12/17/08

The Realized Input

Edward Flemming
Department of Linguistics & Philosophy, MIT

1. Introduction

The P-map hypothesis holds that the acceptability of an unfaithful mapping between phonological forms depends on the perceptual magnitude of the differences between those forms (Steriade 2001a,b). There is now substantial evidence in support of this hypothesis (e.g. Steriade 2001a,b, Fleischhacker 2005, Kawahara 2007, Flemming 2008, among others). For example, Steriade (2001a) shows that the P-map hypothesis accounts for the generalization that the only means by which languages eliminate final voiced stops is through devoicing, e.g. [tab] → [tap], they never employ nasalization of final stops, [tab] → [tam], or epenthesis of a vowel following the stop, [tab] → [tabə]. She provides evidence that the devoicing of a final stop is a smaller perceptual change than the alternative repairs, so, by the P-map hypothesis, it is the means of eliminating marked final voiced obstruents that incurs the least violation of faithfulness constraints, and consequently is always the preferred option.

However, there are still open issues concerning the proper formal implementation of the P-map hypothesis. In this paper we focus on a particular issue that must be addressed in formalizing the theory: perceptual similarity depends on the phonetic details of the forms involved. For example, the extent of the perceptual cues differentiating [tab] and [tap] depends on whether the final stops are audibly released or not (Revoile et al 1982). Languages differ in this respect, for example pre-pausal stops in French are usually released (Tranel 1987), whereas in Korean they are consistently unreleased (Martin 1951). We will see that correspondence constraints are sensitive to these kinds of phonetic differences between languages, for example place assimilation generally does not target released stops (Jun 2002), and that providing a formulation of these constraints that accounts for this sensitivity is not a simple matter because these phonetic details are not systematically distributed in the input. It will be argued here that the effects of phonetic detail on perceptual correspondence can only be accounted for if inputs are mapped onto a phonetically-detailed level of representation, the Realized Input, where language-specific patterns of phonetic realization such as audible release of stops are represented. Perceptually-based correspondence constraints refer to this level of representation rather than referring directly to the raw input, so perceptual correspondence is sensitive to language-specific phonetic detail (cf. Jun 2002, Flemming 2006, Gallagher 2007).

This proposal is motivated by evidence that the magnitude of the faithfulness violations incurred by a change in place of articulation or stop voicing depends on language-specific details of the realization of these contrasts. We will then show that the Realized Input also provides the basis for analyses of a class of ‘displaced contrast’ phenomena where an underlying contrast in one feature is apparently realized in terms of a different feature. For example, in Friulian an underlying obstruent voicing contrast is realized as a contrast in the length of the preceding vowel in a context where obstruents are devoiced. We will see that the apparent displacement of a contrast is more accurately characterized as a shift in the relative contributions of cues that generally differentiate the underlying contrast. Obstruent voicing contrasts are typically accompanied by differences in the duration of the preceding vowel, and in Friulian this difference in vowel duration is exaggerated to maintain contrast where cues from closure voicing are lost. The various cues to a contrast like obstruent voicing are represented in the Realized Input, and it is faithfulness to these cues in the face of contextual markedness constraints that shapes the realization of displaced contrasts.

The next section reviews the evidence that correspondence constraints are sensitive to language-specific phonetic detail, and show that this phenomenon is problematic for most current approaches to the formulation of positional faithfulness constraints. Section 3 develops the notion of the realized input, and its position in phonological grammar. The resulting model is applied to the analysis of the problematic case studies. Positing the realized input as an intermediate level of representation implies an ordering between the phonetic realization processes that map the input onto the realized input and the processes that map the realized input onto the output which predicts the possibility of opaque interactions between the two. Section 4 shows that opaque interactions of this kinds are required to account for displaced contrast phenomena. Section 5 concludes.

2. Faithfulness constraints are sensitive to language-specific phonetic detail

2.1 Stop releases and place assimilation

The first case study demonstrating the sensitivity of faithfulness constraints to language-specific phonetic detail comes evidence that language-specific patterns of stop release affect the typology of place assimilation (Jun 2002).

Jun (1995, 2004) provides evidence that the typology of place assimilation in intervocalic consonant clusters is shaped by the distinctiveness of place contrasts, depending on manner and position. In an intervocalic C_1C_2 cluster, a consonant is more likely to undergo assimilation in place of articulation if the resulting perceptual change is smaller. The most general pattern accounted for in these terms is that C_1 may assimilate in major place of articulation to C_2, but C_2 rarely assimilates to C_1, and then only under restricted morphological conditions (Jun 1995). This reflects the fact that major place contrasts (labial vs. coronal vs. dorsal) are more distinct preceding vowels than preceding consonants (Redford & Diehl 1999, Wright 2001). Given the P-map hypothesis, a change in C_2 place is a greater violation of faithfulness than a change in C_1 place, so a markedness constraint against heterorganic clusters will be satisfied by C_1 assimilation rather than C_2 assimilation, e.g. /atka/ → [akka], *[atta].

This analysis can be implemented in terms of a fixed ranking of correspondence constraints, where constraints against larger perceptual mismatches between corresponding forms are higher ranked (Steriade 2001b:240f.)[1]. IDENT(place)/C_V, which requires faithful realization of major place features in prevocalic contexts, is universally ranked above IDENT(place)V_C, so a change in the place of a pre-consonantal consonant is a lesser violation of faithfulness than a change in the place of a pre-vocalic consonant.

[1] Jun (2004) proposes a slightly different analysis according to which 'constraints preserving perceptually more salient segments must be ranked above those preserving perceptually less salient segments' (p. 73), but the resulting ranking of faithfulness constraints is the same.

Jun accounts for other implicational universals concerning place assimilation in terms of additional universal rankings of faithfulness constraints. For example, if stops undergo place assimilation then nasals undergo place assimilation in the same context (Jun 1995, Mohanan 1993). Jun argues that this is a consequence of the lesser perceptual distinctiveness of post-vocalic nasal place contrasts, and that this is reflected in a universal ranking of faithfulness constraints (1). Ranking AGREE(place) between these two constraints yields place assimilation that targets nasals but not stops.

(1) $\text{IDENT(place)/V}\underset{\text{[stop]}}{\underline{\qquad}}\text{C} \gg \text{IDENT(place)/V}\underset{\text{[nasal]}}{\underline{\qquad}}\text{C}$

This analysis of the role of perceptual factors in constraining place assimilation runs into problems with language-specific differences in phonetic realization, because the existence of these differences imply that it is not possible to determine the cues that are available in a given context on a language-independent basis. As a result, it is not possible to establish a universal ranking of positional faithfulness constraints of the kind illustrated in (1), based on the strength of the cues available in the context specified by each constraint.

For example, Bloomfield (1933:119) observes that languages differ in whether they produce consonant clusters with open or close transition: in open transition the first consonant in the cluster is released before the constriction for the second is formed, whereas in close transition the constriction for the second consonant is formed before the first is released (cf. Catford 1977:220ff.). Where stops are produced in close transition with a following obstruent, this overlap obscures the release burst of the stop, whereas release bursts are audible in open transition. Languages such as Montana Salish (Flemming et al 2008) and Georgian (Robins & Waterson 1952, Chitoran 1998) employ open transitions in stop-obstruent clusters, so stops are released in all contexts, whereas others, such as English (Henderson & Repp 1982) and Korean (Martin 1951), employ close transitions, resulting in no audible bursts for stops preceding obstruents. Stop bursts provide cues to place (Dorman et al 1977, Kochetov & So 2007), so this language-specific difference in the distribution of stop bursts results in language-specific differences in the quality of cues to stop place in pre-obstruent contexts. Accordingly, it is not possible to determine the perceptual magnitude of a change that violates a constraint like IDENT(place)/V_C on a language-independent basis – it depends on whether stops are released in this context.

If language-specific details of realization were irrelevant to phonological patterning, we might be able to resolve this problem devising a measure of perceptual similarity that abstracts away from details such as the presence vs. absence of stop bursts, but the evidence suggests that these details do affect phonological patterning. Jun (2002) provides evidence that major place assimilation generally does not target released stops. He surveys languages where stops are released preceding obstruents (e.g. Arabic, Hindi, Russian), and finds that none of them show place assimilation in this context. Conversely, in languages where stops undergo place assimilation (e.g. Catalan, Korean, English), stops are unreleased preceding obstruents. Note that lack of audible release in the latter set of languages cannot be explained as a consequence of assimilation rather than a cause of assimilation because stops are unreleased in all stop-obstruent clusters, but assimilation only applies to a subset of these clusters. For example, in Korean velar stops do not undergo place assimilation, but they are unreleased in pre-consonantal contexts (Jun 1995).

The generalization that place assimilation applies to unreleased stops but not to released stops implies that a change in the place of articulation of a released stop is a greater violation of faithfulness than a change in the place of articulation of an unreleased stop. In terms of the analysis outlined in (1) above, this suggests that

IDENT(place)/V $\underset{\text{[stop]}}{\underline{\quad\quad}}$C should be replaced by faithfulness constraints that refer to the presence or absence of stop release cues. But, as discussed by Jun (2002), implementing this idea is not simple because the release properties of a stop cannot be determined from the underlying representation. That is, given the standard assumption that there are no language-specific constraints on inputs (Prince & Smolensky's (2004:225) thesis of 'Richness of the Base'), any stop can be specified as released in any context in the input, and any restrictions on the distribution of stop releases must be derived by constraints on output forms. So the language-specific distribution of stop releases, which is relevant to the occurrence of place assimilation, can only be observed in outputs, not in underlying representations. We will see below (section 2.3) that the alternative of trying to formulate faithfulness constraints to refer to the presence of properties such as stop bursts in the output is also problematic, but first we will see that the phonological relevance of stop bursts is not an isolated case. The second case study demonstrating the sensitivity of faithfulness constraints to language-specific phonetic detail concerns variation in the realization of voiced geminates, and its implication for faithfulness to voicing.

2.2 Partial devoicing and faithfulness to voicing

Kawahara (2006) shows that faithfulness to the voicing specification of a singleton consonant outranks faithfulness to voicing of a geminate in Japanese, a phenomenon that he formalizes in terms of the ranking in (2).

(2) IDENT(voice)$_{\text{Singleton}}$ >> IDENT(voice)$_{\text{Geminate}}$

Kawahara further shows that this ranking follows from the P-map hypothesis, but only given some language-specific facts about the realization of voiced geminates in Japanese: Voiced geminate obstruents are realized with partial devoicing and are consequently less perceptually distinct from their voiceless counterparts than voiced singletons, which are fully voiced. Voiced geminate obstruents are fully voiced in other languages such as Arabic, so the ranking in (2) can only be determined based on knowledge of the language-specific patterns of phonetic realization observed in Japanese. As is discussed further below, positing such a dependency between language-specific phonetics and constraint ranking is highly problematic.

The evidence for the ranking in (2) comes from the adaptation of loan words into Japanese. Loan words borrowed from English often contain geminate consonants due to a pattern in which final consonants following a lax vowel are frequently adapted as geminates, e.g. [sutoppu] 'stop'. Voiced stops in the context are regularly adapted as voiced geminates, e.g. [webbu] 'web', [kiddo] 'kid'. However, voiced geminates optionally devoice when another voiced obstruent appears in the word, e.g. [guddo]~[gutto] ‘good’, [doraggu]-[dorakku] ‘drug’. This follows a general pattern according to which native Japanese words cannot contain two voiced obstruents (Itô & Mester 1986), attributed to a constraint OCP(+voice). However, singleton voiced obstruents in loan words do not devoice even in the presence of a second voiced obstruent: [bagii], *[bakii] ‘buggy’, [gibu], *[gipu] ‘give’.

Kawahara (2006) analyzes these patterns in terms of the constraint ranking in (3). Faithfulness to voicing in singletons outranks OCP(+voice), so two voiced singletons can appear in loanwords, but the ranking between OCP(+voice) and IDENT(+voice)$_{\text{Geminate}}$ is variable, so devoicing of geminates is possible to avoid a violation of OCP(+voice). As Kawahara shows, the selective devoicing of geminates cannot be attributed simply to the greater markedness of voiced geminate obstruents compared to voiced singletons. IDENT(+voice)$_{\text{Geminate}}$ clearly outranks any markedness constraint against voiced geminates since these sounds surface faithfully in words like [webbu]. Devoicing only applies in the presence of a second voiced obstruent, and therefore must be motivated by OCP(+voice).

(3) IDENT(+voice)$_{\text{Singleton}}$ >> OCP(+voice), IDENT(+voice)$_{\text{Geminate}}$

Kawahara argues that this difference in ranking of faithfulness to voicing in singletons and geminates is motivated by the P-map hypothesis, since voicing contrasts between geminates are more confusable than voicing contrasts among singletons, as shown by a perceptual categorization experiment. The greater confusability of voicing among geminates is attributed to the fact that voiced geminates in Japanese are realized with substantial devoicing – about 60% of the closure interval is actually voiceless – while voiced singletons are voiced throughout their closure interval (Kawahara 2006:559ff.). However, this pattern of realization of voiced geminates is language-specific. As Kawahara observes, voicing is maintained throughout the closure of voiced geminate stops in other languages including Egyptian Arabic and Madurese. So according to this analysis, the ranking of correspondence constraints in Japanese is determined by a language-specific fact about the phonetic realization of geminates (Kawahara 2006:566).

There is no evidence that voiced geminates are more subject to conditioned devoicing processes than voiced singletons in languages such as Arabic where geminates are regularly fully voiced. For example, a number of Arabic dialects, including Makkan Arabic, have a process of pre-pausal devoicing where final stops lose closure voicing (Watson 2002:248, Kenstowicz, Abu-Mansour & Törkenczy 2003), but word-final geminates are not more prone to devoicing in this context than singletons. On the contrary, in Makkan Arabic the closure of a singleton voiced stops is completely devoiced before pause while voiced geminates retain significant closure voicing in the same context (Abu-Mansour p.c.). Ham (200:1131) reports a similar pattern for a speaker of Levantine Arabic). This pattern does not necessarily imply that faithfulness to voicing in geminates is higher ranked in Arabic, reversing the ranking in (2). It may be that the constraint motivating final devoicing only requires a certain duration of voicelessness preceding pause – the final voiced geminates appear to be partially devoiced since the duration of voicing is significantly less than the duration of an intervocalic geminate. However, the pattern suggests that faithfulness to voicing of geminates is not lower-ranked than faithfulness to voicing of singletons where geminates are regularly fully voiced.

Kawahara (2006) attempts to modify Steriade's (2001a) implementation of the P-map hypothesis to accommodate the effects of language-specific phonetic realization by proposing that correspondence constraints are ranked in each language according to the patterns of phonetic realization that obtain in that language. So IDENT(voice)$_{\text{Singleton}}$ >> IDENT(voice)$_{\text{Geminate}}$ in Japanese because voiced and voiceless geminates are more similar than voiced and voiceless singletons in that language, but this ranking would not obtain in a language like Mekkan Arabic where voiced geminates are fully voiced. However Kawahara does not develop an explicit account of the mechanism by which phonetic realization determines the ranking of phonological constraints. This kind of interaction is difficult to accommodate in models in which phonetic realization is a module that simply applies to the output of the phonological component. In such a model, it is not clear how details of phonetic realization of voiced geminates can be made available to regulate the ranking of phonological constraints that determine whether the language even permits voiced geminates.

As observed by Jun (2002), the sensitivity of faithfulness constraints to language-specific phonetic detail is straightforwardly accounted for if the relevant faithfulness constraints do not refer directly to the input, but instead to a level of representation where the patterns of phonetic realization specified by the grammar have been imposed on the input. Following Gallagher (2007), we will refer to this level of representation as the Realized Input (RI). In RIs pre-obstruent stops are specified as unreleased in Korean, and voiced geminates are specified as partially voiced in Japanese, so the faithfulness violation incurred by place assimilation or devoicing can be determined by assessing the perceptual distance between the RI and these candidate outputs. Before developing this line of analysis further, we consider one alternative that tries to avoid positing an intermediate level of representation by inferring the patterns of phonetic realization that are relevant to faithfulness from the forms of candidate outputs. However, constraints of this form are unable to account for the patterns under discussion here.

2.3 Output-oriented correspondence constraints.

The alternatives to positing an intermediate level of representation are to try to identify the distribution of properties such as stop bursts from the input or the output. We have already observed that it is not feasible to refer to the presence of these predictable properties in the input since they could be specified in any context, given the assumption of Richness of the Base. Correspondence constraints that refer to the presence of predictable properties in the output have been proposed in the context of work on positional faithfulness constraints. For example, Beckman (1998) formulates faithfulness constraints such as IDENT-ONSET which refer to the syllabic position of segments. Reference to syllabification in faithfulness constraints raises many of the same problems as reference to language-specific phonetic detail: richness of the base implies that segments can be assigned any syllabic position in the input, so faithfulness constraints cannot usefully refer to input syllabification, and syllabification is language-specific, so syllable-position cannot be inferred from the input on a language-independent basis. The solution that Beckman proposes is that syllabically-conditioned faithfulness constraints refer to syllabification in the candidate outputs. So IDENT-ONSET requires segments that appear in onset position in the output to match the feature specifications of their input correspondents, regardless of the syllabic status of those input segments. Wilson (2001) shows that this type of constraint fails to derive the desired effects when applied to processes like deletion that can themselves have substantial effects on syllabification in the output. The same problems arise if we try to adopt this output-oriented approach to formulating faithfulness constraints that are sensitive to phonetic cues.

A key problem case is deletion in medial /-VC_1C_2V-/ obstruent clusters, which universally targets the pre-obstruent consonant, C_1 (Wilson 2001). This pattern cannot be accounted for in terms of privileged faithfulness to consonants that are syllabified in onset on the surface because whichever consonant is undeleted will be syllabified as an onset: $[V<C_1>.C_2V]$/ $[V.C_1<C_2>V]$. A similar issue arises if we analyze this pattern in terms of the P-map hypothesis. Intuitively, only C_1 deletes because $[VC_1C_2V]$ is more similar to $[VC_2V]$ than to $[VC_1V]$ since C_2 is prevocalic in the input. The properties of prevocalic consonants are marked by more, and more salient, cues which are lost if it is deleted, so deletion of a prevocalic consonant is a greater violation of faithfulness than deletion of a pre-consonantal consonant (Steriade 2001). As Wilson points out, this privileged faithfulness to prevocalic consonants cannot be established at the output because whichever consonant is undeleted is prevocalic on the surface.

A comparable problem arises if we attempt to account for the resistance of released stops to place assimilation by formulating a faithfulness constraint which refers directly to the presence of a release in the output, e.g. IDENT(place)/released, which requires consonants that are released in the output to have the same place of articulation as their input correspondent (cf. Padgett 1995:19). The problem arises in languages like Russian where stops are released in pre-consonantal contexts when they appear in heterorganic clusters, but are not generally released in homorganic stop-stop clusters (Zsiga 2000). Accordingly, if assimilation applies in a stop-stop cluster, the first stop is expected to be unreleased, and therefore would not be subject to an IDENT(place)/released constraint. So a constraint of this type would to be unable to block assimilation in a language like Russian because, given an input like /-kt-/, candidate outputs [-kt-] and [-t̚t-] would both satisfy IDENT(place)/released: [-kt-] satisfies the constraint because [k] is released and faithful, and [-t̚t-] satisfies the constraint because the unfaithful, assimilated consonant is unreleased. The problem in both cases is that the presence of the crucial cues in the output depends on whether or not the relevant process (assimilation or deletion) has applied, so just as the prevocalic status of a consonant cannot be assessed in the output after deletion, so the greater violation of faithfulness that is involved in changing the place of a released stop compared to an unreleased stop cannot be evaluated by reference to release bursts in the output.

Output-oriented faithfulness constraints are also problematic because they create a motivation to place marked material outside of the protection of high-ranking faithfulness constraints where it can be eliminated. This can generate unattested processes. For example, IDENT-ONSET(voice) in conjunction with a markedness constraint against voiced obstruents, *[+voice, -son], can favor parsing voiced obstruents in coda so that they can be devoiced without the protection of IDENT-ONSET(voice) (Beckman 1998:36 fn.). The ranking in (44) derives an unattested pattern in which intervocalic voiced obstruents are parsed as syllable codas and devoiced, e.g. /aba/ → [ap.a], since this satisfies *[+voice, -

son] and only violates lower-ranked ONSET and NOCODA. This is a general problem with output-oriented faithfulness constraints whether they refer to syllabification or the availability of cues. For example, IDENT(place)/released would be able to motivate realizing marked consonants as unreleased so that their marked properties could be eliminated.

(4) IDENT-ONSET(voice) >> *[+voice, -son] >> IDENT(voice), ONSET, NOCODA

2.4 Interim summary

The case studies reviewed here show that faithfulness is sensitive to language-specific patterns of phonetic realization such as the distribution of stop releases and partial devoicing of geminate obstruents. We have seen that it is not possible to account for the influence of these phonetic details on faithfulness violations by requiring input-output correspondence with respect to these details because the language-specific distribution of phonetic details cannot be observed in input representations, since these are governed by Richness of the Base. The distribution of phonetic details also cannot be inferred from the input on a language-independent basis precisely because the details are language-specific, so it is not possible to replace direct reference to phonetic details by reference to segmental context. E.g. IDENT(place)/_V is not equivalent to a constraint that refers to released consonants because in some languages consonants are also audibly released word-finally.

Language-specific patterns of phonetic realization are apparent in the output, but input-output correspondence constraints that refer to the presence of properties such as release bursts in the output cannot derive the required effects because the distribution of these properties can be affected by the very processes that input-correspondence constraints need to regulate.

These considerations indicate that there must be an intermediate representation, the Realized Input, between the 'unprocessed' input, which must meet the requirements of Richness of the Base, and the output, and that predictable properties that are relevant to the evaluation of faithfulness constraints are derived at the RI (cf. Jun 2002). Perceptual faithfulness constraints then require minimal perceptual change between the RI and the output.

The Realized Input (RI) is derived within the phonology, so language-specific differences in phonetic realization such as release vs. non-release of stops preceding obstruents or partial devoicing of geminates are apparent in the RI. Consequently the patterns analyzed by Kawahara (2006) are unproblematic: an input voiced geminate is mapped to a partially devoiced geminate in RI, so complete devoicing is a modest violation of faithfulness to the RI, given the perceptual similarity facts demonstrated by Kawahara. In the same way, the reduced cues to presence of C_1 in a VC_1C_2V cluster are apparent in the RI, so if a consonant must be deleted in the output, [VC_2V] is preferred over [VC_1V] because it is more faithful to the RI. This avoids the problem of trying to determine which consonant is better cued in the output, where deletion has already applied.

The following section outlines a proposal for incorporating this level of representation into phonological grammar, adapting the model presented in Flemming (2006).

3. Deriving Realized Inputs

The proposed organization of the phonology can be understood in terms of some basic desiderata that the RI must meet to serve the purpose outlined for it above. The basic requirement for the RI is that it must be phonetically detailed in order to allow for evaluation of perceptual similarity between RI and output. For example, we have seen that it must represent properties such as stop releases and partial devoicing of obstruents. It is also crucial that the distribution of these phonetic details must be grammatically regulated, for example in Russian RIs, stops preceding heterorganic obstruents are released, while in Korean they are not. These restrictions on distribution must be imposed by markedness constraints, so the RI must be derived from the input through the interaction of markedness constraints and correspondence constraints. But the RI is intended to serve as the input to phonological processes such as place assimilation, so restrictions on possible sound sequences are not derived at this level. For example, heterorganic clusters must be permitted in RI in all languages even if they are excluded in the output. The role of the RI is to specify the realization that the cluster would have if it were permitted, so that the faithfulness constraints can evaluate whether assimilation would constitute an acceptable deviation from that realization.

This conception of the RI is formalized in terms of a division of phonology into two basic components: the first, the Inventory, determines the basic inventory of contrasting segments types, akin to a phoneme inventory, while the second, the Phonotactics, determines the permissible sequences of those sounds and their ultimate realization. The possible inputs to the phonotactics consist of all sequences of segments drawn from the inventory. The RI specifies the expected phonetic realization for a given segment sequence. Input sequences are mapped onto RIs by a Phonetic Realization component which derives the specific coordination of articulatory gestures that would be used to realize the input string and their perceptual consequences. So there are no restrictions on possible segment sequences at the level of the RI, but the RI conforms to grammatically determined patterns of phonetic realization. Faithfulness constraints then require that the output should be perceptually as close as possible to the RI.

In a sense this arrangement reverses the usual ordering between phonology and phonetic realization in which the output of the phonology is passed to phonetic realization. Here phonetic realization applies before most phonological processes. This arrangement is motivated by evidence reviewed above that faithfulness constraints are sensitive to language-specific details of phonetic realization, so information about the patterns of phonetic realization must be accessible during the derivation of phonological output forms.

The proposed structure of the phonology is summarized in (5): There are three representations, the Input, the Realized Input, and the Output, with three associated components of the phonology: the Inventory, Phonetic Realization, and the Phonotactics. All draw upon a single ranking of constraints, but only a subset of constraints are applicable in any given component. The division of constraints between components is summarized in (6-7), and discussed in more detail below. The workings of each component are illustrated through analyses of the two phenomena discussed above: devoicing of geminates in Japanese loans, and the influence of stop releases on place assimilation.

(5)

'Rich Base'
↓ Inventory
Input
↓ Phonetic Realization
Realized Input
↓ Phonotactics
Output

(6) Markedness constraints that are applicable in each component are marked ✓

	Distinctiveness constraints	Segment-internal markedness constraints	Context-sensitive articulatory constraints	Prosodic constraints	Other
Inventory	✓	✓			
Phonetic realization		✓	✓	✓	
Phonotactics	✓	✓	✓	✓	✓

(7) Correspondence constraints that are applicable in each component are marked ✓

	Cue realization constraints	P-map correspondence constraint
Inventory		
Phonetic realization	✓	
Phonotactics		✓

3.1 Inventory

The selection of voicing and length contrasts among Japanese stops provides a simple illustration of the workings of the Inventory component. Japanese loanwords permit voiced and voiceless singleton and geminate stops, e.g. /k, g, kk, gg/, but, as discussed above, the voiced geminates are partially devoiced.

Inventories of contrasting sounds are shaped by three basic preferences: to maximize the perceptual distinctiveness of contrasts, minimize articulatory effort, and maximize the number of contrasting sounds (Flemming 2004, cf. Lindblom 1986). The preference to maximize the distinctiveness of contrasts is implemented in terms of a set of constraints that each set a minimum perceptual distance that must be satisfied by contrasts between sounds. Perceptual distances are calculated from representations that locate sounds in a multi-dimensional perceptual space, where the dimensions correspond to properties such as Voice Onset Time (VOT), formant frequencies, durations, etc. We will first lay out the dimensions that are relevant to voicing and gemination contrasts, then formulate distinctiveness constraints with respect to these dimensions.

The primary dimension relevant to the gemination contrast is duration. The perceptual magnitude of a change in speech segment duration is approximately proportional to the size of the duration change as a proportion of the initial duration (e.g. Klatt 1976:1219), so we focus here on the relative durations of consonants. The duration of a singleton stop is taken to be 1, and the duration of long consonants are specified as multiples of this duration. Kawahara (2006) found that geminates were about 2.5 times longer than singleton stops.

Japanese contrasts voiced and voiceless unaspirated stops.. Two of the dimensions relevant to stop voicing contrasts are the duration of voicing during the stop closure and VOT. It is closure voicing that is of most interest here, so for the VOT dimension we will simply distinguish VOT of 0 for voiced stops and 1 for voiceless stops. VOT does not differ between singletons and geminates in Japanese (Hirata & Whiton 2005). With regard to voicing during the stop closure, we hypothesize that the relevant perceptual dimension is the Closure Voicing Ratio (CVR) – i.e. the ratio between the duration of closure voicing and the duration of the consonant. This measure of voicing accounts for the fact that voiceless and partially devoiced geminates are more confusable than voiceless and fully voiced singletons, as discussed by Kawahara (2006). Kawahara found that the absolute duration of voicing during stop closure was very similar during voiced singletons and geminates, but that this constituted voicing throughout the closure in the case of singleton /g/, but yielded voicing of only 40% of the closure of geminate /gg/. So , voiced and voiceless singleton stops differ in having 100% vs. less than 20% closure voicing, whereas voiced and voiceless geminates differ by the smaller margin of 40% vs, less than 10% closure voicing. Note that the voiceless stops can have as much as 20% closure voicing because voicing from a preceding vowel persists for about 10ms into the stop closure. The absolute duration of closure voicing does not reflect the observed difference in distinctiveness between geminate and singleton voicing contrasts because both geminate and singleton stops have about the same duration of closure voicing. Snoeren, Hallé & Segui (2006) also found CVR to be a good measure of the perceived degree of voicing of partially voiced stops in French.

Distinctiveness constraints set minimum differences between contrasting sounds on these perceptual dimensions. So voicing contrasts are subject to constraints of the form shown in (8). A constraint MINDIST = CVR:0.4 requires that contrasting sounds differ by at least 0.4 units on the CVR dimension.

(8) MINDIST = CVR:0.4, MINDIST = CVR:0.6, MINDIST = CVR:1

The tendency to maximize the distinctiveness of contrasts can be limited by articulatory effort constraints. This is the case with geminate voicing contrasts in Japanese. The constraints in (7) favor a contrast between fully voiced and fully voiceless geminates, but sustaining vocal fold vibration during an obstruent is difficult, and the longer the duration of the obstruent, the more difficult it is to sustain voicing through the complete closure (Ohala 1983). These considerations motivate an effort constraint against voiced obstruents, *VOICED OBSTRUENT, and a higher ranked constraint specifically against long voiced obstruents *LONG VOICED OBSTRUENT (Hayes & Steriade 2004). The latter constraint bans voiced obstruents with a duration greater than 1.

Finally, it would be possible to satisfy all of the MINDIST constraints and constraints against voiced obstruents by allowing only voiceless geminate obstruents. In general, all distinctiveness constraints and effort constraints penalize contrasts and so would favor empty segment inventories if unopposed. These constraints are kept in check by a constraint that favors maximizing the number of contrasting segments, MAXIMIZE CONTRASTS. This constraint can be understood as expressing a requirement of efficient communication, since a larger number of contrasting sounds means that each sound spoken can distinguish more words, so each sound conveys more information. This is a positive constraint that selects the largest inventory of contrasts that has not been eliminated by higher-ranked constraints.

The required rankings among these constraints are summarized by the stratified ranking in (9): all the constraints in the top stratum must outrank both of the constraints in the lower stratum. The operation of the ranking is illustrated in tableau (10). In the Inventory component, candidates are sets of contrasting sounds. Here we consider a subset of the inventory, labial geminate stops, although we can generalize the results to all geminate obstruents since place is not relevant to these constraints. The tableau shows that the ranking in (9) derives a contrast between partially voiced and voiceless geminates (candidate b). The duration of a stop is indicated by subscripting, so [$p_{2.5}$] is a voiceless labial stop with a duration of 2.5. Partially voiced stops are transcribed as a sequence of a voiced stop followed by a voiceless stop with each component subscripted with its duration, so [$b_1p_{1..5}$] is a stop with a total duration of 2.5 whose voiced portion has a duration of 1. Here we only consider geminates with a duration of 2.5.

Maximization of distinctiveness favors a contrast between fully voiced and voiceless geminates (candidate a) – i.e. this candidate satisfies both MINDIST constraints. However, this candidate is dispreferred because fully voiced geminates violate higher-ranked *LONG VOICED OBS. The winning candidate (b), satisfies this effort constraint since the duration of voicing during the voiced geminate is only 1. Finally, the MINDIST effort constraints could be satisfied by not selecting a voicing contrast (candidate c), but this option violates high-ranked MAXIMIZE CONTRASTS.

(9) MAXIMIZE CONTRASTS, MINDIST = CVR:0.4, *LONGVOICEDOBS
>>
MINDIST = CVR:1
>>
*VOICEDOBS

(10) Inventory

	MAXIMIZE CONTRASTS	*LONG VOICEDOBS	MINDIST = CVR:0.4	MINDIST = CVR:1	*VOICED OBS
a. $p_{2.5}$, $b_{2.5}$	2	*!			*
b. ☞ $p_{2.5}$, $b_1p_{1.5}$	2			*	*
c. $p_{2.5}$	1!				

Note that the duration of voicing is only reduced to 1 because MINDIST = CVR:0.4 outranks the constraint against singleton voiced obstruents. Voicing of duration 1 during a geminate with duration of 2.5 yields a CVR of 0.4 (=1/2.5), so any further reduction in the duration of voicing, favored by *VOICE OBS, would violate MINDIST = CVR:0.4.

The same ranking derives fully voiced singleton stops (11). Fully voiced singletons do not violate *LONGVOICEDOBS, so the only effort constraint opposing maximizing the distinctiveness of the voicing contrast is *VOICED OBS which is outranked by MINDIST=CVR:1.

(11) Inventory

	MAXIMIZE CONTRASTS	*LONG VOICEDOBS	MINDIST = CVR:0.4	MINDIST = CVR:1	*VOICED OBS
a. ☞ p_1, b_1	2				*
b. p_1, $b_{0.5}p_{0.5}$	2			*!	*
c. p_1	1!				

Note that the analysis here abstracts away from the contribution of VOT to the voicing contrasts: voiced and voiceless stops differ in VOT as well as closure voicing ratio. So presumably the MINDIST constraints properly require differences on both dimensions, e.g. MINDIST = CVR:0.4 & VOT:1. However, VOT is unaffected by partial devoicing of geminates, judging from the spectrograms shown in Kawahara (2006), so this factor is orthogonal to the analysis of partial devoicing, which is our main concern here.

To summarize, the selection of an inventory of contrasts involves a compromise between MAXIMIZE CONTRASTS, the MINDIST constraints and effort constraints. Since inventories involve individual segments, only segment-internal effort constraints are applicable. So *VOICED OBS applies, but a context-specific constraint against voiced obstruents following voiceless obstruents would not be relevant. The output of the inventory component is a set of contrasting sounds, specified in terms of target values on the various perceptual dimensions, such as CVR and duration. This set of sounds is then used to construct the inputs to the rest of the phonology.

3.2 Phonetic Realization

Input sequences of segments drawn from the inventory are mapped onto RIs by a process of phonetic realization. This step does not play an important role in the analysis of Japanese geminate devoicing – the partial devoicing of geminates derived in the Input is simply carried over into the RI – so we will briefly outline the workings of this component here, then return to it in more detail in the analysis of the distribution of stop bursts in consonant clusters.

Phonetic Realization involves resolving the conflict between the need to realize the perceptual targets of the input segments with constraints on articulatory effort and timing. Realization of perceptual targets is required by a set of correspondence constraints which we will term Cue Realization constraints to distinguish them from the correspondence constraints that evaluate the perceptual similarity between RI and Output.

The RI is conceived as the expected phonetic realization of the input segment sequence, without substantial changes to the input, such as neutralizing deletions, assimilations etc. Accordingly, we hypothesize that only markedness constraints that directly affect articulation apply in Phonetic Realization. In addition to effort constraints, this includes any constraints on the coordination of articulatory gestures (cf. Gafos 2002), and constraints on segment duration. Constraints on metrical and prosodic structure must apply in this component, given that these properties interact directly with segment duration and timing, but other markedness constraints are restricted to the Phonotactics.

In particular, distinctiveness constraints do not apply in Phonetic Realization. This is central to the difference between phonetic realization and phonotactics because distinctiveness constraints provide the motivation for neutralization processes. That is, contrasts are neutralized in contexts where they would be insufficiently distinct (Steriade 1997, Flemming 2002). The RI is the most faithful realization of the cues to the contrasting segments that is possible given the restrictions imposed by articulatory and prosodic constraints, so in general it preserves all of the segments from the input, although possibly in a rather reduced form. As will be illustrated in the section 3.4, where constraints on phonetic realization yield a contrast that would violate high-ranking distinctiveness constraints in the phonotactic component, the contrast may be neutralized in the output.

In Japanese, the closure of voicing of the voiced stops is faithfully realized in RI, i.e. voiced singletons are fully voiced, and voiced geminates are partially voiced. This is derived by ranking IDENT-CR(CVR) above *VOICED OBS, the effort constraint that penalizes voiced stops. IDENT-CR(CVR) is a correspondence constraint that requires that the CVR values of corresponding sounds in the input and RI must be the same. The constraint name includes –CR to indicate that it is a cue realization constraint.

3.3 Phonotactics

The devoicing of geminates in presence of a second voiced obstruent is derived in the phonotactic component. The analysis is very similar in outline to Kawahara's: Devoicing is motivated by a general constraint against the cooccurrence of voiced obstruents in a stem, OCP(+VOICE). This markedness constraint conflicts with correspondence constraints requiring faithfulness to the RI with respect to voicing. Constraints on correspondence between RI and output require that the output deviate minimally from the RI in perceptual terms, implementing the P-map hypothesis. The notion of perceptual distance that is relevant to the evaluation of correspondence is the same as the perceptual distance that is relevant to evaluating distinctiveness of contrasts (Flemming 2008). Accordingly, faithfulness to closure voicing is assessed in terms of the Closure Voicing Ratio, so devoicing of a fully voiced obstruent involves a change from [CVR 1] to [CVR 0] and is this a greater violation of faithfulness than fully devoicing a partial devoiced obstruent (a change from [CVR 0.4] to [CVR 0].

A set of correspondence constraints set thresholds for the maximum acceptable change in closure voicing ratio between RI and output: IDENT(CVR)<*d* penalizes differences of *d* or greater in CVR between corresponding segments. The crucial constraints here are IDENT(CVR)<1 and IDENT(CVR)<0.5. OCP(+voice) is variably ranked with respect to IDENT(CVR)<0.4. When the markedness constraint ranks higher, it is optimal to devoice a partially voiced geminate, because this is a change of 0.4 on the CVR dimension which only violates the lower-ranked faithfulness constraint (12i). The reverse ranking derives the possibility of retaining partial voicing of the geminate. IDENT(CVR)<1 is always ranked above OCP(+voice), so it is preferable to violate the coocurrence constraint rather than devoicing a singleton obstruent (12ii).

(12) IDENT(CVR)<1 >> OCP(CVR), IDENT(CVR)<0.4

(13) Phonotactics

i.	/bagku/	IDENT (CVR)<1	OCP (CVR)	IDENT (CVR)<0.4
a.	☞ bagku		*	
b.	☞ bakku			*
c.	pagku	*!		

ii.	/bagu/	IDENT (CVR)<1	OCP (CVR)	IDENT (CVR)<0.4
a.	☞ bagu		*	
b.	baku	*!		*
c.	pagu	*!		

The effort constraints, *VOICED OBS and *LONG VOICED OBS must remain in force in the Phonotactics because they must continue to restrict patterns of realization in the final output. Without these effort constraints, distinctiveness constraints on stop voicing contrasts would favor restoring full voicing to geminate obstruents in the output. The ranking of these constraints is the same as in the inventory. That is there is no re-ranking of constraints between components, although only subsets of the full constraint set apply in any given component. For example, segment-internal effort constraints and distinctiveness constraints apply in both the Inventory and the Phonotactics, and these constraints are ranked in the same way in both components. In this case the ranking *LONGVOICEDOBS >> MINDIST = CVR:1, motivated for the inventory component in (8) above, also applies in Phonotactics, preventing full voicing of geminates. In addition IDENT(CVR)<0.4 ranks above *VOICED OBS, preventing this markedness constraint from motivating devoicing of obstruents in the Phonotactics.

This analysis implements Kawahara's analysis of the difference in patterning of geminate and singleton voiced stops, but the implementation developed here avoids the problem of having to derive the ranking of faithfulness constraints from language-specific facts about phonetic realization. Instead the language-specific partial devoicing of geminates is specified in the RI. The fact that geminates are subject to devoicing where singletons are not then follows from the general principle that larger perceptual deviations from the input violate more faithfulness constraints, so devoicing a fully voiced stop is a greater violation of faithfulness than devoicing a partially devoiced stop.

3.4 Stop releases and place assimilation

We turn now to our second case study illustrating that faithfulness constraints are sensitive to language-specific phonetic detail: the effect of stop releases on place assimilation. As discussed above, Jun (2002) provides evidence that the possibility of place assimilation targeting stops depends on whether the stop is audibly released or not. The RI allows for a straightforward analysis of this effect because the distribution of stop releases is derived in phonetic realization, so the faithfulness violation incurred by a change in place of articulation of a stop depends on whether the stop is released or not in RI.

Stop place contrasts are distinguished by formant transitions and the quality of the stop burst, so these properties are specified in the Inventory. The transition features of a consonant are associated with 'sub-segments' distinct from the consonantal constriction, so the full specification of a consonant in the inventory consists of at least three feature matrices. In the case of a plosive, there is an additional matrix for the burst (49, cf. Flemming 2002:23ff.). Features of the burst specify target properties, such as the frequency of the peak in the burst spectrum (Noise Frequency) (Flemming 2002:23).

(14) $$\underset{\textit{closure transition}}{\begin{bmatrix} F1 & 1 \\ \mathrm{F2} & 4 \end{bmatrix}} \underset{\textit{closure}}{\begin{bmatrix} \text{Loudness} & 1 \\ & \end{bmatrix}} \underset{\textit{burst}}{\begin{bmatrix} \mathrm{NF} & 6 \\ & \end{bmatrix}} \underset{\textit{release transition}}{\begin{bmatrix} \mathrm{F1} & 1 \\ \mathrm{F2} & 4 \end{bmatrix}}$$

Not all of these cues are realized in every context. As discussed above, stop bursts are often absent preceding obstruents, and transition cues can generally only be realized where there is an adjacent vowel. Loss of burst cues results when stops are produced in close transition with a following obstruent. This pattern of realization is analyzed in terms of a constraint OVERLAP, which requires consonants that are adjacent in the input to be articulatorily overlapped– i.e. the formation of the constriction for the second consonant must precede the release of the first (cf. Gordon 2001). This constraint is essentially a cue realization constraint requiring realization of the adjacency of consonants in the input by avoiding generating any distinct interval between the consonant constrictions that could be misconstrued as an intervening segment.

Satisfying OVERLAP results in failure to realize release cues to C_1, violating Cue Realization constraints. Cue realization constraints are formalized in terms of a MAX/DEP/IDENT model, along the lines presented in McCarthy & Prince (1995). That is, the realization of input segments and sub-segments is required by MAX constraints and the realization of individual perceptual targets is required by IDENT(F) constraints, e.g. IDENT(CVR). That is, the target for a stop specifies both the presence of a burst, and particular properties of that burst. For example, stop bursts differ in NF – labial bursts have low NF while alveolars have high NF bursts. The proposed system of cue realization constraints distinguishes failure to realize a burst at all, violating MAX BURST, from realization of a burst which deviates from the specified target properties in some respect. For example the NF value of the burst spectrum can vary due to coarticulation with a following vowel, deviating from its target in violation of IDENT(NF).

Realization of the transition sub-segments is required by two cue realization constraints, MAX TRANSITION, which requires the realization of at least one transition per consonant, and MAX RELEASE TRANSITION, which specifically requires the realization of release transitions (15). Realization of bursts is required by MAX BURST.

(15) MAX TRANSITION: For each consonant, at least one transition segment must have a correspondent in the Realized Input.
MAX RELEASE TRANSITION: The release transitions of a consonant must have a correspondent in the Realized Input.
MAX BURST: A burst segment in the input must have a correspondent in the Realized Input.

The mapping from Input to Realized Input is illustrated in (16). In the ranking shown there, OVERLAP outranks MAX RELEASE TRANSITION and MAX BURST so consonant clusters are overlapped in the RI, indicated in the tableau by the 'unreleased' diacritic on the first stop. Neither a transition nor an audible burst can be generated if the release of the first consonant follows the formation of the constriction for the second, so realization of the burst of the first consonant (candidate b), or its burst and release transition (candidate c) violates OVERLAP. If MAX BURST ranks above OVERLAP, we derive a pattern where stop-obstruent clusters are realized with open transition, realizing the stop burst. Realization of the release formant transitions requires epenthesis of vowel between the consonants of the cluster, violating both OVERLAP and DEPV.

(16) Phonetic Realization

	apga	OVERLAP	MAX RELTRANS	MAX BURST	DEPV
a.	☞ ap̚ga		*	*	
b.	$\text{ap}^{\text{h}}\text{ga}$	*!	*		
c.	apəga	*!			*

The presence or absence of stop releases in the RI then affects the possibility of neutralizing place assimilation occurring in the Phonotactic component. We follow Jun (1995, 2004) and Flemming (2008) in analyzing major place assimilation as resulting from neutralization of place contrasts in C_1 position of a C_1C_2 cluster. Neutralization is motivated where the place contrasts would otherwise be insufficiently distinct, as determined by distinctiveness constraints. It is common for C_1 to assimilate in place to obstruent C_2 because fewer cues to place can be realized in this context compared to the prevocalic context. In prevocalic contexts stop place is cued by the burst and release formant transitions, but only the burst can be realized in pre-obstruent contexts, and then only if open transition is acceptable.

The relevant distinctiveness constraints can be given a simplified formulation as in (17) (cf. Flemming 2008). The first ranking reflects the fact that bursts contribute to the distinctiveness of place contrasts (e.g. Kochetov & So 2007), while the second ranking reflects the generalization that release transitions are perceptually more distinct than formant transitions into the closure (Fujimura et al 1978).

(17) MINDIST = {closure transitions} >>
MINDIST = {closure transitions & burst} >>
MINDIST = {release transitions & burst}

A contrast like [at˺ka] vs. [ak˺ka] violates MINDIST = {closure transitions & burst} and MINDIST = {release transitions & burst}. This violation can be eliminated by neutralizing the contrast through assimilation, i.e. [at˺ka] → [ak˺ka], however this violates faithfulness constraints because it constitutes a perceptual change to the RI. As above, the magnitude of this perceptual change is measured in the same terms as the distinctiveness of contrasts, so the ranking of constraints on faithfulness to place cues is as shown in (18).

(18) IDENT{release transitions & burst } >>
IDENT{closure transitions & burst } >>
IDENT{closure transitions}

This analysis makes the occurrence of place assimilation dependent on whether pre-obstruent stops are released in the language in question, a property that is represented in RI. For example, if the markedness and faithfulness hierarchies in (17) and (18) are combined into a ranking as in (19), then we derive assimilation if stops are unreleased in clusters in RI, but no assimilation if stops are released in this context.

(19) IDENT{closure transitions & burst },
MINDIST = {closure transitions & burst} >>
IDENT{closure transitions},
MINDIST = {release transitions & burst}

The tableau in (20) illustrates the case where Phonetic Realization derives clusters with close transition, as in (16) above. For MINDIST constraints 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). The tableaux here illustrate the evaluation of a stop place contrast preceding another stop. Phonetic Realization as in (16) above maps these inputs onto RIs with unreleased stops (shown in the top left cell of the tableau). These RIs are realized faithfully in candidate (a), whereas the place contrast is neutralized in candidates (b) and (c). Neutralization is achieved by assimilation of [t] to [k] in candidate (c), but by neutralization to [t] in candidate (b).

A place contrast between unreleased stops, as in (a), is only distinguished by closure transitions so it violates MINDIST={clos trans & burst}. As a result is preferable to neutralize this contrast because the change in place of articulation of an unreleased stop only violates lower-ranked IDENT(clos trans) (candidates b, c). Neutralization of the place contrast results in assimilation because this is the least effort realization of the neutralized stop since the resulting cluster involves a single constriction gesture rather than the two constriction gestures required for a heterorganic cluster (Jun 2004). This analysis is implemented in terms of the constraint *GESTURE, which penalizes consonantal gestures.

(20) Phonotactics – unreleased stops

	I /atka, akka/ RI [at˺ka, ak˺ka]	IDENT {clos trans & burst}	MINDIST = {clos trans & burst}	IDENT {clos trans}	MINDIST = {rel trans & burst}	*GESTURE
a.	at˺ka, ak˺ka		*!		*	***
b.	at˺ka			*		**!
c.	☞ ak˺ka			*		*

Tableau (20) shows that the same ranking of Phonotactic constraints does not give rise to place assimilation where stops are released in RIs. The aspiration diacritic [^{h}] has been used to emphasize the presence of release bursts. The faithful realization of the contrast between these RIs satisfied MINDIST={clos trans & burst} (a), while neutralization would violate higher-ranking IDENT{clos trans & burst}, so it is preferable to maintain the contrast.

(21) Phonotactics – released stops

	I /atka, akka/ RI [at^{h}ka, ak^{h}ka]	IDENT {clos trans & burst}	MINDIST = {clos trans & burst}	IDENT {clos trans}	MINDIST = {rel trans & burst}	*GESTURE
a.	☞ at^{h}ka, ak^{h}ka				*	***
b.	at^{h}ka	*!		*		**
c.	ak^{h}ka	*!		*		*

Representation of stop bursts in the RI makes it possible for faithfulness constraints to be sensitive to the distinction between released and unreleased stops. Jun's (2002) generalization that released stops do not undergo place assimilation can then be analyzed in terms of a fixed ranking of IDENT{clos trans & burst} >> MINDIST = {clos trans & burst}.

This analysis also illustrates the importance of restricting distinctiveness constraints in the phonotactic component but not in phonetic realization. In (20), stop place contrasts are neutralized in C_1 position of the C_1C_2 cluster because the place contrast would be insufficiently distinct in the absence of release burst cues – i.e. the contrast would violate MINDIST = {clos trans & burst}. The absence of stop bursts in this context is derived in phonetic realization (16), so if the distinctiveness constraints applied at this level, they could motivate neutralization of place contrasts in RI. But the point of the RI is to serve as a point of reference in evaluating the faithfulness violations incurred by neutralization. If the neutralization is derived in RI, it obviously could not serve this function. Without distinctiveness constraints, there is no motivation to neutralize contrasts, and in general all segments from the input are realized in some form.

3.5 The Inferred Input

Jun (2002) proposes a closely related solution to the problem of allowing faithfulness constraints to be sensitive to phonetic details that are predictable on a language-specific basis like stop releases. He proposes an intermediate representation between the input and output, the inferred input, which can be referred to by faithfulness constraints. This representation is 'identical with the input except that it includes all phonetic details', so it is very similar in concept to the RI. However the derivation of this representation is rather different: the inferred input is the most harmonic candidate that obeys all input-output faithfulness constraints – that is, among the candidates that obey all input-output faithfulness constraints, the inferred input is the one that also best satisfies the ranking of markedness constraints. This characterization of the inferred input depends on the claim that the language-specific phonetic details that are relevant to faithfulness, like stop release, are not themselves subject to input-output faithfulness constraints.

This assumption is not consistent with the approach developed here. We have proposed that phonological representations specify the perceptual cues that are relevant to evaluating the perceptual distance between forms for the purposes of evaluating correspondence between RI and Output and the distinctiveness of contrasts. Accordingly, phonetic realization concerns the distribution of cues to contrasts according to their contexts. Realization of cues specified in the inventory is required by cue realization constraints. These correspondence constraints must favor the realization of all cues since they all contribute to the distinctiveness of contrasts. So we have analyzed realization of an unreleased stop as violation of faithfulness to the release burst specified in the inventory (e.g. (16) above). So the RI is not characterized in terms of satisfaction of faithfulness constraints, instead it is characterized by the limited set of markedness constraints that apply at that level – i.e. constraints on articulatory effort, gestural coordination and segmental timing.

4. Opacity and displaced contrasts

Positing the RI as an intermediate representation between input and output predicts the possibility of opaque interactions between phonetic realization and phonotactics, since the output could be faithful to properties that are derived in phonetic realization although the conditioning environment for those properties is lost in the output (cf. Gallagher 2007). In this section we will see that this sort of opacity is attested, and provides the basis for an account of patterns in which an underlying contrast is displaced or transformed in the output. An example comes from Friulian, where an underlying voicing contrast is realized as a vowel length contrast in word-final position (Baroni & Vanelli 2000). That is, obstruents are devoiced in word-final position but the underlying contrast is preserved because vowels preceding devoiced obstruents are lengthened (22). The suffixed forms in (22) show that there is an underlying voicing contrast, /lad/ vs. /lat/, but where the stops are word-final, they are devoiced, but the contrast is preserved as a difference in vowel duration.

(22)	laːt	'gone (masc.)'	lad-e	'gone (fem.)'
	lat	'milk'	lat-a	'to breast-feed'

This pattern looks like an opaque interaction between vowel lengthening before word-final voiced obstruents and final devoicing, but as Baroni & Vanelli point out, a simple rule-ordering account in which vowel lengthening is ordered before final devoicing is not satisfactory because it misses the generalization that this degree of vowel lengthening is only found in precisely the context where obstruents are devoiced, so the word-final context has to be specified independently in the lengthening rule and the devoicing rule. Moreover, this analysis does not connect final vowel lengthening to a more general effect of obstruent voicing on preceding vowel duration in Friulian: vowels are also longer preceding voiced stops than preceding voiceless stops in medial contexts where no devoicing arises, as in many other languages (Chen 1970, Keating 1985). The length difference is smaller in these contexts, so the two lengthening processes cannot simply be unified, but this effect of voicing on preceding vowel duration plausibly explains why a voicing contrast might be realized as a vowel duration contrast.

The existence of a general vowel length difference conditioned by obstruent voicing indicates that the Friulian pattern does not involve displacement of a contrast from one dimension to another so much as a contextual shift in the relative contributions of cues to voicing. In all languages intervocalic voicing contrasts are distinguished on multiple dimensions (cf. Lisker (1986) on cues to voicing in English), but not all of these cues are available or equally important in all contexts. For example, in English, Voice Onset Time is an important cue to voicing in prevocalic stops (e.g. Lisker & Abramson 1970), but this cue is not available in final contexts, and preceding vowel duration becomes a central cue to the contrast in this context (Raphael 1972). The situation in Friulian is actually quite similar: correlates of voicing in intervocalic contexts include presence vs. absence of voicing in closure, closure duration and preceding vowel duration, but in final position closure voicing is not available, and preceding vowel duration appears to be the primary cue to the voicing contrast. So the magnitude of the vowel duration cue is exaggerated in order to compensate for the loss of the closure voicing cue (Baroni & Vanelli 2000:36). The shift in the realization of voicing seems to be more extreme in Friulian, in that closure voicing is completely lost, but the basic pattern is quite similar to the realization of voicing contrasts in English.

The framework proposed here provides the basis for a formalization of this analysis. The fact that a difference in preceding vowel duration is a general correlate of the voicing contrast is represented in the RI, and is thus independent of whether closure voicing actually appears in the output. That is, generation of the vowel duration difference in RI can interact opaquely with devoicing in the output. The fact that the vowel duration difference is exaggerated where voicing cues are lost is accounted for in terms of distinctiveness constraints applying in phonotactics. Distinctiveness constraints can motivate enhancing a difference along one dimension to compensate for loss of cues along another dimension, in order to maintain the overall distinctiveness of the contrast.

In the Inventory, voiced and voiceless obstruents are differentiated by targets for closure voicing and duration, among other dimensions, so [t] has targets of [CVR 0, duration 1.6] and [d] has targets of [CVR 1, duration 1]. A short stop is taken as having a duration of 1, while the other values are ratios of this duration based approximately on measurements reported in Baroni & Vanelli (2000).

A partial analysis of this inventory is shown in (24), using the conventions for transcription of duration summarized in (23). The short duration of voiced stops is enforced by a segment-internal effort constraint, VOICE→SHORT, which requires that voiced obstruents should be short due to the difficulty of sustaining voicing during an obstruent, discussed in section 2.2 above. Specifically this constraint states that voiced segments must have a duration of 1. This outranks a constraint NORMAL C DURATION which stipulates that the preferred duration for a stop is 1.6. Voiceless stops receive this preferred duration. Vowels are assumed to have a duration target of 3 in Inventory.

(23) C duration: / V duration:

short	normal		short	normal	long
1	1.6		2.4	3	4.8
t	tˑ		ă	a	aː

(24) Inventory

		VOICE→ SHORT	MINDIST = CVR:1	MAXIMIZE CONTRASTS	NORMAL C DURATION
a.	☞ tˑ, d			2	*
b.	tˑ, dˑ	*!		2	
c.	tˑ			1!	

The fact that vowels are shorter before voiceless stops than before voiced stops is derived in Phonetic Realization. We adopt the hypothesis that this vowel duration difference has its basis in duration compensation (Maddieson 1997): vowel duration varies inversely with the duration of a following stop due to constraints establishing a preferred rhyme duration (Flemming 2001) or a preferred duration between vowel onsets (McCrary 2004:205ff.). For present purposes this effect is analyzed in terms of a COMPENSATION constraint requiring the duration of a vowel and a following consonant to sum to 4. This constraint competes with a cue realization constraint requiring faithful realization of vowel duration targets, IDENT(Vdur)-CR. VOICE→SHORT and NORMAL C DURATION continue to enforce consonant durations.

(25) VOICE→SHORT: Voiced obstruents must have duration ≤ 1.
COMPENSATION: In a VC sequence, C duration + V duration = 4.

The crucial rankings among these constraints are shown in (25). The tableaux in (26) illustrate how this ranking derives the effect of voicing on preceding vowels in phonetic realization. Where a vowel is followed by a voiceless stop, (26i), the duration targets for vowel and consonant sum to 4.6 (3+1.6) violating COMPENSATION (b). This constraint could be satisfied by shortening the vowel (candidate a) or the consonant (candidate c). The former is preferred because the constraints regulating consonant duration outrank faithfulness to the target vowel duration. A vowel plus a voiced stop satisfies COMPENSATION without deviating from target durations for either segment (3+1) (26ii), so these targets are faithfully realized. Thus vowels are shortened before voiceless stops, but not before voiced stops, deriving the difference in preceding vowel duration as a correlate of the voicing contrast. Exactly the same pattern of realization applies with final stops since they are subject to the same constraints: /latˑ/→[lătˑ], /lad/→[lad].

(26) a. VOICE→SHORT >> NORMAL C DURATION >> IDENT(Vdur)-CR
b. COMPENSATION >> IDENT(Vdur)-CR

(27) Phonetic Realization – medial stops

i.

	latˑa	VOICE→ SHORT	NORMAL C DURATION	COMPENS-ATION	IDENT (Vdur)
a.	☞ lătˑa				*
b.	latˑa			*!	
c.	lata		*!		

ii.

	lade	VOICE→ SHORT	NORMAL C DURATION	COMPENS-ATION	IDENT (Vdur)
a.	☞ lade		*		
b.	lăde		*	*!	*
c.	lădˑe	*!			*

This difference in vowel duration is exaggerated in precisely the context where closure voicing is lost as a cue to the voicing contrast. This effect is attributed to distinctiveness constraints, which apply in the phonotactic mapping. Medial voicing contrasts are adequately distinct based on closure voicing and VOT, but final devoicing leaves the final voicing contrast insufficiently distinct. The final contrast can be made adequately distinct by increasing the difference in vowel duration, lengthening vowels where they are already longer in RI, i.e. preceding underlying voiced obstruents. There are two basic elements to this analysis: first, the output is required to be faithful to the RI, and the greater duration of vowels preceding voiced obstruents is represented in RI, so faithfulness constraints favor preserving this difference even if closure voicing is not maintained. Second, the output is subject to MINDIST constraints which can favor exaggerating a difference between contrasting forms where this is necessary to yield a sufficiently distinct contrast.

The devoicing of final stops is attributed to a constraint against closure voicing in word-final voiced obstruents, *VOICEDFINALOBS, which ranks above IDENT(CVR), which requires faithful preservation of closure voicing. The exaggeration of the vowel length difference in the absence of closure voicing is motivated by a distinctiveness constraint, MINDIST=CVR:1 or Vdur:2.4, which is satisfied by a difference in voicing or a large difference in vowel duration. This aspect of the phonotactic mapping is illustrated in (28). The inputs in these tableaux are the RIs derived above. Only candidates with final voiceless stops are considered here, so the constraints *VOICEDFINALOBS >> IDENT(CVR) are not shown.

Devoicing of final obstruents mean that the distinctiveness constraint MINDIST=CVR:1 or Vdur:2.4 cannot be satisfied by a contrast in CVR, so candidate (b) which has undergone final devoicing, but is faithful to the RI with respect to segment durations, violates this constraint. Enhancing the durational contrast by lengthening the vowel preceding the underlying voiced stop (candidate a) satisfies the distinctiveness constraint. The distinctiveness constraint is undominated, so it can motivate lengthening even though this is unfaithful to the RI [lad], violating IDENT(Vdur), and creates an overlong VC sequence, in violation of COMPENSATION. Enhancing the contrast is also preferable to neutralizing it because neutralization would entail violations of IDENT(Cdur) (candidates c, d).

(28) Phonotactics: final voicing contrasts.

	lad, lătˑ	MINDIST= CVR:1 or Vdur:2.4	IDENT (Cdur)	COMPENS-ATION	IDENT (Vdur)
a.	☞ laːt, lătˑ			*	*
b.	lat, lătˑ	*!			
c.	lătˑ		*!		*
d.	lat		*!		*
e.	lăt, laːtˑ			**!	**

Note that it is crucial that the effect of voicing on vowel duration provides the seeds for the vowel length contrast. Although the inverted pattern in which vowels are lengthened before underlying voiceless stops and shortened before underlying voiced stops would satisfy the distinctiveness constraint (candidate e), this candidate is dispreferred because both realizations violate COMPENSATION and faithfulness to vowel duration. More generally, enhancement of a contrast to satisfy a distinctiveness constraint will only be optimal if it involves a smaller violation of faithfulness than neutralizing the contrast, which will generally only be the case if the enhancement can be achieved by a modest increase in an existing difference (cf. Flemming 2008).

Medial voicing contrasts are not enhanced by further vowel lengthening (29) because they already satisfy the MINDIST constraint (candidate b), so lengthening would be a gratuitous violation of IDENT(Vdur).

(29) Phonotactics: Medial voicing contrasts

	lada, lătˑa	MINDIST= CVR:1 or Vdur:2	IDENT (Cdur)	COMPENS-ATION	IDENT (Vdur)
a.	laːda, lătˑa			*!	*
b.	☞ lada, lătˑa				
c.	lătˑa		*!		*
d.	lada		*!		*

The full ranking of constraints in the phonotactic component is shown in (30) (dominance relations are indicated by downward arrows). VOICE→SHORT also applies in the phonotactic component, but its ranking is unimportant to the analysis since short duration of voiced stops is also enforced by IDENT(Cdur). Since rankings are the same in all components, we assume that is ranked above NORMAL C DURATION, as required for phonetic realization (27).

(30)

$$\begin{tikzcd}
\text{*FINALVOICEDOBS} \arrow[d] \arrow[dr] \arrow[drr] & \text{MINDIST=CVR:1 or Vdur:2.4} \arrow[d] \arrow[dr] & & \text{IDENT(Cdur)} \arrow[dl] \arrow[d] \\
\text{IDENT(CVR)} & \text{IDENT(Vdur)} & \text{COMPENSATION} & \text{NORMAL C DUR}
\end{tikzcd}$$

This line of analysis is applicable to a range of phenomena involving contextual variation in the contribution of cues to the realization of a contrast, ranging from 'low-level' phonetic effects to apparent transformations of contrasts, as observed in Friulian.

As observed above, the role of vowel duration as a cue to voicing in languages like English varies according to contexts in a way that looks very similar to Friulian, although less extreme. That is, the difference between the durations of vowels preceding intervocalic voiced and voiceless obstruents is relatively small, but there is a much larger difference preceding word-final obstruents (Port 1981). The difference in vowel duration preceding intervocalic voiced and voiceless stops is reasonably well accounted for as a duration compensation effect, i.e. the difference in vowel duration between pairs like *dipper* and *dibber* is about the same as the difference in the duration of the following stop, so the duration of the vowel+stop sequence is similar in both words. However, the difference in vowel duration preceding word-final voiced and voiceless consonants is larger than can be explained in terms of compensation for the duration of the following obstruents. That is, the difference in vowel duration between pairs like *dip* and *dib* is larger than the difference in the durations of the stops, [b] vs. [p]. This larger difference in vowel duration before word-final obstruents can be analyzed as an exaggeration of a difference derived by duration compensation in order to maintain the distinctiveness of the voicing contrast in a context where VOT cues are not available.

This analysis of voicing effects on vowel duration closely parallels the analysis of Friulian, but differs in the magnitude of the effects: in Friulian, final obstruents are completely devoiced, whereas in English it is primarily VOT that is lost in this position (although devoicing of fricatives is possible (Smith 1997)). The voicing-related difference in vowel duration before final obstruents is correspondingly smaller in English compared to Friulian – vowels are about 30%-50% longer before final voiced obstruents than before voiceless obstruents (Port 1981, Peterson & Lehiste 1960), while the difference is close to 100% in Friulian (Baroni & Vanelli 2001).

Another example in which the realization of a contrast appears to be more radically changed in a particular context involves the realization of low tone as downstepping of a following high tone. For example, in Lama (Ourso 1989), a sequence of a falling tone and a high tone /HL.H/ is realized as a high tone followed by a downstepped high, so the contrast between falling and high tones is realized by downstep in this context (31).

(31)	H stem:	wáːl	'husband'	wáːl tɛ́ 'at the house of husband'
	HL stem:	nâː	'cow'	náː tꜜɛ́ 'at the house of cow'

This kind of downstep is often analyzed as the surface realization of a floating Low tone – according to this analysis, the low tone from the underlying HL fall on the word 'cow' is not deleted in non-final position, it is delinked from the vowel but remains floating in the output representation, and a floating low tone is phonetically realized as downstepping of a following high tone (e.g. Clements & Ford 1979). This analysis has the disadvantage that it gives tone an exceptional status, since there is little evidence for positing other unassociated features than can affect phonetic realization (Clark 1993:32f.). In the present framework, downstep can be analyzed as an opaque interaction between downdrift and contour simplification. Downdrift is a phenomenon by which high tones are phonetically lowered following a low tone, so the H tones in an HLHLH sequence are produced with progressively lower f0 (Hombert 1974). This pattern is very common in African languages, being found in Twi, Hausa and Igbo among many others (Hyman & Schuh 1974), and may be related to the fact that a fall in f0 (from H to L) can be produced more quickly than an equivalent rise in f0 (from L to H) (1978, Xu & Sun 2002), so given equal time for L to H and H to L transitions, the L to H rises will be smaller, resulting in progressive reduction in the height of H tones (Hombert 1974). These constraints on the rate of f0 rises and falls can derive downdrift in the RI, so the RI of an HLH sequence assigns a lower pitch to the second H compared to the first. Downstep as in Lama arises where faithfulness to pitch favors preserving this difference in the pitch of high tones in the output, even where contour simplification constraints in the phonotactic component enforce elimination of the f0 minimum associated with the L tone.

This approach to the analysis of displaced or transformed contrasts has advantages over an alternative proposed by Lubowicz (2003) based on contrast preservation constraints. Lubowicz proposes that Friulian vowel lengthening is motivated by constraints that require the preservation of input contrasts but are indifferent as to the form in which the contrast is realized. This analysis captures the idea that lengthening applies precisely where devoicing threatens to neutralize a contrast, but it offers no account of the fact that the contrast is preserved through lengthening of vowels preceding underlying voiced stops rather than by introducing a difference in some other property, e.g. nasalization of the preceding vowel (Lubowicz 2003:150). That is, /lat, lad/→[lat, lãt] would satisfy a contrast preservation constraint just as well as /lat, lad/→[lat, laːt], but the former is unattested. In the same way, low tone can be realized by reducing the pitch of a following high tone, but not by raising the high tone, nasalizing a vowel etc.

The present analysis accounts for the relationship between voicing and vowel duration by positing an intermediate level of representation, the RI, where the effect of obstruent voicing on preceding vowel duration is represented. In the same way, low tones can lower a following high tone in RI through the coarticulatory effects outlined above. We have argued that exaggeration of differences that are present in RI can be motivated by distinctiveness constraints in the phonotactics if a contrast would otherwise be insufficiently distinct. The P-map hypothesis predicts that the preferred means to salvage an insufficiently distinct contrast should be to enhance cues that are already present in RI, rather than introducing completely new cues because exaggerating a difference that is already present in RI requires a smaller perceptual change.

5. Conclusions

We have seen evidence that faithfulness constraints are sensitive to language-specific phonetic details such as stop releases and partial devoicing of stops. We have argued that this effect can only be accounted for in terms of a phonetically detailed level of representation, the Realized Input, which mediates between the input and the output. That is, the input is mapped onto its expected phonetic realization, the RI, and the output is then required to be faithful to this intermediate representation. Language-specific phonetic details such as the distribution of stop releases and partial devoicing of geminate stops are represented in the RI. Since faithfulness constraints require that the output should be perceptually similar to the RI, they are necessarily sensitive to these phonetic details. It is not possible to account for the effects of phonetic detail in terms of faithfulness constraints that refer to the unrealized input because language-specific phonetic details are not systematically distributed at this level. Faithfulness constraints that refer to the presence of phonetic details in the output are also inadequate because the distribution of phonetic properties can be affected by the very processes that the faithfulness constraints are supposed to regulate.

According to this analysis, the mapping from input to output involves two steps: a mapping from input to RI (Phonetic Realization) and a mapping from RI to output (Phonotactics). This arrangement allows for opaque interactions between processes derived in phonetic realization and processes derived in the phonotactic component. We have seen opaque interactions of this kind are attested and can give rise to displaced contrasts where a cue that is derived in phonetic realization is preserved in the output, although its conditioning environment is lost.

References:

Baroni, Marco, & Laura Vanelli (2000). The relationship between vowel length and consonantal voicing in Friulian. Lori Repetti (ed.) *Phonological Theory and the Dialects of Italy*, 13-44. John Benjamins, Amsterdam.

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

Bloomfield, Leonard (1933). *Language*. London: Unwin.

Catford, J.C. (1977). *Fundamental Problems in Phonetics*. Indiana University Press, Bloomington.

Chitoran, Ioana (1998). Georgian harmonic clusters: phonetic cues to phonological representation. *Phonology* 15, 121-141.

Clark, Mary (1993). Representation of downstep in Dschang Bamileke. Norvin van der Hulst & Keith Snider (eds.) *The Phonology of Tone*. 29-73. Mouton de Gruyter, Berlin.

Clements, George N. & Kevin C. Ford (1979). Kikuyu tone shift and its synchronic consequences. *Linguistic Inquiry* 10, 179-210.

Dorman, M. F., Studdert-Kennedy, M., & Raphael, L. J. (1977). Stop-consonant recognition: Release bursts and formant transitions as functionally equivalent, context-dependent cues. *, 22*, 109-122.

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 (2006). The role of distinctiveness constraints in phonology. Ms., MIT, Cambridge.

Flemming, Edward (2008). Asymmetries between assimilation and epenthesis. Ms, MIT, Cambridge.
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.
Gafos, Adamantios (2002). A grammar of gestural coordination. *Natural Language & Linguistic Theory* 20, 269-337.
Gallagher, Gillian (2007). Pharyngeal preservation in West Greenlandic. Handout, LSA Annual Meeting 2007.
Ham, William H. (2001). *Phonetic and Phonological Aspects of Geminate Timing*. Routledge, New York.
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.
Henderson, Janette B., & Bruno H. Repp (1982). Is a stop consonant released when followed by another stop consonant? *Phonetica* 39, 71-82.
Hirata, Yukari & Jacob Whiton (2005). Effects of speaking rate on the single/geminate stop distinction in Japanese. *Journal of the Acoustical Society of America* 118, 1647-1660.
Hombert, Jean-Marie (1974). Universals of downdrift: their phonetic basis and significance for a theory of tone. *Studies in African Linguistics, Supplement* 5, 169-183.
Hyman, Larry M. & Russell G. Schuh. 1974. Universals of tone rules: Evidence from West Africa. *Linguistic Inquiry* 5, 81–115.
Itô, Junko, & Armin Mester (1986). The phonology of voicing in Japanese: Theoretical consequences for morphological accessibility. *Linguistic Inquiry* 17. 49–73.
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.
Jun, Jongho (1995). *Perceptual and Articulatory Factors in Place Assimilation: An Optimality-Theoretic Approach*. PhD dissertation, UCLA.
Jun, Jongho (2002). Positional faithfulness, sympathy, and inferred input. Ms, Yeungnam University, Daegu, Korea.
Jun, Jongho (2004). Place assimilation. Bruce Hayes, Robert Kirchner, and Donca Steriade (eds.) *The Phonetic Bases of Markedness*. Cambridge University Press, Cambridge. 58-86.
Kawahara, Shigeto (2006). A faithfulness ranking projected from a perceptibility scale: The case of [+voice] in Japanese. *Language* 82, 536-574.
Kenstowicz, Michael, Mahasen Abu-Mansour & Miklós Törkenczy (2003). Two notes on laryngeal licensing. Jonathon Kaye & Stefan Ploch (eds.) *Living on the Edge: 28 Papers in Honour of Jonathon Kaye*. Mouton de Gruyter, New York.
Klatt, Dennis H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America* 59, 1208-122l.

Kochetov, Alexei and Connie K. So (2007). Place assimilation and phonetic grounding: A cross-linguistic perceptual study. *Phonology* 24, 397-432.

Lindblom, Björn (1986). Phonetic universals in vowel systems. J.J. Ohala and J.J. Jaeger (eds.) *Experimental Phonology*. Academic Press.

Lisker, Leigh (1986). "Voicing" in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech* 29.3-11.

Lisker, Leigh, & Arthur S. Abramson (1970). The voicing dimension: some experiments in comparative phonetics. *Proceedings of the Sixth International Congress of Phonetic Sciences, Prague, 1967*. Academia, Prague.

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

Maddieson, Ian (1997). Phonetic universals. William J. Hardcastle & John Laver (eds) *The Handbook of Phonetic Sciences*. Blackwell, Oxford.

Martin, Samuel E. (1951). Korean phonemics. *Language* 27, 519–533.

McCarthy, John J., & Alan Prince (1995). Faithfulness and reduplicative identity. *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*, 294-384.

McCrary, Kristie (2004). *Reassessing the role of the syllable in Italian phonology: An experimental study of consonant cluster syllabification, definite article allomorphy and segment duration*. PhD dissertation, University of California, Los Angeles.

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.

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. (1978). Production of tone. InVictoria. A. Fromkin (ed.), *Tone: A linguistic survey*. New York: Academic Press, 5–39.

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.

Ourso, Meterwa (1989). *Lama Phonology and Morphology*. Ph. D. dissertation, University of Illinois at Urbana-Champaign.

Padgett, Jaye (1995). Partial Class Behavior and Nasal Place Assimilation. Rutgers Optimality Archive 113, http://roa.rutgers.edu/.

Peterson, Gordon, & Ilse Lehiste (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America* 32, 693-703.

Port, Robert (1981). Linguistic timing factors in combination. *Journal of the Acoustical Society of America* 69, 262-274.

Prince, Alan, & Paul Smolensky (2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Blackwell, Malden, MA.

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.

Revoile, S., .M. Pickett, Lisa D. Holden, & David Talkin (1982). Acoustic cues to final stop voicing for impaired- and normal-hearing listeners. *Journal of the Acoustical Society of America* 72, 1145-1154.
Robins, R.H., & Natalie Waterson (1952). Notes on the phonetics of the Georgian word. *Bulletin of the School of Oriental and African Studies* 14, 55-72.
Smith, Caroline L. (1997). The devoicing of /z/ in American English: effects of local and prosodic context. *Journal of Phonetics* 25, 471-500.
Snoeren, Natalie, Pierre A. Hallé & Juan Segui (2006). A voice for the voiceless: Production and perception of assimilated stops in French. *Journal of Phonetics* 34, 241-268.
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, Doncca (2001b). Directional asymmetries in place assimilation. Elizabeth Hume and Keith Johnson (eds.) *The Role of Speech Perception in Phonology*. Academic Press, New York.
Tranel, Bernard (1987). *The Sounds of French: an Introduction*. Cambridge University Press, Cambridge.
Watson, Janet (2002). *The Phonology and Morphology of Arabic*. Oxford University Press, New York.
Wilson, Colin (2001). Consonant cluster neutralization and targeted constraints. *Phonology* 18, 147-197.
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.
Xu, Y. and Sun X. (2002). Maximum speed of pitch change and how it may relate to speech. *Journal of the Acoustical Society of America* 111: 1399-1413.
Zsiga, Elizabeth C. (2000). Phonetic alignment constraints: consonant overlap and palatalization in English and Russian. *Journal of Phonetics* 28. 69-102.