's corpus: it is likely that the fricative's noise spectrum alone provided place information that was comparable to that encoded in the stop's VC transitions. If so, then the members of the [ʈʂ] cluster were equally well cued, although in different ways, and any assimilatory realization — [ʈʂ] or [ʂʈ] — would have been avoided as too dissimilar to the lexical form.

(9) Manner effects in Sanskrit apical assimilation (Allen 1962, Whitney 1889)

---

[8] Place neutralized apicals can occur in the C_2 position of heterorganic clusters: they do so in 7 (9%) of the languages in the corpus.

[9] Cf. LaRiviere, Winitz and Herriman 1975 on noise spectrum as the major place cue in fricatives.

(a) same manner apical clusters: **progressive assimilation**

  i. fricative-fricative: ʂs → ʂʂ   (*jyotiʂ-su → jyotiʂʂu*)
  ii. stop-stop: ḍ-dʰ → ḍḍʰ   (*piṇḍ-dʰi → piṇḍḍʰi*)
                ḍ-t → ṭṭ   (*iḍ-te → iṭṭe*)
  iii. nasal-nasal: ŋ-n → ŋŋ   (*ßaŋ-na:m → ßaŋŋa:m*)

(b) fricative-C clusters: **progressive assimilation**

  i. fricative-stop: ʂt → ʂṭ   (*iʂ-ta → iʂṭa*)
  ii. fricative-nasal: ʂn → ʂṇ   (*uʂ-na → uʂ-ṇa*)

(c) nasal-obstruent clusters: **regressive assimilation**

  i. ŋ-t [nt]   expected *ka«-aŋṭi*, actual *ka«-anti*
  ii. ŋ-dʒ [≠dʒ] expected *dʒaŋdʒaŋa,* actual *dʒa≠dʒaŋa*

(d) stop-fricative clusters: **no assimilation**

  ṭs [ṭs]   *viṭ-su*

I have been unable to locate other languages with a sound pattern comparable to that of Sanskrit—esp. /ṭs/ and /ʂt/ lexical sequences - and thus cannot determine how representative this pattern of apical assimilation is.

7. The grammar of perceptibility effects

In the remainder of this study, I consider two questions relating to the analysis of the assimilation patterns discussed so far. First, is it necessary that the perceptual account of assimilation have some synchronic counterpart? Second, what would a perceptually-based, synchronic analysis of assimilation look like?

It is easier to address the second question first. To allow perceptibility effects into the analysis of place assimilation, we must adopt two assumptions, one justifiable and the other unoriginal. The first assumption is that speakers can compute relatively consistent similarity values for sound differences and that this computation of similarity takes into account cue distribution. This assumption is defended in section 9. The unoriginal second assumption is that constraint rankings in an Optimality Theoretic model of phonology can be indexed to phonetic scales (Prince and Smolensky 1993): the rankings we will discuss are those of correspondence constraints and the scales these rankings are indexed to are scales of perceived similarity. The idea is that if two contrasts a-b and x-y are known to differ in degree of similarity, with a-b more similar than x-y, then correspondence constraints that prohibit an a-b difference between input (a pronunciation norm) and output (a modified realization of it) are lower ranked than correspondence constraints that ban an x-y difference between input and output. What is needed

is to link similarity scales to the ranking of correspondence constraints. This point is sketched in section 10 and developed in more detail in Steriade 2000.

The issue of necessity—must similarity scales play this role in the grammar of assimilation internalized by native speakers?—is addressed in the next section by considering the incipient stages in an assimilatory sound change.

8. The innocent misapprehension theory of assimilation

The argument advanced here is that the speakers who initiate assimilation as a sound change select a specific modification of a lexical norm on the basis of two factors: perceived similarity to the original form and optimized articulation. Any modification must be tolerably similar to the original; and must involve an improvement, in articulation, perception or paradigm structure, over the original. This echoes Lindblom's et al. (1995), Kohler's (1990) and Hura et al.'s (1992) view of assimilation as "perceptually *tolerated* articulatory simplification". The simplification, in the case of place assimilation, consists of eliminating one of the two original constrictions. The critical assumption then is that speakers exert some control over the incipient sound change: that they do so, in part, by computing the distance in perceptual space between a lexical norm and potential modifications of it. In what follows, references to *the synchronic analysis of assimilation* are references to this computation.

An alternative view of assimilatory sound change has been put forth by Ohala in passages like this:

> "A non-teleological view of sound change [...]: neither the speaker nor the hearer chooses—consciously or not—to change pronunciation. [...] Rather variation occurs due to innocent misapprehensions about the interpretation of the speech signal [...] [Sound change] does not optimize speech in any way: it does not make it easier to pronounce, easier to detect, or easier to learn." [Ohala 1990:266]

The perceptually based asymmetries in place assimilation reported in this study would be consistent with Ohala's view, if the relative frequencies of sound changes that turns a's into b's matched the likelihood of a perceptual confusion between a and b. If that were the case, we could conclude that listeners mistake an intended a for a b and, if the b for a mistake is frequent, the b-forms become lexically entrenched and the a $\rightarrow$ b change become part of the grammar.

However, the patterns of perceptual confusion observed in the laboratory do not exactly match attested sound changes. More specifically, the experimental evidence on perceptual confusion correctly identifies strings more likely to be modified by sound change but does not match the *actual* modification. Recall for instance the place-confusability hierarchy nasals > stops > fricatives established by Hura et al. (1992). This matches observed tendencies in the selection of assimilation targets. But the patterns of confusion observed by Hura et al. were mostly non-assimilatory, with a bias in favor of the alveolar stops and nasals regardless of context. Thus the greater inclination of nasals to assimilate relative to stops cannot be attributed to higher rates of grammaticalized misperception for nasals vs. stops. Nasals do tend to be misperceived but not primarily in assimilatory ways. Therefore bare misperception is unlikely to be the root of assimilation. Asymmetries in assimilation may arise from the fact that speakers are aware of the relative rates of confusion by manner class and by position and deliberately put this

knowledge to use in their own production: they target the nasals more and the stridents less in place assimilation because they know that it is safer—perceptually more tolerable, as Hura et al. put it—to modify the nasals' place features than to modify the stridents. This may be an example of knowledge of perceptibility used as a phonological tool.

Continuing the defense of the optimizing intent in sound change, I consider now a different process, which plausibly involves an improvement not in articulation but in perception. Hume (1997) has discussed metathesis as a case of perceptual optimization. I would like to spell out this argument for a goal-directed sound change using the example of stop-sibilant metathesis (as in Southern English *ask* → *aks*). Central to the discussion is the different perceptibility of place in stops and sibilants: place identification in stops is dependent on transitions and burst quality while sibilants benefit, in addition to the vocalic transitions, from the place cues inherent in their noise spectrum. In general this means that the constriction information survives, for sibilants, in contexts where it could not be maintained for unreleased stops: initially before a stop, finally after a stop, or interconsonantally. There are four different cases of metathesis seen below (data from Brugmann 1933, Grammont 1933, Stroop 1981, Hume 1997 and Harkema 1999):

(10) Metathesis in stop+s, s+stop clusters

(a)   VTsV → VsTV   (T = stop):

        rural Latin *ipse* → *ispe; wepsa* → *wespa;*

        A. Greek: *euk^{h}samenos* → *eusk^{h}amenos*

        Old Dutch: *wepse* → *wes*p*e*

        19th cent. Parisian French: *fiksə* → *fiskə, ɛ̃dɛksə* → *ɛ̃dɛskə*

(b)   #TsV → #sTV

        A. Greek: p^h*syk^{h}e:* > *s*p^h*yk^{h}e:*

        Latin: *psallere* → *spallere*

        Dutch (children): *psychologe* → *spychologe*

(c)   VsTC → VTsC

        Lithuanian: *dresk-ti:* → *dreks-ti*

(d)   sT# → Ts#:

        Dutch (dialects and children): *wes*p → *weps*

        Southern American English *wasp*→ *waps*

Hume points out that, in the VTsV → VsTV case (10.a), the stop moves to a position where it will have CV transitions, the preferred source of place information for major place distinctions. More systematic is the case of initial metathesis, (10.b): #Ts → #sT. Here the stop is moving from a position where it possesses no transitional cues at all and where, if it were to remain, its common diachronic fate would be deletion. In cases of #Ts to #sT metathesis, a minor perceptual loss for [s] is offset by a major gain for the stop. Metathesis operates in the opposite

direction in cases (10.c) and (10.d) where the stop is trapped between a sibilant and another obstruent or the end of the word. Here too the movement positions the stop so that it will be flanked at least on one side by a vowel: this too is a major gain for the stop and a minor loss for [s]. To summarize, the common types of s//T metathesis seem to have the consequence of providing the stop with the best transitional cues locally available.

In principle, however, metathesis as a sound change can arise as a form of listener error. Given this, we ask a second question: Is s//T metathesis equally well attested in both directions? Note that confusability is, in principle, symmetric: if *aks* is confusable with *ask* then *ask* is also confusable with *aks* and any bias in favor of one of these forms requires explanation. If the sound change is initiated as misperception there would be no reason to expect metathesis in one direction and not in the other. In fact however the direction of metathesis is highly constrained. Only certain types of reversal, which can be identified as perception-optimizing, are frequent and systematic, as shown both by Grammont's comments and by individual studies like Stroop 1981 and Harkema 1999.

(11) *Ts-sT* Metathesis in the Dutch children's speech (Harkema 1999)

- Word-final   sT→ Ts common: *asterisk ~ asteriks, wesp ~ weps* 'wasp'

  Ts → sT not found: *tʃips* 'chips', never **tʃisp*,

  gips 'plaster', not **gisp*

- Word-initial Ts → sT attested: *psychologe → spychologe*

  sT→ Ts not found: *spriŋhaan* 'grasshopper' not **psriŋhaan*

The same can be said about most other cases of systematic metathesis: when the reordering of C's becomes a regular sound change rather than lexical fluctuation, it systematically locates the stop in a position of improved perceptibility in its local context. This generalization fits the data in Grammont's 1933 and Hume's 1997 surveys[10].

The metathesis data is consistent in two respects with the view of sound change presented here: on the one hand, it does suggest an attempt at perceptual improvement. Metatheses that involve no gain or a net loss of perceptibility are not systematic. On the other hand, drastic dislocations—such as *psykʰe → *pykʰes* or *æsk → *kæs—* are strictly unattested even when they do improve the perceptibility of all consonants involved.

---

[10]Grammont is an explicit defender of the teleological aspect of CC metathesis (although he assumes that metathesis optimizes syllable structure rather than perceptibility, an assumption that is difficult to defend in detail). Here is a typical quote stressing teleology: "L'interversion est toujours déterminée opar un principe d'ordre et de moindre effort. Elle a souvent pour object de réparer les désastres causés par les évolutions brutales [...] Elle ne crée jamais des monstres mais elle les redresse quand il s'en présente" (1933:249) [BH's high school French; sorry for errors: "Metathesis is always governed by a principle of order and of least effort. It often has the object of repairing disasters created by violent [phonetic] evolution. It never creates monsters but it repairs them when they arise."]. Grammont identifies a few unusual cases of metathesis (such as Sorabian initial ʃk becoming kʃ: kʃit 'shield' from earlier ʃkit, a loanword) that are said to be motivated by the need to avoid clusters unusual in the language. The Sorabian case - about which we lack further details - belongs in this class.

These remarks do not exclude the possibility of sound change originating as genuine, non-optimizing, misperception but make it likely that some modifications of lexical norms are *selected* by speakers because they pass the two tests mentioned earlier: perceived similarity to the original form and improved functionality.

9. The P-map

The idea that potential sound innovations are subjected to a similarity test is inspired by Lindblom's proposal (1990, Hura et al. 1992, Lindblom et al. 1995) that speakers constantly tune their articulation to the perceptual needs of their listeners. As Lindblom puts it (1990:403), the speaker is guided by his "tacit awareness of the listener's sources of information independent of the signal", and compensates for the absence of this signal-independent information or for anticipated gaps in the signal itself. The idea central to Lindblom's H&H hypothesis is that speaker behavior is guided by a *model of the listener's lexical access mechanisms*, which draw on both signal-dependent and signal-complementary information. But in addition to this model of lexical access, it is likely that the speaker's activity is also guided by *a model of the generic listener's perceptual abilities and biases*.

I call this second model the P-map. The P-map is the repository of speakers' knowledge, rooted in observation and inference, that certain contrasts are more discriminable than others, and that the same contrast is more salient in some positions than in others. One function of the P-map is to identify the margins of articulatory freedom for the speaker: the regions of relative safety within which he may deviate from established pronunciation norms while minimizing the risk of being "found out". Thus, if an innovative speaker contemplates articulatory simplification in a VC_1C_2V cluster, it is the P-map that will identify the optimal target of simplification: the consonant whose modification is least likely to be detected by his interlocutors. Other P-map functions include identifying more vs. less salient morphological alternations (Steriade 1999b), and generating the judgments of similarity needed for rhyming, loan adaptation, speech disguise, and in experimental situations.

One can conceive of the P-map as a set of statements, each of which assigns a similarity value to a contrast realized in a specific context. By *contrast* I mean a perceived difference between two strings, regardless of its phonemic status. I leave open the source of the similarity knowledge contained in the P-map: for our immediate purposes it matters only that this knowledge exists in the minds of speakers. Whether it results from observations of confusion rates (Shepard 1972), or is deduced from a similarity computation, the end result of interest here is just the actual set of similarity statements. We will however have to assume that among the factors that shape similarity judgments, relative perceptibility plays a role: we see below that contrasts realized in less informative contexts, where they lack some of their perceptual correlates, are judged more similar than the same contrasts realized in more informative contexts. This may indicate that a deductive theory of similarity – one that attributes to speakers the ability to anticipate similarity relations among pairs of strings, without necessarily relying on observed confusion rates – will be based on a calculus of perceptual correlates to contrasts, rather than on a calculus of distinctive features. The following P-map fragment (12) illustrates this hypothesis. The vertical axis lists segmental contrasts generated by anteriority differences among apicals (t/ʈ, s/ß, n/ɳ etc.). The horizontal axis lists six of the contexts where each of these sounds might in principle occur: the contexts are arrayed from left to right in the order of the number and likely weight of potential cues to retroflexion available in each position. The cells

thus defined are labeled with characters whose sizes code hypothesized similarity values: larger characters stand for contrasts assumed to be less similar, hence more salient.

(12) Hypothetical P-map fragment: similarity of apical pairs by context

• letter size reflects hypothesized similarity: bigger letter = less similar pair

	V_V	V_#	V_C	#_V	C_V	C_C
s/ʂ	S/ʂ	S/ʂ	S/ʂ	s/ʂ	s/ʂ	s/ʂ
t/ʈ	t/ʈ	t/ʈ	t/ʈ	t/ʈ	t/ʈ	t/ʈ
n/ɳ	n/ɳ	n/ɳ	n/ɳ	n/ɳ	n/ɳ	n/ɳ

A number of basic observations, all independent of the issue of assimilation, lead one to postulate the P-map. I outline these below.

9.1. Poetic equivalence and the P-map

Speakers can provide relatively consistent judgments of sound similarity, in experimental situations or in versification. Some similarity judgments are reflected by linguistic mechanisms already in place but others cannot be so understood. Thus *cumulative* similarity effects – e.g. the fact that [ɪn]-[ɛn] is a more similar pair than [ɪn]-[ɛd] – are explained by any framework that uses the features [±high] and [±nasal]. On the other hand, if the pair [ɪm]-[ɪn] emerges as systematically more similar than [ɪb]-[ɪd], then some supplement to distinctive feature theory is needed to record the judgment, as the same feature of labiality separates [b]-[d] and [m]-[n].

The perceptual similarity literature and the literature on poetic uses of similarity contain enough examples of the second sort to suggest that knowledge of similarity does not emerge straightforwardly from the feature count. English rhyming practices (Zwicky 1976) disregard place of articulation differences in final nasals much more frequently than in final stops, revealing that [ɪm]-[ɪn] is indeed judged more similar than [ɪb]-[ɪd]: this fact mirrors differences discussed earlier in the perceptibility of place distinctions between nasals and stops. Further, pairs of front lax vowels that differ in height ([ɪ]-[ɛ]) rhyme much more frequently than corresponding tense pairs ([i]-[e]) (Zwicky 1976) and this effect corresponds to differences in rates of perceptual confusion (Peterson and Barney 1951): note that the same height difference contributes more to dissimilarity in a longer vowel than in a shorter one. A third example is the fact that the contrast between a C and its absence evokes much stronger dissimilarity judgments in positions adjacent to a vowel: pairs like [drɪft]-[drɪf] are judged more similar than [drɪft]-[drɪt] (Wingstedt and Schulman 1988; Fleishhacker 2000). Correspondingly, poets frequently count as equivalent, in rhyme and assonance, VC_0C-$VC_0\emptyset$ pairs like *man-hand,* while VCC_0-$V\emptyset C_0$ pairs like *loud-ground* seldom function in this way (Zwicky 1976). Something like the P-map must be

assumed as a supplement to distinctive feature theory, as the latter fails to draw any of the distinctions observed here.[11]

9.2. Loan adaptation and the P-map

The ability of speakers to compute the closest equivalent, in their native inventory, to a non-native string is a further reason to postulate the P-map. A single example of this sort is mentioned here, drawing on research by Silverman (1992) into loan adaptation patterns from English into Cantonese. The case of interest to us is a similarity ranking between two contrasts: stop vs. zero and [s] vs. zero. This ranking is revealed by the different treatment of sibilants and stops in complex syllable margins. Sibilants surface in all contexts, including in pre-obstruent onsets and post-obstruent codas: stamp is borrowed as [sit^{h}am], *tips* as [t^{h}ipsi], *forecast* as [fok^{h}asi]. Stops, in contrast, surface only when adjacent—in the English word—to a vowel or liquid: *post* becomes [p^{h}osi], not *[p^{h}osit(i)], *lift* [lip], not *[lipt^{h}i] or [lip^{h}it(i)]. An optimality-theoretic analysis of this data will treat sibilant recovery (formalized as context-free MAX(strident)) as undominated; but corresponding faithfulness conditions for stops must rank lower and must depend on the input context. In particular, MAX (stop)/C_# ranks below DEP (cf. *lift* → [lip]) and below Contiguity (cf. *post* → [posi]). But both DEP and Contiguity are outranked by MAX (strident) (*stamp* → [sit^{h}am]). The different ranking of various MAX(C) constraints suggests that contrasts between different segment classes and Ø have different similarity values. I summarize this point below using the notation in (12).

(13) P-map reflecting the treatment of two C/Ø contrasts in Cantonese loans:

	_ (L)V	V(L)_	#_T	VN_	T_#
s/ Ø	s/Ø	s/Ø	s/Ø	s/Ø	s/Ø
T/ Ø	T/Ø	T/Ø	_	T/Ø	T/Ø

The similarity rankings observed here do not mirror Cantonese alternations or distributional asymmetries and thus could not have been projected from the speakers' knowledge of their native sound system. Nor do they reflect processes at work in English: English labials do not delete in complex codas (cf. *stamp*→[sit^{h}am]). Rather, the rankings are likely to reflect evaluations by Cantonese speakers of the distinct degrees of auditory salience of the C/Ø contrasts listed. And, once again, these ranked similarity relations do not derive from a distinctive feature count: [s] does not possess more features than [p], on any version of feature theory, but [s] is preserved in contexts where [p] is not.

10. P-map-based analyses of place assimilation

---

[11]Only a subset of the similarity facts mentioned here can be accomodated by Broe's (1993) model, which builds into the computation of similarity the effect of phonological redundancy. See also Frisch, Broe and Pierrehumbert 1997 for an application of this theory and an extension that encodes the effect of temporal distance. This model is an important advance in the study of similarity, but it does not attempt to encode the effects of context on perceived similarity and the link between similarity and perceptibility.

In the early part of this study, I had documented the fact that the direction and incidence of place assimilation are influenced by perceptibility factors. These factors are contrast specific - they operate differently for apical and major place assimilation – and thus spelling out their effect on place assimilation requires enumerating the perceptual cues to different place categories as affected by context, internal and external to the segment. In a later section I noted that an analysis of place assimilation in which perceptibility plays an explicit role is necessary. We must explain the basis on which innovative speakers select modifications of lexical norms: a factor in this selection must be the similarity between the assimilated form and the unassimilated original. Finally I have shown in the last section that speakers are indeed able to perform similarity computations in which perceptibility factors play a role. The result of such computations is the P-map. The final question is how the contents of the P-map can be linked to the grammar, so as to control the functioning of assimilation.

The answer is the idea that rankings among correspondence constraints (McCarthy and Prince 1995) must be indexed to the perceived similarity of the input-output differences they refer to. A correspondence constraint prohibits a certain type of difference between a lexical form and its surface realization: but, as we have seen, some differences are more salient than others. The proposal is that if two contrasts a-b and x-y differ in perceived salience in a given language, the correspondence constraints "a must not surface as b" and "x must not surface as y" are predictably ranked relative to each other: the more salient contrast projects the higher ranked constraint. Thus if s-Ø is a more noticeable contrast than t-Ø then both MAX(s) ranks above MAX(t) and DEP(s) ranks above DEP(t). The Cantonese data discussed earlier requires a ranking of this type: linking up the correspondence system to the P-map explains where the speakers' knowledge about the ranking comes from, a fact that up to now has remained mysterious.

Similarly, assume that the t-ʈ/V_C (the apical contrast as realized in V_C) is more salient than t-ʈ/C_V (the apical contrast realized in C_V), an assumption built into the P-map fragment in (12). We are justified to suppose this, given the difference between the information relevant to apical identification that is carried by VC vs. CV transitions. Then the constraint IDENT(anterior)/ V[_, apical, stop]C – which requires identity of anteriority values between apical stops in V_C– must rank above IDENT(anterior)/C[_, apical, stop]V – which refers to apicals in C_V. Under this fixed ranking, the direction of apical assimilation is invariably progressive. The illustrations below simplify matters by using a blanket constraint (*Agree*, cf. Lombardi 1999) against heterorganic C's.

(14) Apical assimilation:

a. /t-ʈ/ → [tt]

/pat-ʈal/	Agree	Ident [ant]/ V[_apical, stop]C	Ident [ant]/ C[_apical, stop]V
☞ pattal			*
patʈal		*!	
patʈal	*!		

b. /ʈ-t/ → [ʈʈ]

/- eʈ-tu /	Agree	IDENT [ant]/ V[_apical, stop]C	IDENT [ant]/ C[_apical, stop]V

☞ - eṭṭu			*
- ettu		*!	
- eṭṭu	*!		

Note that we continue to predict the possibility of cross-linguistic variation in the application of apical assimilation: that is a function of the ranking of *Agree* relative to the lower IDENT[anterior] constraint. However, if apical assimilation is to occur at all, then the effect of cue distribution on perceived similarity guarantees under our proposal that the assimilation will be progressive.

The same mechanism that projects rankings of correspondence constraints from differences in contrast salience insures the progressive direction of apical assimilation and, at the same time, the regressive direction of major place assimilation. Here I assume that the terms of major place contrasts are perceived as more similar in post-V position than pre-V, an assumption partially justified by the results on cue weighting in place perception due to Fujimura et al. (1978). It is these differences in distinctiveness that will insure that IDENT[place]/ _V systematically outranks Ident[place]/V_ (where *place* refers to major place features): and this in turn will guarantee the regressive direction of assimilation.[12]

 (15)   Major place assimilation:

a. /tp/   → [pp]

/atpa/	AGREE	IDENT [place]/ C_V	IDENT [place]/ V_C
☞ appa			*
atta		*!	
atpa	*!		

b. /pt/   → [tt]

/apta /	AGREE	IDENT [place]/C _V	IDENT [place]/ V_C
☞ -atta			*
appa		*!	
apta	*!		

Note that the IDENT constraints we must use are highly specific: they do not simply refer to identity of, say, anteriority values, but rather to anteriority values as realized in specific segment types (stops, nasals, fricatives) and specific contexts. There is no question, given the data reviewed earlier, that only these highly specific constraints can characterize complex assimilation patterns like those of Sanskrit, in which the apical's stricture degree and the external context interact. The question is rather where the knowledge of this vast set of correspondence constraints is coming from. The P-map provides a plausible answer here too if we assume that any two P-map cells with distinct similarity indices project distinct, and ranked, correspondence constraints. Thus the speaker derives his knowledge that there exist at least the two Ident[anterior] constraints used in (14) from his knowledge, encoded in the P-map, that the pair

---

[12] No evidence has been presented so far that listeners have any awareness of their own CV bias in place perception. The argument here rests heavily on the fact that listeners show awareness of *some* of their own perceptual biases: the conjecture is that the CV bias is among these.

ʈ-t sounds significantly different after a vowel, but quite similar in other contexts. A complementary assumption is that P-map cells with identical similarity indices are unranked relative to each other and identically ranked relative to other correspondence constraints. Under this assumption, the learner who knows that the similarity of the t/ʈ contrast is identical to that of d/ɖ in all contexts also expects, without further investigation, that patterns of apical assimilation should not be affected by voicing differences.

Earlier I had asserted that the P-map contains sets of statements that assign specific similarity values to contrasts. One reason to make this assumption is that absolute values of similarity can easily translate into relative similarity rankings for any pair of P-map cells: this is very useful in understanding how the choice is made between multiple repair strategies for a given phonotactic violation (Steriade 2000). But for the cases discussed here, the idea that the P-map contains not just similarity rankings among constraints but rather absolute similarity values is potentially useful in explaining the role of perceptibility in determining the rate of assimilation. So far we have employed the constraint AGREE to force place-assimilation in any CC sequence: and we have observed that AGREE should be able to outrank or be outranked by the conflicting Ident constraints, in order to characterize assimilating as well as non-assimilating languages. Then we still need to explain the asymmetries in assimilation rates: for instance the fact that apical clusters almost always assimilate, especially when compared to the very low rate of assimilation in apical+non-apical clusters. Such facts suggest a partition of P-map cells into three classes: those with very high similarity indices, whose terms are perceived as nearly identical; those with very low similarity indices, whose terms are perceived as highly distinct; and all others. The assumption we must make is that the correspondence constraints matching two of these three classes have nearly invariant rankings relative to phonotactic constraints like *Agree*. First, we must assume that all phonotactics typically outrank correspondence constraints projected from the high similarity class. This will insure that the goal of phonotactic improvement will typically be pursued if the similarity cost - in terms of deviation from the input - is low. An example of this low ranking constraint corresponding to a high-similarity P-map cell is IDENT [anterior]/[_apical, stop]V, which is systematically violated in virtually all Australian languages possessing retroflexes. Second, we must assume that all phonotactics are typically outranked by correspondence constraints projected from the low similarity class of contrasts. This will insure that the phonotactic improvements, no matter how dramatic, simply cannot be sought at the expense of a highly noticeable deviation from the original. Possible examples of the low similarity-high ranking class of constraints are variants of the Linearity constraint that prohibit certain long distance reordering of segments: we observed earlier that local *ps-sp* reversals are well documented, while distal reversals of the form *asp-pas* seem impossible, despite the fact that they turn highly marked into nearly optimal syllables. Our suggestion was that this and other sound changes are missing because there is *an absolute dissimilarity limit to phonotactic improvement*: to encode this we will need absolute similarity values.   Finally, the third class of correspondence constraints – all others – are the ones whose ranking relative to phonotactics is genuinely unpredictable. This proposal must remain sketchy in the absence of a model of similarity computation. If implemented, it may allow us to place realistic limits on the still excessive amount of cross-linguistic variation predicted by allowing free ranking of correspondence and phonotactics.

I close by identifying one more of the many question that remain open in this investigation of the interplay between perceived similarity and phonological patterns. Recent work by Harnsberger 1999 and Hume, Johnson, Seo and Tserdanelis 1999 demonstrates that the perception of similarity can be influenced by language specific factors. It follows then that

certain P-map properties should be expected to differ from language to language. Such cases have not been discussed here for two reasons. First, no testable phonological consequences could be computed from the language-specific similarity effects documented so far. Second, the results of Fujimura et al. 1978 reviewed earlier, suggest that language specific factors (e.g. the rich range of place contrasts available before C's in English) do not necessarily obliterate language independent perceptual biases (in this case the bias in favor of CV transitions). Cases that may eventually shed light on the interplay between language specific and language independent similarity factors are those in which the two types of factors enter in conflict. These must be left for future work.

References

Allen, W.S 1962 *Sandhi*, Mouton.

Anderson, V. 1997 "The perception of coronals in Western Arrernte*" Proceedings of Eurospeech*, Rhodes, Greece 1.389-392

Anderson, V. and I. Maddieson 1994 "Acoustics of Tiwi consonants*," UCLA Working Papers in Linguistics*, vol.87.

Asher, R.E. and T.C. Kumari 1997 *Malayalam*. Routledge, London.

Beckman, J. 1998 *Positional Faithfulness*, University of Massachusetts, Amherst PhD thesis.

Bhat, S. 1967 *Descriptive analysis of Tulu*. Deccan College, Poona.

Bhat, D.N.S. 1973 "Retroflexion: an areal feature" Working Papers on Language Universals, 13, 27-58

Blake, B.J. 1979 *A Kalkatungu grammar*. Pacific Linguistics, series B, no. 57.

Bloch, J. 1970. *The formation of the Marathi language.* Delhi, Motilal Banarsidass

Broe, M. 1993   *Structured specification*, University of Edinburgh Ph.D.thesis

Brugmann 1933 *Kurze Vergleichende Grammatik der Indorgermanischen Sprachen* Truebner: Strassburg (1st printed 1903)

Butcher, A. 1995 "The phonetics of neutralization: the case of Australian coronals," in Windsor-Lewis, J. (ed.) *Studies in general and English phonetics*, Routledge, London.

Casali, Roderic 1996 "Vowel elision in hiatus: which vowel goes?" *Language* 73, 493-533.

Dart, S. 1991 *Articulatory and Acoustic Properties of Apical and Laminal Articulations*, UCLA Working Papers in Phonetics, vol. 79

Dave, Radhekand 1976 "Retroflex and Dental Consonants in Gujarati: a Palatographic and Acoustic Study*," Annual Report of the Institute of Phonetics, University of Copenhagen* (ARIPUC) 11, pp. 27-155.

Dench, A. 1995 *A grammar of Martuthunira* Pacific Linguistics Series C vol. 125.

Fleishhacker, Heidi 2000 "The Location of Epenthetic Vowels with Respect to   Consonant Clusters: An Auditory Similarity Account" UCLA MA thesis.

Flemming, Edward 1995 *Perceptual Features in Phonology*, UCLA Ph.D. dissertation

Frisch Stefan, Michael Broe and Janet Pierrehumbert 1997 "Similarity and Phonotactics in Arabic" *Rutgers Optimality Archive 223*

Fujimura, O. 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.

Glasgow, K. 1981 "Burarra phonemes," in B.Waters (ed*.) Australian Phonologies*, collected papers, Darwin: SIL, pp.63-90

Grammont, M. 1933 *Traité de phonétique*   Paris: Librairie Delagrave.

Harkema, H. 1999 "Dutch schwa is not a long vowel." UCLA ms.

Hamilton, P. 1996 *Constraints and markedness in the phonotactics of Australian Aboriginal languages*, Ph.D. Dissertation, University of Toronto

Harnsberger, J. 1999 "The effect of linguistric experience on perceptual similarity among nasal consonants," *Research Report on Spoken* Language Processing, Progress Report 23.Indiana University.

Henderson, J. 1997 *Topics in Eastern and Central Arrernte Grammar*, Ph.D thesis, The University of Western Australia.

Hume, E. 1997 "The role of perceptibility in Consonant/Consonant Metathesis, in Blake, S., Eun-Sook Kim and Kimary Shahin (eds.) *WCCFL XVII Proceedings.* Stanford: CSLI. 293-307

Hume, E., K. Johnson, M. Seo and G. Tserdanelis 1999 "A cross-linguistic study of stop place perception," *Proceedings of the 14th International Congress of Phonetic Sciences.*

Hura, S., B.Lindblom and R.Diehl 1992 "On the role of perception in shaping phonological assimilation rules*," Language and Speech* 35, 59-72

Jun. Jongho 1995 *A Constraint Based Analysis of Place Assimilation Typology*, UCLA Ph.D.Dissertation.

Keen, S. 1983 "Yukulta," in R.M.W.Dixon and B.*Blake Handbook of Australian Languages* vol. 3.

Kohler, K. 1990 "Segmental reduction in connected speech in German: phonological facts and phonetic explanation:" in Hardcastle, W.J. and A. Marchal (eds*.) Speech Production and Speech Modeling*, Kluwer, Dordrecht, 69-92.

Kostic',D. A.Mitter & B.Krishnamurti 1977 *A short outline of Telugu phonetics*, Calcutta

Kothandaraman, P. 1997 *A Grammar of Literary Tamil,* International Institute of Tamil Studies, Chennai, India.

Ladefoged, P. and I. Maddieson 1986 *Some of the Sounds of the World's Languages,* UCLA Working Papers in Phonetics, vol. 64.

Lal, S.M. 1991 *A descriptive analysis of Urali : speech of a Dravidian hill tribe* Mysore : Central Institute of Indian Languages.

LaRiviere, C. H. Winitz and E. Herriman 1975 "The distribution of perceptual cues in English prevocalic fricatives," *Journal of Speech and Hearing Research*18, 613-622

Lindblom, B. 1990 "Explaining phonetic variation: a sketch of the H&H theory", in Hardcastle, W.J. and A. Marchal (eds*.) Speech Production and Speech Modeling*, Kluwer, Dordrecht, 403-439.

Lindblom, B. S. Guion, S. Hura, S.-J.Moon, R. Willerman 1995 "Is sound change adaptive?" *Rivista di Linguistica* 7, 5-37.

Lombardi, Linda 1999 "Restrictions on the direction of voicing assimilation, direction in assimilation*" Rutgers Optimality Archive* 246

Malik, A. N. 1995 *The phonology and morphology of Panjabi* New Delhi: Munshiram Manoharlal Publishers.

McCarthy, John J. and Alan Prince 1995 "Faithfulness and Reduplicative Identity," *University of Massachusetts Occasional Papers in Phonology*, vol. 5.

McGregor W. 1990 *A Functional Grammar of Gooniyandi*. John Benjamins'. Amsterdam; Philadelphia

Mojumder, A. 1972 *Bengali language; historical grammar*. [1st ed.] Calcutta, Firma K. L. Mukhopadhyay.

Nayar, V.R.P. 1973 *Malayalam: a linguistic description,* National Research Publishing Company, Trivandrum.

Nordlinger, R. 1998 *A Grammar of Wambaya* Pacific Linguistics Series C, 140.

Ohala, J. 1981 "The listener as a source of sound change," in C.Masek, R.A.Hendrick and M.F.Miller (eds.) *Papers from the Parasession on Language and Behavior*, Chicago Linguistic Society, University of Chicago. 251-274

Ohala, J. 1990 "The phonetics and phonology of aspects of assimilation," in John Kingston and Mary Beckman (eds.) *Papers in Laboratory Phonology I: Between the grammar and the physics of speech.* Cambridge University Press, Cambridge. 258-275

Ohala, M. and J. Ohala 1999 "Correlations between Consonantal VC transitions and degree of perceptual confusion of place contrast in Hindi," Proceedings of the 5th International Conference on Spoken Language Processing, Syndey. 2795-2798

Peterson, G. and H. Barney 1951 "Control methods used in the study of vowels," *JASA* 44.2.401-407

Poser, W. 1982 "Phonological representations and action at a distance," in van der Hulst, H. and N.Smith (eds.) *The structure of phonological representations*: part 2. Foris, Dordrecht, 121-158.

Prince, Alan and Paul Smolensky 1993 *Optimality Theory*, RuCCS TR-2, Rutgers University, to appear MIT Press.

Rajapurohit, B.B. 1982 *Acoustic Characteristics of Kannada*, Central Institute of Indian Languages, Mysore.

Reddy, J. 1979 *Kuvi Grammar*, Central Institute of Indian Languages, Mysore.

Shepard, Roger 1972 "Psychological representation of spech sounds," in E.E.David and P.B.Denes (eds.) *Human Communication: a Unified View*. New York, McGraw-Hill.

Silverman, D. 1992 "Multiple scansions in loanword phonology: evidence from Cantonese" *Phonology* 9, 289-328.

Sridhar, S.N. 1990 *Kannada*, Routledge. London.

Steriade 1995 *Positional neutralization*, UCLA ms.

Steriade, Donca 1999a "Alternatives to the syllabic analysis of consonantal phonotactics," in Osamu Fujimura, Brian Joseph and Bohumil Palek (eds.) *Proceedings of the 2nd Linguistics and Phonetics Conference*, The Karolinum Press, Prague, 205-245.

Steriade, Donca 1999b "Lexical Conservatism in French Adjectival Liaison," in B.Bullock, M. Authier and L. Reed (eds.) *Formal Perspectives in Romance Linguistics*, John Benjamins, pp. 243-270.

Steriade, Donca 2000 "P-map effects on constraint organization" UCLA ms.

Stevens Kenneth and Sheila Blumstein 1975 "Quantal aspects of consonant production and perception," *Journal of Phonetics* 3, 215-233.

Street, C.S. and G.P. Mollinjin 1981 "The Phonology of Murinbata" in B.Waters (ed.) Australian Phonologies, collected papers, Darwin: SIL, pp.183-244.

Stroop, J. 1981 "Metathesis van s en p" *Spektator*, 11, 224-248.

Wackernagel, Jacob, 1896. *Altindische Grammatik*. Göttingen, Vandenhoeck und Rupricht.

Waters, B. 1979 *A distinctive features approach to Djinang phonology and verb morphology*. Summer Institute of Linguistics, Australian Aboriginal Branch. Darwin.

Whitney, William Dwight 1889*Sanskrit Grammar*, Harvard University Press.

Wingstedt, Maria and Richard Schulman 1988 "Listener's judgments of simplifications of consonant clusters," *Linguistics* 26, 105-123

Zwicky, A. 1976 "Well, this rock and roll has got to stop, Junior's head is hard as a rock," in S.Mufwene, C.Walker and S.Steever *Proceedings of CLS 12,* University of Chicago, 676-697.

Appendix: patterns of word-internal apical assimilation

Language	Source	Retro- Same manner	Alveolar Different manner	Alveolar- Retro Diff. manner
Kuvi	Reddy 1979	/uɳ-t-e/e)/ [uʈʈe/e)] (drink-past-1sg)	/hoɳ-du/ [hoɳɖu] (run-pl)	
Kannada	Sridhar 1990 Rajapurohit 1982	/hee -ₐl-illa/ [hee illa] (tell-inf-neg)	•/toʈʈilu/ [toʈ u] 'cradle' •/kaaɳ-d-e/ [kaɳɖe] (see-past-1sg)	/maɖ-id-ₐ u/ (do-past- 3sgfem) [maɖidlu]
Urali	Lal 1991	•/eɳ-nuuru/[eɳɳuuru] (hundred-8= '800') •/keɖ-t-a-/ [keʈʈa-] (spoil-intrans')	• /iru -ti/ [iru ʈi] (Irula-woman)	
Telugu	Kostic', Mitter & Krishnamurti 1977	• /guɖilu/ [gu u] (temples) • /waaɖini/ [waaɳɳi] (him)		
Tulu	Bhat 1967		•/uɳ-dE/ [uɳɖE] (I eat) •/uɳ-la/ [uɳ a] (eat!) •/paaɖɪ-la/ [paaɖ a] (put!) •/kaʈɪ-la/ [kaʈ a] (tie!)	
Tamil	Kothandaraman 1997	/viʈ-t-aan/ [viʈʈaan] (read-past-3sg)		
Malaya-lam	Asher & Kumari 1997 Nayar 1973:41	•/kee -tu/ [keeʈʈu] (hear-past) •/peʈ-tu/ [peʈʈu]	/uɳ-tu/ [uɳʈu] (eat-rice-past)	

Indic			
Language	Source	Retro- Same manner	Alveolar Diff. manner	Alveolar-Retro Same manner
Middle Indic	Mojumder 1972	Skt. va«dʰaɳa MI baɖɖaɳa 'growth'		Skt. suut«am MI suttam 'sutra'
Marathi	Bloch 1973:173	" -l is almost "		
Sanskrit	Whitney 1889, Wackernagel 1958, Allen 1962	• /av-iɖ-dʰi/ 'favor!' [aviɖɖ̥ʰi] •/ßaɳ-na:m/ 'of six' [ßaɳɳa:m] • /jyotiß-su/ 'in planets' [jyotißßu]	• /iß-ta!/ 'sacrificed' [ißʈa!] • /uß-na!/ 'hot' [ußɳa!] •/gi:«-su!/ 'in songs' [gi:«ßu!]	

Australian				
Language	Source	Retro-Alveolar Diff. manner	Alveolar- Same manner	Retro Diff. manner
E.Arrernte	Henderson 1997: 63	/ar-ə -t-an-əm/ 'see-pl' [arə ʈanəmə]		
Murinbata	Street and Mollinjin 1981	• /ŋud̪u- E -nu/ [ŋud̪u E ŋu] 'roll-FUT'		/pan-ʈal/ [pantal] 'he/she cut it' /ma-n- arᶉ-nu/ [manlarᶉnu] 'wrap-them -FUT'
Yukulta	Keen 1983	/mipu -ti/ [mipu ʈi] (threat-verb)		
Wambaya	Nordlinger 1998	• /RED-labaɳga/ [laba(ɳ)- abaɳga] (branch of tree)		
Kalkatungu	Blake 1979	• /ulaaɳ-ta/ [ulaaɳʈa] (sun-locative)		
Burarra	Glasgow 1981:85			/an-ɖeta/ [andeta] 'strong one'