This paper investigates the perceptual effects of varying coarticulation within consonant clusters which consist of labial and dorsal stops. Of central interest is the effect that overlap between labial and dorsal articulations has on the perception of place of articulation in coda position. Previous work on the perception of gestural overlap between stop consonants has focused primarily on clusters containing coronal stops; this paper therefore redresses a data gap in this literature. The results of three perception experiments are reported here, which show definitively that a high degree of coarticulation perceptually favors labials over dorsals in coda position. This result corroborates other studies indicating that labial place cues are more robust in coda position than dorsal place cues are. Some consequences of these findings for the relationship between gestural overlap and place assimilation are discussed, with particular attention to phonological systems in which labials undergo place assimilation, such as Seoul Korean.

1 Introduction

A seminal observation guiding inquiry into the relationship between phonetics and phonology is that, to the extent that they can even be reliably distinguished, the two domains of description often resemble each other quite closely. In the phonological domain, the process displaying perhaps the most evident similarity to phonetic patterning is that of consonantal place assimilation. The triggers, targets, and largely regressive directionality of categorical place assimilation processes have all been argued (by e.g. Ohala 1990, Jun 1995, Steriade 2001) to parallel the types of low-level misperceptions and articulatory perturbations found in online language use.

One view that has gained ground over the last two decades is that no account of the phonetic grounding of place assimilation is complete without reference to coarticulatory overlap between consonants.¹ Gestural overlap variation is intrinsic

¹ Many thanks go to Jaye Padgett, Grant McGuire, Abby Kaplan, and Dan Silverman for their generous donation of time and expertise in answering questions, designing experiments, and coding scripts. I am also grateful to Dani Byrd, Louis Goldstein, and Ben Parrell of USC for help with TADA, and to Dan Jurafsky and Yun-Hsuan Sung of Stanford for help with MATLAB. All errors remain my own.

¹ Specifically, the overlap we are interested in is typically defined for a two-consonant (C₁C₂) cluster as the portion of C₁’s total gestural movement produced after the onset of C₂’s gestural movement.
to speech production (Nolan 1992, Byrd 1996, Byrd & Tan 1996, Ellis & Hardcastle 2002) and can induce perceptual masking of speech gestures, resulting in cases where a gesture may not be audible or recoverable despite being articulatorily present (as with the [t] in Browman & Goldstein’s (1990) famous example ‘perfect memory’). Misperception asymmetries in this gestural masking have also been argued to pattern with asymmetries in place assimilation systems (Byrd 1992, Surprenant & Goldstein 1998). Thus, misperception due to overlap represents a real-world phonetic difficulty that appears to parallel the form of phonological systems. Depending on one’s theory of phonetic grounding in phonology, such difficulty can either constrain synchronic grammars (e.g. Lindblom 1990, Flemming 1995, Boersma 1998, Hayes & Steriade 2004, Steriade 2009) or shape the direction of sound change (e.g. Ohala 1981, Blevins 2004, Silverman 2006).

The apparent completeness of the affinity between overlap-induced misperception and place assimilation has engendered a pair of reinforcing strands of research in phonological theory. The first is empirical, concerned with experimentally testing just how deep the affinity runs, and is based in the techniques and outlook that have come to be known under the name “laboratory phonology”. The second is theoretical, and takes up the question of how to properly interpret the nature of the affinity and the fact of its existence. Cases where phonetic and phonological patterns don’t align with one another are particularly informative for accounts of how such alignments arise, since theories differ in how easily they accommodate exceptional behavior.

This paper investigates the empirical basis of one case of apparent phonetic misalignment in a phonological system. Seoul Korean exhibits an asymmetrical place assimilation system, shown in (1), in which (unreleased) labial codas optionally assimilate to following dorsals, but (unreleased) dorsal codas do not assimilate to following labials ((1b) vs. (1c)) (Kim-Renaud 1991, Jun 1995).\(^2\)

(1) Korean place assimilation processes

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Surface forms</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. /mit + ko/</td>
<td>[mikk’o] or [mitk’o]</td>
<td>‘believe and’</td>
</tr>
<tr>
<td>/mit(h) + pota/</td>
<td>[mipp’ota] or [mitp’ota]</td>
<td>‘more than the bottom’</td>
</tr>
<tr>
<td>b. /ip + ko/</td>
<td>[ikk’o] or [ipk’o]</td>
<td>‘wear and’</td>
</tr>
<tr>
<td>/ip + ta/</td>
<td>[ipt’a] (*[itt’a])</td>
<td>‘wear + SENTENCE ENENDER’</td>
</tr>
<tr>
<td>c. /nok + ta/</td>
<td>[nokt’a] (*[nott’a])</td>
<td>‘melt’</td>
</tr>
<tr>
<td>/kuk + pota/</td>
<td>[kukp’ota] (*[kupp’ota])</td>
<td>‘more than soup’</td>
</tr>
</tbody>
</table>

Recent research on the perception of place of articulation (POA) in unreleased coda consonants indicates that labial codas do not have a greater propensity to be misperceived than dorsal codas, making this process appear phonetically unnatural (Kochetov & So 2007; contra Jun 1995, 2004). Other authors, however,

\(^2\)Coronal codas assimilate to both labial and dorsal onsets, but we are not concerned with them here. Coronal onsets never trigger place assimilation in preceding labial or dorsal codas. All stop clusters and geminates are also subject to a process of post-obstruent tensification, shown in surface forms with [‘].
argue that Korean’s place assimilation system appears phonetically natural once Korean-specific articulatory settings for gestural overlap are taken into account (Son et al. 2007, Son 2008). Gestural overlap is thus at the center of the debate over labial assimilation in Korean, both for language-specific reasons and because of the general affinity between overlap-induced misperception and assimilation mentioned above. The perceptual consequences of gestural overlap between labial and dorsal stop consonants are not clearly understood, though, due to a lack of coverage in the phonetics literature. In this paper, I present three perceptual experiments designed to mend this data gap by gauging how fine-grained changes in overlap affect the perception of coda POA in labial-dorsal and dorsal-labial stop clusters. The results, I argue, lend further support to the conclusion that Korean is synchronically unnatural with regard to how its gestural overlap settings respond to perceptual difficulty.

Investigation into apparently unnatural assimilation systems stands to advance our understanding of how naturalness considerations come to constrain place assimilation in the first place. In the context of the wider debate between synchronic and diachronic construals of markedness, unnatural systems represent a sort of functional “opacity” which is hard for theories of strictly synchronic phonetic grounding (e.g., the Phonetically Based Phonology of Hayes & Steriade 2004) to accommodate. In such theories, the drive for optimality is codified as a restriction on the form of learnable grammars (and/or on types of valid sound changes), and therefore every reduction of fitness along some dimension must be counterbalanced by an increase of fitness along another. As I argue below, Korean presents us with a case in which sound changes led only to decreases in overall perceptual fitness, with no concomitant increases elsewhere in the system. The resultant functional opacity is due to the “non-adaptive” nature of the proposed sound changes, which can only be comfortably accommodated within theories that do not constrain sound change (or synchronic grammars) to be strictly optimizing, such as Evolutionary Phonology (e.g., Ohala 1981, Blevins 2004). I submit that the Korean overlap/assimilation system can be best understood within such emergentist phonological frameworks, and elaborate my proposal using insights from evolutionary exemplar theory (Pierrehumbert 2001, Hume & Johnson 2001, Blevins 2004, Wedel 2004, 2007, Silverman 2006, Bybee 2007). The upshot is a theory that still predictably traces place assimilation to phonetic causation, while tolerating sometimes non-optimizing outcomes of language change.

The roadmap of the paper is as follows. Before presenting the experiment designs and results in §§2, 3 and 4, I finish out this section with a discussion of previous work bearing both generally and specifically on the present topic. The theoretical consequences of the experimental results outlined in the preceding paragraph are discussed in §5, and §6 concludes.

---

3 Some exceptions are discussed below in §1.3.
1.1 Perception of coda consonant place of articulation

In this section we review how coda POA is perceived in clusters of stop consonants, in order to determine how it is dependent on and can be influenced by gestural overlap.

1.1.1 Primary cues

Perception of POA in coda stop consonants depends primarily on three cues: (i) VC formant transitions, (ii) coda release burst spectra, and (iii) closure duration in clusters (cf. e.g. Wright 2004). Coda POA information can also be transmitted in following onset cues such as CV formant transitions and aspiration (Surprenant & Goldstein 1998), but for our purposes these effects are secondary. Onset cues are in general much more robust than coda cues (Ohala 1990), due to an asymmetrically strong response in the auditory nerve to acoustic changes immediately after silence or low amplitude (Wright 1996). Coda perception may also be hampered by the lack of one or more of the general cues listed above: unreleased codas have no burst spectra, and word- or phrase-final codas are unaided by closure duration information.

The most important “direct” cues to coda POA are VC formants and coda release bursts (if present). The relative importance of each of these cues depends heavily on POA, as summarized in table 1.1. Perceptual studies have shown that formant transitions are most robust for labials, slightly less so for coronals, and least robust for dorsals, while release bursts are by far most robust for dorsals, less so for coronals, and least for labials (Wright 2001, Winters 2001, Kochetov & So 2007). 4

<table>
<thead>
<tr>
<th>VC formant transitions:</th>
<th>Labials &gt; Coronals &gt; Dorsals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coda burst spectra:</td>
<td>Dorsals &gt; Coronals &gt; Labials</td>
</tr>
</tbody>
</table>

The same studies found that the most robust released codas are dorsals, indicating that release bursts have an overall greater perceptual impact than formant transitions do. Cue-mixing studies that present listeners with mismatched VC formants and bursts support the same conclusion, as subjects reliably hear codas with POA matching that of the bursts, but not that of the formant transitions (Malécot 1958, Repp 1978).

Closure duration in stop clusters is somewhat distinct from bursts and formant transitions in that it is not a direct cue to any particular coda POA, but rather serves as a cue to whether a singleton stop is responsible for the closure. Since singleton stops are by their nature “monorganic”, the perception of a heterorganic cluster as a singleton stop amounts to assimilatory misperception. Several studies

4 These perception results are valid independently of gestural overlap between codas and onsets; the studies that derived them used non-overlapped codas in their testing materials.
have found that short closures can override diverging heterorganic VC/CV
formant transitions, leading only C₂ in a C₁:C₂ cluster to be perceived (Repp 1978,
Dorman et al. 1979). Thus, closure duration is a kind of “meta-cue” that informs a
listener how to weight the transition information into and out of a consonant
closure (cf. Gow & Im 2004).

1.1.2 Overlap dependencies among coda cues
How are the above coda POA cues affected by changes in gestural overlap? In
general, greater gestural overlap leads to loss of release bursts, adulteration of VC
formant transitions with onset POA information, and shorter closure durations, all
of which can adversely affect the recovery of coda POA (Byrd 1992, Zsiga 1994,
Zsiga 2000, Chitoran et al. 2002, Kochetov et al. 2007). While VC formant
transitions and release bursts are both directly dependent on gestural overlap for
either production, closure durations can be controlled through independent
gestural planning (Jun 1995:134-135). That is to say, the closure in a C₁C₂ cluster
can be made arbitrarily long regardless of high overlap between C₁ and C₂. Barry
(1991), for instance, has found evidence for such lengthening in overlapped
clusters in English casual speech, presumably as a means of preserving closure
length in order to cue the presence of a consonant cluster. On the other hand,
Zsiga (2000) found that greater gestural overlap is correlated with shorter closure
durations in the comparison between Russian and English consonant clusters; Son
(2008) found the same for Korean. This dichotomy suggests that individual
speakers can indeed control closure duration independently of overlap, but that
closure duration and overlap may evolve together diachronically.

Of the three cues discussed here, coda release bursts are the most sensitive to
changes in overlap. Release bursts depend on the venting of a pressurized oral
cavity to the atmosphere. The presence of a full coda burst in a stop cluster
requires that the release of the coda stop precede the achievement of closure for the
onset stop. This necessitates a very low level of overlap, so that the entire oral tract
can be vented and a loud burst be produced. However, as pointed out by Byrd
(1996), audible bursts can also be produced from the venting of a smaller anterior
cavity between two highly overlapped stop closures. In particular, if a coda stop is
articulated closer to the front of the mouth than the following onset stop (in so-
called front-to-back order), then the release of the coda closure will vent the oral
cavity between the two stops to the atmosphere. If the cavity is under great
enough pressure, this venting will be audible in front-to-back clusters, even at high
levels of overlap. The opposite place ordering of coda and onset closures, namely
back-to-front order, is incapable of producing an audible burst under high overlap,
since the more posterior coda release can only equalize pressure within the oral
tract. This asymmetry between the presence vs. absence of coda release bursts
under high overlap in front-to-back vs. back-to-front sequences has been
experimentally verified by Zsiga (2000) and Chitoran et al. (2002). One predictable
effect of this ordering asymmetry is to facilitate the perception of codas in
overlapped front-to-back [pt], [pk] and [tk] clusters, and to inhibit coda perception
in overlapped back-to-front [tp], [kp] and [kt] clusters. Combined with the VC
formant transition strengths listed in table 1.1, this finding suggests that labial-stop
codas may be at a significant perceptual advantage over dorsal-stop codas when overlap levels are high.

Several studies have looked into the effects of gestural overlap on VC formant transitions. In English, Repp (1983) and Zsiga (1994) found evidence that VC formant transitions contained POA information from onset consonants even in careful speech. Byrd (1992) and Surprenant & Goldstein (1998) determined that coronal VC formant transitions undergo significant perceptual changes when overlap with a following labial consonant is increased, but that labial formant transitions are relatively stable under all but the highest degrees of overlap with a following coronal. Similarly, Zsiga & Byrd (1990) showed that coronal formant transitions are significantly influenced by even modest amounts of overlap with a following dorsal. Only two studies I have seen (Jun 1996, Son et al. 2007) explicitly investigate how labial codas respond to overlap with a following dorsal, but neither study isolates how VC formants are affected; they are discussed below in §1.3.1. No study that I am aware of systematically investigates the effect of gestural overlap on dorsal VC formant transitions, although Byrd (1992:19) speculates that overlapped dorsals are as perceptually stable in coda position as overlapped labials are. Her explanation (ibid.) for the particular malleability of coronal formant transitions, following an argument of Browman & Goldstein’s (1990), is that coronals are intrinsically swifter, and therefore shorter and more easily overlapped, than labials and dorsals due to the tongue tip’s greater stiffness and low mass (cf. Kuehn & Moll 1976). This account also implies that clusters with overlapped coronal codas will tend to have more singleton-like closure durations, independently increasing the likelihood of coda misperception. Note that Byrd’s explanation only applies to coronal/non-coronal asymmetries in overlap masking though, since labials and dorsals have similar, relatively sluggish gestural durations.

Table 1.2 summarizes the above research on coda cue vulnerability to overlap-induced misperception.

<table>
<thead>
<tr>
<th></th>
<th>Closure durations:</th>
<th>VC formant transitions:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-coronals &gt;</td>
<td>Non-coronals &gt;</td>
</tr>
<tr>
<td>Production of release</td>
<td>Coronals</td>
<td>Coronals</td>
</tr>
<tr>
<td>bursts:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Labials &gt;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Coronal</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dorsals</td>
<td></td>
</tr>
</tbody>
</table>

1.2 Gestural overlap, misperception and place assimilation

In this section, we review two basic arguments supporting the notion that gestural overlap and place assimilation are causally linked in linguistic systems. The first is

---

5 The swiftness of coronal gestures is hypothesized by Byrd to be responsible not only for their vulnerability as codas to perceptual masking, but also for their inability as onsets to greatly influence the VC formant transitions of overlapped labial and dorsal codas.
that, cross-linguistically, the degree of gestural overlap in consonant clusters appears to be correlated with the prevalence of place assimilation in those clusters. The second is that place assimilation systems appear to pattern with overlap-induced POA misperception in codas, but not with general misperception of POA in codas.

1.2.1 Overlap and assimilation in linguistic systems
In finding explanations for the form of place assimilation systems, it is first necessary to note that not all assimilation systems are categorically neutralizing. Much evidence has been collected showing that assimilation processes can be optional in their application and gradient in their spatial extent. Coronal place assimilation processes in English and Korean, for example, are two cases where instances of categorically-neutralizing and gradient assimilation coexist and apply stochastically (Barry 1985, Browman & Goldstein 1990, Nolan 1992, Ellis & Hardcastle 2002, Kochetov & Pouplier 2008). In gradient assimilation processes, increased overlap plays a directly measurable role in the perceptual masking of speech gestures, alongside other casual-speech effects like gestural reduction (Surprenant & Goldstein 1998). What’s more, the coexistence of categorical and gradient behavior in some processes suggests that fully neutralizing assimilation is not fundamentally different from gradient assimilation, but rather is simply at one end on a spectrum of variation.

Just as place assimilation is not a set of purely “phonological”, categorical processes, so too are gestural overlap levels not purely “phonetic”, stochastic facts, but rather vary systematically both within and across languages. Within individual languages, gestural overlap between consonants has been found to have mean settings that are consistently different from cluster to cluster (Byrd 1996, Wright 1996, Surprenant & Goldstein 1998, Zsiga 2000, Chitoran et al. 2002, Son 2008). Languages can also differ from one another in having distinct systemic gestural overlap settings (Silverman 1995). This can perhaps most easily be seen from the fact that coda consonants are consistently released in some languages but not released in others (Zsiga 2000, Kochetov et al. 2007), since coda releasedness depends on gestural overlap settings in clusters. Taken together, these observations support the increasingly well-accepted conclusion that linguistic competence encompasses not just categorical, contrastive knowledge, but also non-contrastive variation and detail.

Addressing these two facets of linguistic variation, recent studies have found evidence both within and across languages that the degree of mean gestural overlap in a consonant cluster is correlated with that cluster’s proneness to undergo place assimilation. Within individual languages, studies have shown such a correlation for assimilating clusters with coronal codas in English (Byrd 1996, Surprenant & Goldstein 1998) and labial codas in Korean (Son et al. 2007, Son 2008). Comparing between languages, Kochetov et al. (2007) note that languages

---

\(^6\) Note that such correlations are only meaningful for non-categorical assimilation. The output of fully neutralizing assimilation, namely a geminate, does not involve overlap between separate gestures.
with consistently released codas, like Russian, tend to have fewer place assimilation rules than languages with more frequently unreleased codas, like English and Korean. Using EMMA articulatory data, they verify that clusters in Russian are also far less overlapped than those in Korean. Unfortunately, the present scope of such cross-linguistic studies is limited by the lack of phonetic data on fine-grained articulatory features like gestural overlap. The upshot of this research is to suggest that diachronic increases in gestural overlap in clusters (whatever the source may be) are associated with the inception of assimilation processes targeting those clusters.

1.2.2 Overlap-induced misperception and the targets of place assimilation

Evolutionary theories of language structure that posit diachronic sources for phonological patterns make frequent mention of the guiding role that misperception plays in language change (cf. e.g. Ohala 1981, Blevins 2004, Silverman 2006). Given the typological generalization made in the preceding subsection, namely that languages with unreleased codas are more prone to host assimilation processes, we can compare two competing hypotheses for the diachronic origins of place assimilation: (i) that misperception of unreleased codas due to high overlap is to blame, as argued here; or (ii) that misperception of unreleased codas in general (regardless of overlap level) is to blame. One way of teasing apart these two hypotheses is to compare asymmetries in the two types of misperception. The basic logic is that, if place assimilation is caused by the diachronic “amplification” of a phonetic process like misperception, then any asymmetries present in the phonetic process should carry over to become asymmetries in place assimilation systems.

Place assimilation systems have long been recognized to asymmetrically target coronal codas, and to generally target labial and dorsal codas only if coronals are also targeted (Paradis & Prunet 1991, Mohanan 1993, Jun 1995, 2004, de Lacy 2002). Do either of the two types of misperception considered above reflect this asymmetry? That is, are coronal codas misperceived significantly more often than non-coronal codas, either when overlapped and unreleased or when unreleased in general? Consulting table 1.1, we see that among unreleased codas, whose main POA cue is VC formant transitions, dorsals are generally misperceived far more often than coronals and labials when overlap is not a factor. However, table 1.2 shows that under increased gestural overlap, unreleased coronals’ VC formant transitions succumb to misperception far more readily than those of non-coronals do. Cluster closure durations also come to resemble those of singleton stops under lower levels of overlap with coronal codas than with non-coronal codas.

---

7 This cross-linguistic pattern could also simply mean that presence vs. absence of bursts is the crucial element in determining how prevalent place assimilation is in a language. This possibility is discussed and argued against in the next subsection.


9 Even if front-to-back cluster release bursts are also considered as a potential POA cue for low-overlap “unreleased” codas, dorsal codas are still at a disadvantage relative to coronals.
Thus, the surprising fact of coda POA perception from table 1.1 is that coronals are generally not the least robust POA in coda position (whether released or not), despite being the most prone to place assimilation cross-linguistically (see Winters 2001 and Kochetov & So 2007 for similar statements of surprise). The codas that are rarest as targets of place assimilation are those that stand up comparatively well to high overlap, like dorsals and labials, rather than those with generally intelligible VC formant transitions, like coronals and labials. This pattern supports the hypothesis that the diachronic precursor to place assimilation systems is perceptual masking induced by gestural overlap variation in language usage (cf. Browman & Goldstein 1990, Pierrehumbert 2001, Silverman 2006, Bybee 2007).

1.2.3 Avenues of sound change
From a sound change perspective, place assimilation’s origin in gestural overlap is unsurprising, given that variations in overlap are an inevitable part of speech implementation. Any language with a moderate mean level of gestural overlap in consonant clusters will therefore present a consonant cluster “ecology” to learners in which some clusters sound perceptually assimilated. All it takes for a phonological process to arise is for some of these learners to reinterpret the perceptual assimilation as articulatory assimilation (Browman & Goldstein 1990, Ohala 1990). One example of how this reinterpretation can happen is proposed by Chen (2003), based on a computational model that estimates vocal tract aperture settings from the acoustics of overlapped consonant clusters. As the amount of overlap masking increased, her artificial articulatory learner tended to derive aperture settings that were most consistent with gesturally reduced codas. Coronals experienced the largest effects. She concludes that high overlap sounds like gestural reduction, especially in coronals.

Another speculative possibility is that, as overlap masking and partial reductions become more common in a language, new learners come to devote fewer cognitive perceptual resources to discriminating the degraded coda POA cues, making them less sensitive (and more tolerant) to small assimilatory changes in codas. Kochetov & So (2005) propose such an effect to explain their finding that Korean speakers make less use of coda release bursts as a POA cue than Russian and English speakers do. The logic is that learners attend to cues proportionally to how useful the cues are at distinguishing native language categories, and Koreans have little use for release bursts since their codas are largely unreleased.

The upshot for both proposals is that several plausible diachronic avenues for articulatory reinterpretation open up once a modest degree of mean gestural overlap gets a toehold in a language’s clusters, leading ultimately to the overlap/assimilation patterns discussed in this section.

---

10 A similar surprising fact was found by Winters (2003) for POA perception in nasals. POA was more accurately identified in nasal codas than in plosive codas, even though nasals are more prone to undergo place assimilation than plosives cross-linguistically (Mohanan 1993).

1.3 Labial assimilation, labial/dorsal overlap and the present experiments

In the last section we saw that major aspects of the shape of place assimilation systems can be explained based on how VC formant transitions (and possibly closure durations) respond to increases in gestural overlap. While this pattern holds well for the majority of place assimilation systems, namely those involving coronal codas, the lack of data on how overlap affects non-coronal codas leaves us unsure of whether the explanation generalizes to cases of non-coronal place assimilation. One case of non-coronal place assimilation in particular that has raised the interest of laboratory phonologists in recent years is labial-to-dorsal assimilation in Korean. Based on the cues to coda POA that we have reviewed so far, there is a prima facie case to be made for labial-to-dorsal assimilation being phonetically unnatural. We examine the evidence a bit more closely now, and outline the goals pursued by the experiments presented in this paper.

1.3.1 Labial place assimilation in Korean

In Korean, labial stops (both plosives and nasals) in coda position optionally assimilate to the place of following dorsal stops, but dorsal coda stops do not assimilate to following labials.12 The fact that labials undergo assimilation to dorsals but do not themselves trigger it in dorsals has led many to speculate, following Jun (1995), that dorsals exert an overall greater influence on POA cues than labials do. Judging from the discussion in §1.1 and from table 1.1, there is strong evidence contradicting this speculation. All codas are reported to be uniformly unreleased13 in Korean (Kim-Renaud 1991), so dorsal codas do not have the advantage of their strong release bursts. By rights, labial codas’ strong formant transitions should decide the issue in their favor. The fact that the apparently perceptually superior labials are in fact the targets of assimilation represents a significant mismatch between phonetic and phonological patterning (cf. Kochetov & So 2007 for the same conclusion).

The discussion in §1.2 points to the effects of gestural overlap as a general explanatory source of assimilatory misperception. But an examination of table 1.2 gives reason to question the viability of this explanation, too. While it is unclear how labial and dorsal VC formant transitions are affected by increased overlap, labial codas have an advantage in that they are (at least in principle) able to preserve a release burst under high overlap, due to necessary front-to-back cluster ordering. This observation should leave us wondering what perceptual advantage, if any, dorsal codas enjoy in Korean. What do Korean codas sound like?

12 Son et al. (2007) and Kochetov & Pouplier (2008) refute Jun’s (1996) conclusion that labial assimilation involves spatially gradient gestural reduction in Korean. In Jun’s data, lip aperture is not measured directly, and Son et al.’s EMMA data show evidence only for categorical replacement of labials. NB: this result also refutes the claim in Jun’s title, that “place assimilation is not the result of gestural overlap.”

13 Kim & Jongman (1996) and Son (2008:80) have found some evidence of coda releases in their Korean elicitation, although other studies show codas to indeed be unreleased (Cho & McQueen 2008).
Luckily enough, labial assimilation in Korean only applies stochastically, allowing unassimilated labial codas to be directly observed, and their properties compared to those of the non-assimilating dorsal codas. EMMA investigations show that labial codas in labial-dorsal (henceforth, PK) clusters exhibit far greater levels of mean gestural overlap than dorsal codas in dorsal-labial (henceforth, KP) clusters do (Son et al. 2007, Son 2008). The investigators argue that the substantially higher level of overlap among labial codas is immediately responsible for their being prone to assimilation, despite the usually high fidelity of labial POA cues. Indeed, Jun (1996) and Son et al. (2007) both find perceptual evidence that codas in Korean PK clusters are less identifiable than “pure” labial codas. Nonetheless, these two papers are the only works I have found that explicitly address the question of how labial codas are perceptually altered by overlap with a dorsal.

1.3.2 Open questions
Without wishing to understate the significance of Jun’s and Son’s (et al.) studies, I believe that they leave several important issues unaddressed. The experiments that I report in this paper are designed to address these concerns head-on, which are as follows:

- The perception of overlapped labial PK codas is not directly compared with that of dorsal KP codas, overlapped or otherwise. Given the overall low fidelity of dorsal formant transitions, it is not clear a priori that low-overlap dorsal codas should be any more robust than high-overlap labial codas.
- It has not been determined whether the perceptual degradation of overlapped PK codas is due to changes in VC formant transitions or to reduced closure intervals.
- A systematic relationship between overlap and labial coda misperception in PK clusters has not been established. In previous work the levels of overlap tested were few, and were collected somewhat at random from natural speech tokens.
  No previous work has systematically examined how overlap affects KP codas.

The first two of the above issues concern Son et al.’s claim that overlap differences in PK and KP clusters are sufficient for explaining the synchronic naturalness of Korean’s labial place assimilation system. Their articulatory data support the notion that high gestural overlap is a precursor to place assimilation, but the data do not seem to make Korean’s place assimilation system fit any better into our phonetically predicted typology of assimilation. If strong VC formant transitions and the possibility for maintaining release bursts under high overlap favor the perception of labial codas over dorsals, then labial-to-dorsal assimilation is still a counterexample to the typological generalization made in §1.2.2. However, we in fact do not know how labials and dorsals respond perceptually to gestural overlap; no studies have yet compared the two. It is still possible that dorsal KP codas are better identifiable than labial PK codas, even when overlap is high.

14 But interestingly, Jun (1996) also found that some highly overlapped labial codas were well identifiable even when the following dorsal onset’s closure preceded that of the labial.
codas maintain their POA character better under overlap than labial PK codas do. For instance, it may be that overlap causes VC formant transitions to undergo more extreme changes in labials than in dorsals. This possibility is meaningful enough to be codified as a hypothesis:

**Hypothesis H1.1**
Under increasing gestural overlap, VC formant transitions undergo greater assimilatory perceptual changes in PK clusters than in KP clusters.

If hypothesis H1.1 were verified, then Korean labial place assimilation would no longer look quite so anomalous. It would in fact fit the typological pattern that place assimilation tends to target clusters that are more drastically misperceived under gestural overlap.

The third issue above is somewhat more concerned with how the Korean cluster overlap/assimilation system diachronically came to be the way it is – that is, which avenues of sound change led it to its present state. This question is important given the “convenient” nature of the Korean overlap system: if language-specific overlap settings are responsible for assimilation, then what explains the overlap settings? Why, for instance, are dorsal codas (but not labials) so non-overlapped? We revisit this issue in §5. I submit that we can only start to answer this question by examining how small, fine-grained changes in overlap are perceived. If language change is gradual and driven by variation and misperception, as is commonly assumed in theories of self-organizing or evolutionary phonology, then we must learn how gradual variations map onto perceptual changes.

1.3.3 Experimental goals
Three experiments were run to investigate the issues outlined in the last subsection. Experiments I (in §2) and II (in §3) are based closely on those of Byrd (1992), who examined the perception of coronal and labial stop clusters under overlap. Their goal is to confirm or disconfirm hypothesis H1.1 – that is, to determine differences in how PK and KP clusters respond to high overlap, and to isolate the effects on VC formant transitions. Experiment III (in §4) is designed to further verify the findings of experiments I and II by factoring out possible subject categorization biases, and also to test how perceptible small changes in overlap are in each cluster. Such tests bear on the question of how the perceptual results of gradual variation affect sound change.

2 Experiment I
This experiment, like Byrd’s (1992), ascertains place-specific asymmetries in the perception of coda stops by way of a forced-choice identification task. Both use artificial stimuli in which gestural overlap between stop consonants has been manipulated; for our purposes, the stops are labials and dorsals. Artificial stimuli are used instead of natural stimuli in order to allow for fine control of gestural
overlap independently of other articulatory variables linked to speech rate, such as gestural reduction. Artificial stimuli also have the advantage of near-complete uniformity, so that acoustic differences between tokens are limited primarily to POA cue differences.

2.1 Stimuli

2.1.1 Method of production

The speech stimuli used in the experiments reported in this paper were modeled using Haskins Labs’ TADA (TAsk Dynamic Application; Nam et al. 2004) articulatory simulator. The articulatory models produced by TADA were fed through the Sensimetrics HLSyn speech synthesizer (Bickley et al. 1997, Hanson et al. 1997) to create high quality, realistic utterances complete with air-stream features such as bursts and aspiration. Four different single-word [VC₁:C₂V] utterance types were generated, with a medial consonant cluster consisting of a /p/ or /k/ in coda position and a /p/ or /k/ in onset position, and flanked by the vowel /a/. The four experimental word-types thus featured both heterorganic and homorganic sequences, and can be broadly transcribed as /ap.ka/, /ak.pa/, /ap.pa/, and /ak.ka/. These word-types were chosen to be non-lexical items and to minimize the effects of vowel quality on consonant perception. Henceforth, the consonant sequences associated with these word types are referred to as PK, KP, PP, and KK sequences respectively.

Within each sequence type, 11 different levels of CC-overlap were modeled. In the TADA articulatory simulator, gestural activation timing is controlled in 10 ms “frames”, thus setting 10 ms as the smallest programmable temporal resolution. The 11 overlap settings were chosen based on this resolution, in the following manner. TADA’s default setting for C₁–C₂ phasing marked the lowest level of overlap, and each subsequent level of overlap was 1 frame (10 ms) greater than the next lowest one. At the lowest (default) setting, C₁’s activation gesture overlaps that of C₂ by 3 frames (30 ms). At each higher level of overlap, C₂’s activation gesture is shifted 1 frame earlier, while C₁’s timing, as well as that of the vowels, remains unchanged throughout. Overall, the 11 gestural settings range between 3 and 13 frames of overlap for the coda and onset activation gestures. For the two lowest levels of overlap, i.e. 3 and 4 frames, enough of a delay between coda and onset target attainment exists to permit a coda release burst in the heterorganic

---

15 Many thanks to the USC phonetics laboratory group, especially Dani Byrd, Louis Goldstein, and Ben Parrell, for making the relevant software, equipment, and expertise available to me to produce the experimental stimuli.

16 Syllable affiliations are specified in the input of the articulatory model, and so are listed within slash brackets here.

17 The TADA mass-spring model decomposes articulatory motions into two parts: an onset gesture and an offset gesture. For reasons of disambiguation, I will refer to these here as “activation” and “deactivation” gestures. Note that “active” gestural deactivation is necessary to capture the quick return to spatially neutral positions characteristic of actual articulations. Without deactivation, an articulator returns to neutral much more slowly, its motion depending on the ambient level of spring tension elsewhere in the system.
clusters. At 5 frames of overlap, this burst disappears for both /ap.ka/ and /ak.pa/, and microbursts are not modeled in these simulations — therefore, closures are completely silent at 5 frames and above. By 12 frames of overlap, C₁ and C₂ are activated simultaneously, and finally, at 13 frames, the activation gesture of C₂ precedes that of C₁ by 1 frame. Two simplified sample gestural scores illustrating this manipulation of overlap are shown below in figure 2.1. While the timing of C₁’s gestures (including both oral constriction and glottal abduction) is varied across the 11 overlap degrees, the overall timing of C₁ and the vowels remains constant. This choice was made in order to decrease the number of articulatory parameters being directly manipulated.

**Figure 2.1.** Condensed gestural scores for a KP sequence under low and high overlap.

<table>
<thead>
<tr>
<th>Activation gestures</th>
<th>/ak.pa/, low overlap (~3 frames)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lips</td>
<td></td>
</tr>
<tr>
<td>Constriction</td>
<td></td>
</tr>
<tr>
<td>Tongue body</td>
<td></td>
</tr>
<tr>
<td>Constriction</td>
<td></td>
</tr>
<tr>
<td>Vowel</td>
<td></td>
</tr>
<tr>
<td>/a/</td>
<td></td>
</tr>
<tr>
<td>/a/</td>
<td></td>
</tr>
<tr>
<td>Glottis</td>
<td></td>
</tr>
<tr>
<td>Abduction</td>
<td>[-voice] [-voice]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Activation gestures</th>
<th>/ak.pa/, high overlap (~12 frames)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lips</td>
<td></td>
</tr>
<tr>
<td>Constriction</td>
<td></td>
</tr>
<tr>
<td>Tongue body</td>
<td></td>
</tr>
<tr>
<td>Constriction</td>
<td></td>
</tr>
<tr>
<td>Vowel</td>
<td></td>
</tr>
<tr>
<td>/a/</td>
<td></td>
</tr>
<tr>
<td>/a/</td>
<td></td>
</tr>
<tr>
<td>Glottis</td>
<td></td>
</tr>
<tr>
<td>Abduction</td>
<td>[-voice]</td>
</tr>
</tbody>
</table>

---

18 Naturally, no coda release burst is present in any of the homorganic sequences; these are produced as geminates.

19 To minimize clutter, no deactivation gestures have been indicated in figure 2.1. The gestures for /p/ are shown as longer than those of /k/ because TADA implements onsets with longer activation gestures than codas.
2.1.2 Acoustic characterization

The onset of modal voicing for the initial vowel is uniformly set at 80 ms after the beginning of each articulatory simulation. The resulting stimuli thus contain 80 ms of pre-utterance silence, and all are between 1.01 and 1.02 seconds long. The CC closure intervals for all clusters with non-released codas are silent; closure intervals for CC sequences with released codas are silent except for the release burst. For nearly all of the stimuli, then, identification of coda POA depends primarily on information carried by VC formant transitions and consonantal closure durations.

As expected, stop closure durations become shorter as gestural overlap is increased; closure times are charted in Table 2.1.

<table>
<thead>
<tr>
<th>Sequence type</th>
<th>APPA</th>
<th>AKKA</th>
<th>APKA</th>
<th>AKPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>151</td>
<td>170</td>
<td>171</td>
<td>169</td>
</tr>
<tr>
<td>4</td>
<td>146</td>
<td>167</td>
<td>160</td>
<td>159</td>
</tr>
<tr>
<td>5</td>
<td>134</td>
<td>159</td>
<td>151</td>
<td>144</td>
</tr>
<tr>
<td>6</td>
<td>127</td>
<td>149</td>
<td>140</td>
<td>135</td>
</tr>
<tr>
<td>7</td>
<td>115</td>
<td>140</td>
<td>131</td>
<td>123</td>
</tr>
<tr>
<td>8</td>
<td>105</td>
<td>130</td>
<td>120</td>
<td>114</td>
</tr>
<tr>
<td>9</td>
<td>95</td>
<td>119</td>
<td>110</td>
<td>106</td>
</tr>
<tr>
<td>10</td>
<td>86</td>
<td>110</td>
<td>102</td>
<td>94</td>
</tr>
<tr>
<td>11</td>
<td>78</td>
<td>101</td>
<td>97</td>
<td>85</td>
</tr>
<tr>
<td>12</td>
<td>72</td>
<td>93</td>
<td>92</td>
<td>78</td>
</tr>
<tr>
<td>13</td>
<td>67</td>
<td>90</td>
<td>90</td>
<td>76</td>
</tr>
</tbody>
</table>

It bears pointing out that the programmed overlap levels only represent the inputs to the physical model – the actual overlapping motion of articulators depends on the inertial state of each articulator within the overall system. Thus, two articulations that are planned simultaneously do not necessarily achieve their targets simultaneously. Consonant production also varies by syllable position and POA; we can see from Table 2.1, for instance, that closure durations in the output depend on both of these factors. We can discern the underlying lengths of closure gestures by examining low-overlap PK and KP tokens (<5 frames), where a coda release burst divides the consonantal closure into its two component parts. These values are averaged and shown in Table 2.2. Closure durations for onsets are longer than those for codas, and labials have consistently shorter closure durations than dorsals, modulo syllable position.
Table 2.2. Closure durations for labials vs. dorsals in coda and onset positions.

<table>
<thead>
<tr>
<th></th>
<th>Labial (coda)</th>
<th>Dorsal (coda)</th>
<th>Labial (onset)</th>
<th>Dorsal (onset)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>49 ms</td>
<td>70 ms</td>
<td>76 ms</td>
<td>101 ms</td>
</tr>
</tbody>
</table>

To give an impression of the effects of overlap on VC formant transition shape, each token's F1, F2 and F3 values were measured in Praat (Boersma & Weenink 2009) throughout a 38 ms window immediately prior to C₁ closure silence. Figures 2.2 and 2.3 show these formant shapes for each of the heterorganic cluster tokens, along with a sample of the homorganic clusters’ VC formants for reference. As can be expected, the formant transitions for the homorganic cluster tokens do not vary noticeably with overlap, and so are not displayed individually.

Figure 2.2. VC formant transitions for PK tokens at each level of overlap, with example PP and KK formants given for comparison. Level of overlap is listed as a superscript over the formant curves, and a curve’s horizontal extent represents its duration. Time is not quantified on the horizontal axis, though, since formant curves are lined up left-to-right for exposition.

Figure 2.3. VC formant transitions for KP tokens at each level of overlap, with example KK and PP formants given for comparison. Presentation is the same as in figure 2.2.

20 Because Praat’s measurement of formant values becomes erratic near consonantal closures, I chose to drop the spectral slice nearest to the closure in order to smooth the displays in figures 2.2 and 2.3.
Judging from the formant transitions, increasing levels of modeled overlap affect the acoustic properties of the two heterorganic sequences in different ways. At low levels of overlap, VC transitions for PK and KP closely resemble those of pure labial and pure dorsal sequences, respectively. As overlap increases, the F1 and F3 transitions for PK fluctuate slightly and become somewhat more KK-like, but PK’s F2 transition does not appear to become more KK-like with higher overlap. In fact, at medium to high levels of overlap, PK’s F2 diverges markedly from that of both KK and PP, though it ultimately returns to a more PP-like shape at 13 frames of overlap. For KP sequences, the generalization is quite different: F1 and F3 show little acoustic assimilation to PP-like values (except perhaps at the highest levels of overlap), but F2 becomes significantly PP-like by around 11 frames of overlap. At higher levels of overlap, KP’s F2 continues to become less KK-like, but also diverges steadily from that of a “pure” PP. It is unclear why the two sequences’ VC transitions depend on overlap in such qualitatively different ways.

From the differences between PK and KP formant shapes in figures 2.2 and 2.3, we expect increasing gestural overlap to affect the perception of PK and KP clusters in distinct ways, depending on the individual importance of F1, F2, and F3 as cues to labial or dorsal POA. Note in particular that increasing overlap appears to create lower F2 transition values for both PK and KP, though at different rates. Assuming that falling F2 transitions are a primary cue for labial POA, we therefore anticipate that higher overlap will lead to more misidentifications of dorsal codas than of labial codas, contra hypothesis H1.1.

2.2 Method

2.2.1 Participants
Fourteen subjects participated, of whom ten were undergraduates drawn from UC Santa Cruz linguistics classes and compensated with class credit, and four were undergraduate volunteers. All possessed normal hearing, were naïve to the nature of the experiment.

2.2.2 Presentation
The experiment was administered by computer in a soundproof booth in the UC Santa Cruz phonetics lab, using the SuperLab (version 4.0.7b) experiment design software. The 44 experimental stimuli (4 word-types × 11 overlap degrees) were presented over headphones in randomized blocks within the framework of a forced-choice identification task. Within every block, each of the 11 tokens of the heterorganic sequences /ap.ka/ and /ak.pa/ was repeated twice, for a total of 44 presentations (11 overlap degrees × 2 sequences × 2 repetitions); and 9 of the 11 homorganic sequences /ap.pa/ and /ak.ka/ were presented once, totaling 18 presentations (9 overlap degrees × 2 sequences); overall there were a total of 62 presentations per block.²¹ Presentation order was randomized anew for each

---

²¹ Two out of 11 levels of overlap were omitted from the homorganic sequences in each block, though the choice of which two to omit changed from block to block in order to spread the omissions evenly across the 11 levels. For each of the two homorganic sequences, one level of
block, and blocks were separated by breaks whose lengths were controlled by the participants, during which they were free to move around. Six blocks were run in total, thus yielding 372 token presentations per experiment.

Subjects read a screen displaying the experiment’s instructions. After being asked by the experimenter if any clarification was needed, subjects began the experiment at their leisure; the instruction screen was replaced with a blank screen once any button was pressed, and stimuli started being presented 600 ms thereafter. The subjects’ task was to identify each stimulus token as one of the four word-types listed in §2.1.1. Upon hearing a token, subjects responded by pressing one of four buttons labeled “AKA”, “APA”, “AKPA”, or “APKA” on a button box. No time limit was placed on their response time; the program simply waited for a response before moving on to play an inter-stimulus interval of 600 ms and then another token.

2.2.3 Data collection
Raw response choices were collected for each subject into percentage correct figures using scripts in R (R Development Core Team 2009). Responses with reaction times under 10 ms and over 10000 ms (10 s) were discarded; this amounted to very few trials both overall and for each subject.

2.3 Results

Figure 2.4 shows the mean percentage of correct responses for each sequence type as a function of gestural overlap, ranging between 3 and 13 activation frames. Subjects’ individual rates of response accuracy are reported in appendix A. Responses to homorganic sequences were generally more accurate than responses to heterorganic sequences, averaging 82% vs. 62% overall, though the homorganic sequences fared worse than the heterorganic ones at the lowest levels of overlap. The overlap dependencies of the heterorganic and homorganic sequences in fact showed opposite trends: PP and KK identification became more accurate as overlap increased, while PK and KP identification became less accurate. Among homorganic sequences, KK was identified overall less accurately, averaging 78% correct vs. 85% correct for PP.

As for heterorganic sequences, despite much similarity at low overlap levels the accuracy curves for PK and KP exhibit qualitatively different shapes, as expected. Responses to both sequences tended strongly to get less accurate as overlap increased, and both display a notable drop in accuracy between 4 and 5 frames, corresponding to the overlap level where coda release bursts are lost.

overlap ended up with one fewer presentation than the others. For KK this was overlap level 8, and for PP it was overlap level 9. Thus, over the course of 6 blocks, each KK and PP token came to be repeated 5 times, except for the two tokens just mentioned, which came to be repeated 4 times. This design feature is an obsolete carry-over from a previous version of the experiment which was perpetuated by accident. Though odd in appearance, this distribution of homorganic presentations should not foreseeably invalidate any of the conclusions regarding the heterorganic sequences, which are the true focus of the experiment.
Figure 2.4. Mean response accuracy for experiment I as a function of overlap level and cluster type. Mean correct responses to homorganic clusters are shown on the left, and to heterorganic clusters on the right.

The two sequences exhibit very little difference in behavior under low levels of overlap (<11 frames). If anything, it appears that PK sequences were very slightly more accurately identified than KP sequences. At high levels of overlap (11-13 frames) though, the two diverge sharply. Between 10 and 12 frames, PK responses increased in accuracy, and returned to the greater-than-chance level of 52% at 13 frames. Between 11 and 13 frames of overlap, responses to KP, on the other hand, abruptly and consistently fell to around 8% accuracy (well below the chance response rate of 25%).

2.3.1 Analyses of variance

A repeated measures ANOVA was conducted to verify these distinct types of overlap dependency for the heterorganic clusters, with cluster type and level of overlap as independent variables and accuracy as the dependent variable. As anticipated, a highly significant main effect for overlap was found [F(10,130) = 41.134, p < 0.001***], as well as a significant main effect for cluster type [F(1,13) = 9.027, p = 0.0102*]. The interaction between overlap and cluster type was highly significant [F(10,130) = 16.448, p < 0.001***], indicating that accuracy depended on level of overlap in significantly different ways between the two sequences. Within each level of overlap, Bonferroni-corrected post hoc tests22 verified that the two cluster types did not significantly differ at any level less than 11 frames (p > 0.32 in each case), but KP responses were less accurate than PK responses to a highly

---

22 For post hoc testing within overlap degree, 11 pair-wise comparisons were made, necessitating p-values of p<[0.0045*, 9.1e-4**, 9.1e-5**] rather than the usual p<[0.05*, 0.01**, 0.001***].
significant degree at 11, 12 and 13 frames of overlap [F(1,13) = 42.786, p < 9.1e-5***; F(1,13) = 81.213, p < 9.1e-5***; and F(1,13) = 42.681, p < 9.1e-5***, respectively].

2.3.2 Discussion

Some aspects of experiment I’s results accord with previous studies, but other aspects are quite unexpected.

- Like Byrd (1992), we find that greater overlap has the effect of rendering categorization of heterorganic clusters less accurate. Both studies also find an asymmetry in how coda POA is perceived under high levels of gestural overlap. Byrd found that coronal codas perceptually assimilate more readily before labials than vice versa, and in this study we find that dorsal codas do as well.

- Unlike Byrd (1992), however, we find a further asymmetry in how abruptly labials and dorsals succumb to this perceptual masking. While labial codas gradually decline in perceptibility as overlap with a following dorsal increases, dorsal codas quickly become unrecoverable once a following labial is overlapped more than a certain critical amount. At no point did overlapped labials become unrecoverable. This result is partially in agreement with the outcome of Son et al.’s (2007) perceptual study of PK sequences. They found that overlapped, naturally produced PK sequences were perceived with quite low mean sensitivity (d’ ≈ 0), but that individual subjects’ sensitivity to these sequences varied widely, ranging from -3.17 to 4.65, indicating that labial codas were not rendered unrecoverable when overlapped by a following dorsal.23

- As Son et al. (2007) anticipated, low-overlap dorsal codas did prove to be more identifiable than some labial coda tokens at the highest levels of overlap, supporting their claim that highly overlapped labials can be “natural” targets of assimilation in a system where dorsals are not targeted.

- Somewhat unexpectedly and unlike Byrd (1992), we found that homorganic sequences tended to improve in accuracy as overlap increased.

Byrd (1992) argued that the coronal/labial perceptual masking asymmetry she found is plausibly related to gesture duration: the briefness of coronal coda gestures leads them to be fully obscured at a lower level of overlap than it takes to obscure labials. Can we call on a similar explanation for the masking results for experiment I? I.e., are dorsals more effectively masked by following labials because TADA models dorsal codas with shorter-duration gestures than labials? Judging from output closure durations, the answer appears to be “no”: table 2.2 shows that labials have consistently shorter closure durations than dorsals do. The fact that labials are shorter than dorsals should lead them to be the more easily masked of the two in coda position. That labials still survive better under high overlap indicates that our results are not dependent on the articulatory speed of dorsal and labial gestures realized in TADA, making them less likely to be mere artifacts of the simulation.

23 This sensitivity pattern was found for both native English- and Korean-speaking listeners.
The F2 VC transition values in figures 2.2 and 2.3 appear to be well correlated with perceptual assimilation in PK and KP clusters, and corroborate the stark difference in recoverability observed for these clusters under high overlap. The rise in PK accuracy at 12 frames may be related to the behavior of PK’s F2 transitions at that level of overlap: the value of F2 for PK at 12 frames is similar to the F2 values measured at much lower levels of overlap, indicating that PK formants at 12 frames are still quite labial-like. It is unknown why PK at 13 frames is not perceived with similar accuracy, despite a similar F2 transition value. Comparing the KP response curve with the VC formant measurements in figure 2.3, we see that the drop in KP accuracy at around 10 frames of overlap coincides with the KP stimuli’s F2 values becoming more PP-like than KK-like.

However, formant transitions do not explain why accuracy for homorganic clusters also depends on overlap. Since PP and KK formant transitions do not vary with overlap, this dependency must be related to closure duration. The homorganic response trend could arise from the labeling of the PP and KK response buttons as “APA” and “AKA”, leading subjects to listen for singleton stops rather than stops of geminate length. This would explain why PP and KK response accuracy increases with greater overlap: as overlap increases, the closure durations of PP and KK sequences become closer to those of singleton [p]s and [k]s, which are themselves identifiable through robust onset POA cues (Repp 1978, Dorman et al. 1979). By this logic, subjects faced with a low-overlap, long-closure homorganic sequence only discover that it is homorganic after interpreting the VC formant transitions and finding that they match the onset POA cues. Their overall categorization accuracy will then reflect (in fact, it will be roughly upper-bounded by) their accuracy at categorizing coda POA. This hypothesis is tested in experiment II. That KK responses were overall less accurate than PP responses may also be attributable to the greater (i.e., more geminate-like) closure length of KK clusters relative to PP clusters. It may also be a reflection of the greater indistinctness of unreleased dorsal codas in general.

2.4 Conclusions

Experiment I suggests a rejection of hypothesis H1.1: recovering coda place in PK sequences does not prove more difficult than in KP sequences under high levels of gestural overlap. At most levels of overlap the two are similarly recoverable, with neither sequence being perceived significantly more reliably than the other. At the highest levels of overlap, subjects remain able to identify PK signals while losing all ability to identify KP signals.

However, a confound remains in that we cannot be fully certain how much the present findings are attributable to variation in VC formant transitions or to variation in closure durations. Because the two physical dimensions co-vary in

24 Geminate-length stops are also, of course, not attested lexically in English, which could have led them to simply sound more unfamiliar or more like heterorganic clusters. Byrd (1992) gets around this problem by having her subjects treat /b.b/ and /d.d/ clusters as though they cross a word boundary; as a consequence, her accuracy results for homorganic sequences do not show a dependency on overlap.
our heterorganic stimuli, experiment I’s design cannot tease apart which is the source of the asymmetry between labial and dorsal codas. It could be the case, for instance, that the perceptibility of KP declines faster than that of PK because TADA arbitrarily created closures that are shorter for KP clusters than for PK clusters. An examination of table 2.1 shows that this possibility cannot be ignored: KP closures are indeed shorter than PK closures. If this is the cause of the asymmetry, then we cannot claim to have found an authentic difference in how labials and dorsals react to gestural overlap in coda position. A difference in closure duration could just be an artifact of TADA’s gestural model. To eliminate this unwanted possibility, we must verify that VC formant transitions are the true source of the asymmetry found in experiment I.

3 EXPERIMENT II

The purpose of experiment II is to isolate the source of the perceptual asymmetry found among overlapped labial and dorsal codas in experiment I. In particular, we want to know whether an asymmetry persists among heterorganic sequences when only VC formant transitions are available as a cue to coda POA. If such an asymmetry does persist, then we can claim to have found support for rejecting hypothesis H1.1.

3.1 Method

3.1.1 Stimuli

This experiment used stimuli which were adapted from those in experiment I. Since the purpose of experiment II is to isolate the perceptual influence of gestural overlap on VC formant transitions, it is necessary to eliminate any information conveyed by closure duration, bursts, and CV formant transitions. This is achieved by truncating the experiment I stimuli just after the onset of C1 closure silence, so that only their VC portions – without C1 release – remain (see Byrd 1992 and Surprenant & Goldstein 1998 for similar use of truncation). The resulting VC stimulus tokens were all around 240 ms long (range: 225 ms – 248 ms; see §4.1.1 for further discussion).

Since the VC formant transitions of the homorganic sequences PP and KK do not vary significantly with overlap, only one truncated token of each was used in experiment II, namely the tokens produced at 3 frames of overlap. These two homorganic tokens, along with the 22 heterorganic tokens adapted in the above fashion, make a total of 24 stimuli used in experiment II.

These truncated stimuli will be referred to henceforth as AP(KA), AK(PA), AP(PA), and AK(KA) tokens, to underscore their distinctness from the full tokens used in experiment I.

3.1.2 Subjects

Eighteen subjects participated in experiment II, all of them UC Santa Cruz undergraduates from linguistics courses who were compensated with course
credit. The data from three subjects had to be excluded because of the loss of a large number of trials (> 15%) due to a computer error. Of the remaining fifteen, two were excluded because they reportedly did not hear any AK(PA) or AK(KA) tokens and thus responded ‘AP’ for every trial. This left thirteen subjects whose data could be used for analysis.

3.1.3 Presentation
Experiment II was administered largely as in experiment I, with the following exceptions. Each block of trials involved two repetitions of each of the 22 truncated heterorganic stimuli (44 trials) and 10 repetitions of both truncated homorganic stimuli (20 trials) for a total of 64 trials per block, whose presentation order was randomized by block. The number of stimulus repetitions was chosen to keep the ratio between heterorganic and homorganic trials close to that of experiment I. Five blocks were presented, bringing the total number of trials in experiment II to 320.

Subjects were given only two response choices in experiment II, having been instructed to identify what they heard by pressing either ‘AP’ or ‘AK’ on the button box. After pressing a button for their choice, an ISI of 800 ms was played before the beginning of the next trial.

3.1.4 Data collection
Raw response choices were collected for each subject into percentage correct figures using scripts in R (R Development Core Team 2009). Responses with reaction times under 10 ms and over 10000 ms (10 s) were discarded; this amounted to very few trials both overall and for each subject.

3.2 Results
Figure 3.1 shows the mean percentage of correct token categorizations in experiment II. Individual subjects’ performance is reported in appendix B. Several important generalizations can be made about the results in figure 3.1. First, there does not appear to be a consistent dependency between overlap and accuracy for AP(KA) tokens, which are correctly identified at an essentially uniform high rate relative to overlap, averaging 81% correct across all levels. AK(PA) tokens, on the other hand, display a steep drop-off in identifiability between 11 and 13 frames of overlap, bottoming out at 17% correct. At less than 10 frames of overlap, AK(PA) tokens are still identified less accurately than AP(KA) tokens, averaging 68% correct overall. The generally lower response accuracy for AK(PA) tokens is consistent with the high confusability of unreleased dorsal codas found by Winters (2001) and Kochetov & Šo (2007). It also coincides with the fact that two subjects responded ‘AP’ to every trial, but no subjects responded ‘AK’ to every trial – unreleased dorsal codas are more confusable than labials independently of gestural overlap.

While the latter two subjects’ ‘AP’-only results are excluded from the data analysis, the fact that they both massively misperceived in the same direction is discussed further below.
Figure 3.1. Mean response accuracy for experiment II as a function of overlap level and cluster type. Mean correct responses to homorganic clusters are shown on the left, and to heterorganic clusters on the right. Since homorganic stimuli at only one level of overlap (level 3, namely) were presented in experiment II, the results are shown as a bar graph.

As for homorganic sequences, AP(PA) tokens were identified more accurately than AK(KA) tokens, at 79% and 76% correct respectively, but a two-sample t-test showed that this difference is not significant (p = 0.74). Homorganic tokens were not identified at a significantly greater rate than their low-overlap heterorganic token counterparts, though AK(KA) accuracy is around 8% higher than AK(PA) accuracy. The difference between AK(PA) and AK(KA) is not significant for degrees of overlap less than 11 (p = 0.24), and neither is the difference between AP(KA) and AP(PA) across all degrees of overlap (p = 0.64).

3.2.1 Analyses of variance
The distinct overlap dependencies for AP(KA) and AK(PA) sequences found in this section were verified using a repeated measures ANOVA. Similarly to experiment I, highly significant main effects were found for cluster type \(F(1,12) = 54.321, p < 0.001^{***}\) and level of overlap \(F(10,120) = 15.536, p < 0.001^{***}\), and a highly significant interaction was found between these two variables \(F(10,120) = 15.883, p < 0.001^{***}\). These results validate the impression from figure 3.1 that AK(PA) sequences vary with increasing overlap in a distinct way from AP(KA) sequences. It also appears that AK(PA) sequences were categorized less accurately than AP(KA) sequences by an overall greater margin than in experiment I. Bonferroni-corrected post-hoc tests within each level of overlap were run to determine whether these larger accuracy differences were significant\(^{26}\); the differences between AK(PA) and AP(KA) at levels 3, 11, 12, and 13 were significant.

\(^{26}\) For n=11 post-hoc comparisons, the adjusted significance levels are \(p<(0.0045^{*}, 9.1e-4^{**}, 9.1e-5^{***})\).
First and most importantly, the steep drop-off in accuracy found for KP sequences in experiment I was reproduced in experiment II, in the absence of cues to closure duration. This persistent result indicates that a qualitative change in VC formant transitions is responsible for the drop in accuracy, as was predicted from direct examinations of the formant transitions themselves in figures 2.2 and 2.3. It is important to eliminate closure duration as a possible source of the change in KP categorization, since closure duration does not physically depend on overlap per se, while VC formant transitions do.27

Interestingly, the steady decline in accuracy for highly overlapped PK clusters in experiment I was not reproduced in experiment II, indicating that it was entirely due to decreasing closure durations. The non-dependency of AP(KA) accuracy on overlap in experiment II shows that labial VC formant transitions are extremely effective at masking the influence of an overlapping dorsal articulation.

27 That is, closure duration can be controlled at least partly independently of degree of overlap. The closure for a highly overlapped pair of stops can be freely lengthened, for instance.
identified at 73% accuracy. Collectively, these natural-speech studies indicate that the labial preeminence found in experiments I and II is not a fluke of the articulatory modeling in TADA, but rather a physically meaningful property of overlapped labials and dorsals.

3.3 Further questions

The preceding two experiments support the conclusion that codas in KP clusters evolve perceptually under increasing gestural overlap quite differently than codas in PK clusters do. Not only does the identifiability of a coda suffer more heavily under overlap in a KP cluster than in a PK cluster, but the rate of identity loss in KP codas is more sudden and drastic.

However, there is another possible explanation for why this asymmetry exists which is worth considering, given the asymmetry's unexpected nature. It might be argued that the effect is due to a categorization bias towards /p/ in coda position, rather than a true asymmetry in the perceptual properties of these two sequences under overlap. We have not ruled out the possibility that labial codas in PK clusters undergo just as drastic an overlap-driven perceptual change as dorsal codas in KP clusters do, but that the phonological category /p/ simply admits more allophonic variation in coda position than /k/ does. After all, labial codas in PK clusters do exhibit changes in their VC formant transitions under increasing overlap (cf. figure 2.2). Since the designs of experiments I and II depend on the activation of listeners' phonological categories rather than their primary perceptual impressions, our results may simply reflect a fact of their native phonology – rather than a language-independent fact of the mapping between articulatory output and perceptual input. If the phonological model that we are ultimately investigating is one in which perception drives language change, then we want to be as sure as possible that our results reflect perceptual facts, and not simply English-specific categorization biases.

The way to reject the possibility of such a categorization confound is to give listeners a task not involving categorization, but one which forces them to respond from their primary perceptual impressions of the stimuli. If we can measure these perceptual impressions, then it will be useful to answer a pair of questions relating to the results of experiments I and II:

- How perceptually “steep” is the drop-off in identifiability of highly-overlapped AK(PA) tokens? Is it any more abrupt than the perceptual changes that AP(KA) tokens undergo as overlap is increased?
- How completely do highly-overlapped AK(PA) tokens perceptually assimilate? Are they indistinguishable from underlyingly homorganic AP(PA)? Is highly-overlapped AK(PA) any less distinguishable from AP(PA) than AP(KA) is from AK(KA)?

---

28 Alternatively, subjects could be biased toward pressing the ‘AP’ or ‘APKA’ buttons, for unknown reasons.
If native-language categorization bias is truly not responsible for the findings in experiments I and II, then the answers to these questions should continue to reflect an asymmetry between AP(KA) and AK(PA). Also, testing for qualities like “steepness” informs our view of how small gestural variations are perceived and potentially incorporated into language change.

4 EXPERIMENT III

Experiment III employs a testing design that is meant to measure perceptual impressions of the truncated VC stimuli from experiment II. Working from the results of experiments I and II, we put forth the following hypotheses:

**Hypothesis H4.1**
Sensitivity to small increases in overlap is greater among AK(PA) tokens than among AP(KA) tokens, particularly at medium to high degrees of overlap.

**Hypothesis H4.2**
As overlap increases, AK(PA) tokens become more AP(PA)-like, but AP(KA) tokens do not become more AK(KA)-like.

Experiment III is organized around a same-different task, with comparisons made between different levels of overlap within a single heterorganic sequence, and between heterorganic and homorganic sequences. These comparisons are meant to circumvent potential biases associated with the activation of linguistic categories. If the codas in AP(KA) and AK(PA) tokens indeed undergo significant perceptual changes when overlap is increased, then differences between “nearby” levels of overlap within both sequences should be discernible upon comparison. Furthermore, comparing overlapped AP(KA) and AK(PA) codas with codas from homorganic AP(PA) and AK(KA) sequences should indicate how similar the perceptually assimilated codas are to “pristine” (homorganic) labial and dorsal codas.

4.1 Method

4.1.1 Stimuli
The stimuli used in this experiment were the same as the truncated VC stimuli used in experiment II. The sound files were all of similar (but not equal) duration, falling between 225 ms and 248 ms, as recorded in table 4.1.²⁹

---

²⁹ The AP(PA) and AK(KA) tokens used were those modeled at 3 frames of overlap, as in Experiment II.
Table 4.1. Durations (in ms) of truncated VC stimuli used in experiment III.

<table>
<thead>
<tr>
<th>Sequence type</th>
<th>AP(PA)</th>
<th>AK(KA)</th>
<th>AP(KA)</th>
<th>AK(PA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>248</td>
<td>241</td>
<td>245</td>
<td>235</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>245</td>
<td>234</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>244</td>
<td>234</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>244</td>
<td>235</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>244</td>
<td>235</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>-</td>
<td>244</td>
<td>235</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>-</td>
<td>244</td>
<td>235</td>
</tr>
<tr>
<td>10</td>
<td>-</td>
<td>-</td>
<td>242</td>
<td>237</td>
</tr>
<tr>
<td>11</td>
<td>-</td>
<td>-</td>
<td>238</td>
<td>235</td>
</tr>
<tr>
<td>12</td>
<td>-</td>
<td>-</td>
<td>233</td>
<td>236</td>
</tr>
<tr>
<td>13</td>
<td>-</td>
<td>-</td>
<td>225</td>
<td>228</td>
</tr>
</tbody>
</table>

The duration differences between the stimuli result from differences in the modeled kinematic properties of labials and dorsals in TADA, and from the attainment of onset closures prior to coda closures at the highest levels of overlap. Though small, the largest of these durational differences (around 23 ms) is slightly above the JND for durations of this magnitude found by Meng et al. (2006): around 7.5% of stimulus duration, or ~19 ms. Thus, duration could potentially (and undesirably) provide a small amount of information for distinguishing some of the experimental stimuli, and is tested in §4.2.5 as a possible confounding factor.

4.1.2 Subjects
Sixteen subjects took part in experiment III, of which one was a non-Linguistics post-doctoral student and the rest were UC Santa Cruz undergraduates enrolled in linguistics classes. Subjects were either given linguistics course credit for participation or were compensated at a standard rate. As in experiment II, a computer error forced the data of two subjects to be discarded, leaving fourteen subjects whose data were analyzed.

4.1.3 Presentation
Subjects were seated in a soundproof booth in front of a computer screen, and stimuli were presented to them over headphones in SuperLab (version 4.0.7b). Stimuli were paired together in AXA presentations in a roving same-different task. In an AXA presentation trial, a single reference stimulus (A) is presented both before and after the focus stimulus (X), which is randomly chosen to be either labial or dorsal.

---

30 This JND value was determined by Meng et al. (2006) for 300 ms white noise bursts; it is assumed that the duration JNDs for white noise and speech-like stimuli do not differ appreciably from one another.

31 A roving task is one in which stimuli are reused for different comparison trials within a single block (Macmillan & Creelman 1991/2005:221).
identical to or different from the reference. The AXA presentation type is a variant of the more common AX presentation type and is intended to make the very close comparisons in this experiment easier to distinguish, essentially by providing two comparison opportunities per trial (i.e., AX and XA).

Subjects were instructed to decide whether the middle focus stimulus was the same or different from the flanking reference stimuli, and to record their decision on a labeled button box. Their reaction speed was timed from the onset of the third stimulus. Subjects were also asked to respond as quickly as possible to avoid getting a warning from the computer, which was administered on the screen if reaction time exceeded 940 ms. Since reaction time is measured from the beginning of the last stimulus, and stimuli are around 240 ms in duration, limiting reaction times to 940 ms effectively means limiting responses to within 700 ms after the last stimulus.

The temporal makeup of trials is depicted in figure 4.2. Each trial is preceded by 1000 ms of silence. Within a trial, each pair of adjacent stimuli is separated by an ISI of 100 ms. With two such ISIs and the roughly 240 ms duration of each of the three stimuli, each presentation ends up having a duration of around 920 ms.

**Figure 4.2. Time-course of AXA trials.**

If we estimate that subjects largely respond within 700 ms of the end of the last stimulus, an average experimental trial can be expected to take under 2.6 seconds.

The choice of A and X stimulus pairings was made in order to provide comparisons of the types mentioned at the beginning of this section, and is summarized in table 4.2. In “Same” comparisons, A = X; each block of trials contains 24 Same comparisons – one for each of the 22 heterorganic and 2 homorganic tokens. Among “Different” trials, two broad types of comparisons are made: (i) within-sequence comparisons between tokens of a single heterorganic sequence at different levels of overlap; and (ii) between-sequence comparisons between a heterorganic token and one of the two homorganic tokens AP(PA) or AK(KA).

Within-sequence comparisons come in two varieties, one juxtaposing heterorganic tokens that are two levels of overlap apart from one another (e.g., comparing AK(PA) tokens at 8 and 10 frames of overlap) and another juxtaposing tokens three levels apart (e.g., AP(KA) tokens at 4 and 7 frames of overlap).

---

32 It is assumed that the crucial information distinguishing A from X comes only at the end of the second stimulus, which renders the beginning of the third stimulus the earliest practical point at which a Same/Different judgment can be made.
The rationale behind having these two varieties of within-sequence comparisons is to be able to suit subjects with a wide range of sensitivities. It is hard to tell, \textit{a priori}, if two frames of overlap distance is too small for a given interesting comparison, or if three frames is too large, so both are included for good measure. For each heterorganic sequence AK(PA) or AP(KA), there are 9 such two-frames-apart comparisons and 8 such three-frames-apart comparisons to be made, yielding 34 total within-sequence comparison trials.

Between-sequence comparisons juxtapose heterorganic tokens at all levels of overlap with either of the two homorganic tokens AP(PA) or AK(KA). 22 such between-sequence comparisons can be made for each of the two sets of heterorganic tokens, totaling 44 trials of this type. A block of trials therefore contains 78 Different comparisons in all, for a total of 102 comparisons of all types per block. Six blocks of trials were presented, resulting in six repetitions of each of the comparisons in table 4.2, for a total of 612 trials. The order of trials was randomized within blocks. Half of the blocks were also used to counterbalance the order in which stimuli were presented within trials. In particular, three blocks contained trials whose stimuli were in XAX order, while the other three contained
trials with stimuli in AXA order. AXA blocks were interspersed evenly with XAX blocks.

The low number of Same trials relative to Different trials summarized in table 4.2 was meant to compensate for the subjective impression that many of the Different trials (e.g., low-overlap within-sequence comparisons) sound indistinguishable from Same trials. This impression was verified in a pilot study: despite three-quarters of the stimuli being Different trials, ‘Same’ judgments accounted for 60% of pilot subjects’ responses.33 Anticipating that a good deal of sensitivity is required to discern some of the important comparisons in Experiment III, it was felt that a large number of Same trials would either discourage subjects from responding ‘Different’ or lead to response errors from the repetitiveness of ‘Same’ button presses. Both cases would prevent the experiment from accurately estimating subjects’ sensitivities to subtle stimulus comparisons.34

4.1.4 Predicted performance
If the results of experiments I and II are due to perceptual changes in the stimuli and not to categorization biases, then we can anticipate a number of effects among the four comparison types described above. If hypothesis H4.1 is correct, then the two within-sequence comparison types should show greater sensitivity (at high overlap) for AK(PA) comparisons than for AP(KA) comparisons. This pattern would indicate that the “abruptness” of overlap-induced perceptual changes was greatest among AK(PA) tokens, as suggested by experiments I and II. In particular, the especially abrupt change in categorization of AK(PA) tokens between overlap levels 10 and 11 should be registered as an increase in sensitivity to within-sequence comparisons across that “critical line”.

If hypothesis H4.2 is correct, then comparisons at high overlap should be more discernible between AP(KA) and AK(KA) than between AK(PA) and AP(PA). This pattern would indicate that the AK(PA) tokens perceptually assimilate under high overlap, but that AP(KA) tokens do not. The remaining between-sequence comparisons, namely AP(KA)/AP(PA) and AK(PA)/AK(KA), should pattern the opposite of the ones just mentioned. Particularly worthwhile is the question of whether AP(KA) sequences are truly unaffected by overlap; if so, then they should be uniformly indiscernible (or nearly so) from AP(PA) at all overlap levels.

4.1.5 Data collection and analysis
Raw response choices were collected for each subject into sensitivity scores using scripts in R (R Development Core Team 2009). Responses with reaction times under 10 ms and over 10000 ms (10 s) were discarded; for all but two subjects, this amounted to very few trials. For these two subjects it led to the discarding of more than 15% of their data, forcing them to be dropped from the analysis, as mentioned in §4.1.2. The underlying computer error forcing these data to be

---

33 This rate of “Same” responses is similar to the rate of 66% observed in Experiment III proper.
34 This approach is not unprecedented: see Guenther et al. (1999) for an example of a same/different experiment also relying on subtle stimulus differences, in which only 25% of all trials were Same trials.
discarded was the same as for experiment II, and was patched after these subjects’ data was lost.

Subjects’ sensitivity scores were measured using the criterion $d'$ (Macmillan & Creelman 2005:8) as a way of controlling for subject-internal response bias. Subject responses are classified relative to the categories hit, miss, false alarm (FA), and correct rejection (CR) in the manner depicted in figure 4.3. The value of $d'$ depends on the hit rate and FA rate, i.e. hits/(hits + misses) and FAs/(FAs + CRs).

Figure 4.3.  Response categories used to compute $d'$ for Different comparisons.

<table>
<thead>
<tr>
<th>Subject response</th>
<th>‘Different’</th>
<th>‘Same’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Different</td>
<td>Hit</td>
<td>Miss</td>
</tr>
<tr>
<td>Same</td>
<td>False Alarm (FA)</td>
<td>Correct Rejection (CR)</td>
</tr>
</tbody>
</table>

In a roving same/different task such as the one employed in experiment III, subjects are assumed to make decisions based on a differencing strategy. This strategy is different than the independent observation strategy standardly assumed to describe subject behavior in yes/no and identification tasks, and necessitates a different calculation of $d'$ than the usual $z(H) - z(FA)$ equation (Macmillan & Creelman 2005:221-228). Differencing model $d'$ values can be determined from look-up tables in appendix 5.4 of Macmillan & Creelman (2005:401-419), or using the Psyphy Package for R (Knoblauch 2009). The latter method is used here.

By the nature of the function used in $d'$ calculation, higher $d'$ values can reflect either higher hit rates or lower FA rates. The effect of this function is to subtract out the subject’s general bias towards reporting a signal even when a non-signal is presented. The function values for hit and FA rates of 0 or 1 are undefined, so any such examples in the data must be shifted up or down by a small amount to give sensible transform values. This amount is commonly chosen to be half of the difference between 0 (or 1) and the next smallest (or largest) possible hit/FA rate found in the data. In this experiment, the smallest possible non-zero hit/FA rate is 8%, so hit/FA rates of 0 and 1 are shifted to 4% and 96%, respectively. Once adjusted in this way, the range of possible $d'$ values falls between 0 and 5.38, with the chance response rate at $d' = 0$, and noticeable perception above around $d' = 1$.

One final data adjustment had to be made to fit the requirements of the differencing model $d'$ function, which is undefined when FA rate > hit rate. Though $d' = 0$ is the lowest theoretically possible value for sensitivity, and assumes at least parity between hit and FA rates, Macmillan & Creelman (2005:15) note that FA rates can often exceed hit rates in experiments with low numbers of observations. It is therefore justifiable to treat such instances as being due to
4.2 Results

Sensitivity to within-sequence comparisons is discussed in §4.2.1, and to between-sequence comparisons in §4.2.2. Individual subjects’ performance for all comparisons is reported in appendix C.

4.2.1 Within-sequence trials

Figure 4.3 lays out the mean sensitivity results for comparisons within-sequence. As described above, these trials test for sensitivity to small changes in overlap, and bear on hypothesis H4.1. Sensitivity for both types of within-sequence trial was generally low, owing to the subtlety of the differences involved, but a noticeable increase at high levels of overlap can be identified in both graphs. Although notable sensitivities can be seen for some of the AP(KA) trials, namely for “11-13” ($d' = 1.59$) in the left-hand graph of figure 4.3, and for the “7-10” ($d' = 1.21$) and “10-13” ($d' = 1.82$) comparisons in the right-hand graph, the greatest overall sensitivity is to AK(PA) trials. Several AK(PA) trials exceed the largest sensitivity recorded for AP(KA) in either condition (namely, $d' = 1.82$). In particular, the AK(PA) comparisons “10-12” ($d' = 2.26$), “9-12” ($d' = 2.74$), and “10-13” ($d' = 3.09$) are all discernible above this level.

Figure 4.3. Mean sensitivity ($d'$) to within-sequence comparisons in experiment III, as a function of the overlap levels being compared and of the “distance” of comparison. Sensitivity to comparisons between stimuli that are 2 frames of overlap apart is shown on the left, and between stimuli 3 frames apart on the right.
Not surprisingly, each of these high-sensitivity AK(PA) trials involved a comparison across the “10-11 critical line” mentioned in §4.1.4.

At low overlap levels, sensitivity to comparisons within both AP(KA) and AK(PA) tended to be around $d' = 0.5$. Examination of the raw data for low-overlap within-sequence comparisons indicated several cases of FA rates exceeding hit rates, which were readjusted to $d' = 0$ as described in §4.1.5. In experiment III, hit rates were calculated from only six observations, suggesting that statistical error is indeed responsible for the unusual hit and FA rates.\footnote{Also potentially to blame is the fact that a single Same comparison often contributes to the FA rate for two $d'$ estimations. This is due to the reuse of many stimuli for both “left-hand” and “right-hand” comparisons, as in comparing e.g. overlap level 5 to levels 3 and 7. The Same trials for overlap level 5 will contribute to the FA rate for the 3-5 comparison and the 5-7 comparison. This may allow anomalies (i.e., high FA rates) in individual comparisons to influence multiple $d'$ estimations.} Note, though, that adjusting $d'$ to zero is equivalent to artificially raising hit rates or lowering FA rates. If “true” hit and FA rates were taken into consideration, it is possible that the aggregate $d'$ values among these comparisons would in fact be lower than the $d' = 0.5$ level seen in figure 4.3. For our purposes, this possibility simply underscores the perceptual uniformity of low-overlap heterorganic sequences: the tokens are almost indiscernible from one another.

All in all, the results in figure 4.3 show that within-sequence AK(PA) comparisons are more discernible under high overlap than within-sequence AP(KA) comparisons are. This implies that small articulatory changes map onto large perceptual changes for overlapped AK(PA) codas, but not for overlapped AP(KA) codas.

### 4.2.2 Between-sequence comparisons

The sensitivity results for comparisons between heterorganic tokens and the two homorganic tokens are shown in figures 4.4 and 4.5. Figure 4.4 groups the results by comparison type (AP(PA) vs. AK(KA)). Figure 4.5 rearranges the results in terms of the comparisons that bear on hypothesis H4.2. That is, figure 4.5 is meant to contrast perceptual assimilation among the heterorganic stimuli, by juxtaposing AP(KA)/AK(KA) comparisons with AK(PA)/AP(PA) comparisons (the “full-assimilation” comparisons), and AP(KA)/AP(PA) comparisons with AK(PA)/AK(KA) comparisons (the “unassimilated” comparisons).

As is to be expected, sensitivity to comparisons where POA does not match in C1 were quite high at low levels of overlap, while comparisons involving matching POA were indistinguishable at low overlap. At high overlap levels, perceptual assimilation appears to reverse this trend: sensitivity to both AK(PA)/AP(PA) and AP(KA)/AK(KA) comparisons drops. Figure 4.5 illustrates that AK(PA) sequences are more greatly affected by increasing overlap than AP(KA) sequences are, as was also found among within-sequence comparisons. Interestingly, neither of the drops in sensitivity to “full-assimilation” comparisons at high overlap quite reach $d' = 0$, instead staying largely above $d' = 1$. 
Figure 4.4. Mean sensitivity ($d'$) to between-sequence comparisons in experiment III, as a function of overlap level and comparison. Sensitivity to comparisons between heterorganic sequences and AP(PA) is shown on the left, and between heterorganic sequences and AK(KA) on the right.

Figure 4.5. Mean sensitivity ($d'$) to comparisons between heterorganic sequences and “full-assimilation”/“unassimilated” sequences in experiment III. The results in figure 4.4 are rearranged in terms of comparisons between sequences sharing either the same $C_2$ POA (i.e., “full-assimilation” comparisons, on the left) or the same underlying $C_1$ POA (“unassimilated” comparisons, on the right).
We can take this behavior as an indication that perceptual assimilation is not absolute. While an overlapped, perceptually assimilated dorsal coda sounds largely labial, it still possesses a discernible non-labial component. This non-labial component appears not to be identifiably dorsal, however, judging from subjects’ great sensitivity ($d' > 2.95$) to the AK(PA)/AK(KA) comparison at high overlap ($> 11$ frames). On the other hand, we can infer that a perceptually assimilated AP(KA) token takes on a distinctly dorsal quality at the highest level of overlap (13 frames). This follows from the fact that AP(KA) becomes only moderately distinguishable from either AP(PA) ($d' = 2.02$) or AK(KA) ($d' = 1.84$) at this overlap level. Against our speculations in §4.1.4, it appears that AP(KA) tokens are not uniformly indiscernible from AP(PA) tokens, but rather tend towards being mildly discernible at medium to high levels of overlap.

4.2.3 Analyses of variance
Repeated-measures ANOVAs were run separately for the results of each of the four comparison types reported in §4.2.1 and §4.2.2. Some of the factors do not reach significance, which we provisionally attribute to the low number of subjects run in experiment III.

Among the within-sequence conditions, no main effect of sequence type (AP(KA) vs. AK(PA)) was found for either two-frames-apart trials [$F(1,13) = 0.661, p = 0.431$] or three-frames-apart trials [$F(1,13) = 3.452, p = 0.086$], although the latter was a nearly significant trend. This result implies that neither of the sequence types diverged greatly from each other or from $d' = 0$ overall, which is to be expected given the large number of indistinguishable low-overlap comparisons contributing to mean sensitivity. Significant main effects of overlap level were found for two-frames-apart trials [$F(8,104) = 11.223, p < 0.001^{***}$] and for three-frames-apart trials [$F(7,91) = 17.055, p < 0.001^{***}$]. Finally, a significant interaction between sequence type and overlap level was found for both two-frames-apart trials [$F(8,104) = 2.911, p = 0.0057^{**}$] and three-frames-apart trials [$F(7,91) = 8.267, p < 0.001^{***}$]. These significant interactions verify the impression that AP(KA) comparisons are affected by overlap differently from AK(PA) comparisons. Post-hoc tests were run to investigate the source of the significant interaction between overlap and sequence type. For two-frames-apart trials, only the “10-12” comparison showed a significant difference$^{37}$ between AP(KA) and AK(PA) [$F(1,13) = 13.758, p = 0.0026^*$], and for three-frames-apart trials only the “9-12” comparison showed a significant difference$^{38}$ [$F(1,13) = 15.795, p = 0.0016^*$]. Note that each of these significant trials involves a comparison across the “10-11 critical line” for AK(PA), exactly where we expect the largest divergence between AP(KA) and AK(PA) behavior. This verifies the “critical line” effect anticipated in §4.1.4.

For the between-sequence trials, it is most sensible to use an ANOVA to examine how AP(KA) and AK(PA) sequences fared in “full-assimilation” vs. “unassimilated” comparisons – that is, to test our impressions from figure 4.5 rather than figure 4.4. A significant main effect of sequence type was found within

---

$^{37}$ For n=9 post-hoc comparisons, the adjusted significance levels are $p<0.0056^*, 0.0011^{**}, 1.1e-4^{***}$.

$^{38}$ For n=8 post-hoc comparisons, the adjusted significance levels are $p<0.0063^*, 0.0013^{**}, 1.3e-4^{***}$. 
the “full-assimilation” comparisons \([F(1,13) = 40.429, p < 0.001^{***}]\), but the main effect for sequence type in “unassimilated” comparisons was a nearly-significant trend \([F(1,13) = 4.615, p = 0.051]\). Judging from the left-hand graph in figure 4.4, this effect seems to reflect a higher sensitivity across all overlap levels to AP(KA)~AK(KA) trials than to AK(PA)~AP(PA) trials, which is surprising given what should be the symmetric nature of the comparisons at low overlap levels. This asymmetry may indicate that the perceptual qualities of either labial or dorsal \(C_1\) differ subtly between heterorganic and homorganic sequences, although such an effect does not seem to be great enough to have influenced performance in the “unassimilated” condition. Highly significant main effects of overlap level were found within both “full-assimilation” \([F(10,130) = 11.631, p < 0.001^{***}]\) and “unassimilated” comparisons \([F(10,130) = 19.965, p < 0.001^{***}]\), verifying that \(C_1\) indeed undergoes a measure of perceptual assimilation at high overlap levels. Significant interactions between overlap level and sequence type were found within both “full-assimilation” comparisons \([F(10,130) = 3.266, p < 0.001^{***}]\) and “unassimilated” comparisons \([F(10,130) = 5.749, p < 0.001^{***}]\), indicating that AP(KA) and AK(PA) sequences perceptually assimilate to different degrees under increasing overlap. These results mirror the perceptual assimilation findings from experiment II, although here we have found evidence that AP(KA) becomes somewhat dorsal-like at the highest level of overlap, despite being categorized reliably as “AP” in experiment II.

4.2.4 Discussion

The results of experiment III confirm hypothesis H4.1, but are somewhat equivocal regarding hypothesis H4.2. The magnitude of overlap-induced changes among AK(PA) tokens was indeed found to be greater than the changes among AP(KA) tokens, as H4.1 predicted, though small changes among the latter were recorded. Both sequences strongly resembled labial-only homorganic tokens at the highest overlap levels, supporting H4.2, though the resemblance was not perfect. Against H4.2, AP(KA) took on a distinctly dorsal character at overlap level 13, which accords with a modest decrease in identifiability for that token in experiment I (from 72% to 52%), though not in experiment II. Perceptual assimilation for AP(KA) at this level of overlap is understandable, given that \(C_3\) closure precedes \(C_1\) closure in this token. Despite the lack of full support for H4.2, experiment III largely verifies that the results of experiments I and II arise from perception and not categorization.

However, some evidence from experiment III is consistent with a categorization bias towards /p/ in experiment II. At the highest level of overlap, AP(KA) tokens were roughly equally confusable with AP(PA) and AK(KA) in experiment III, but were consistently categorized as ‘AP’ in experiment II.\(^{39}\) This suggests that /p/ admits more contextual allophony than /k/, implying that the category boundary is closer to the /k/ prototype than to the /p/ prototype. If

\(^{39}\) As mentioned above though, in experiment I, a drop in identifiability was found for PK sequences at 13 frames of overlap.
present, such a categorization bias in experiment II is at most a small effect – not sufficient to throw doubt on our rejection of hypothesis H1.1.

4.2.5 Correlation with acoustic measures

A final potential confound that must be dismissed is related to the acoustic makeup of experiment III stimuli, in particular to the differences in stimulus length mentioned in §4.1.1. While it was expected that differences in formant values would contribute most to subjects’ sensitivity to the comparisons in experiment III, we cannot rule out that stimulus length may play a role, too. Recall that the largest difference in stimulus duration was 23 ms, on par with the JND for duration in these stimuli. If length differences contribute a large perceptual effect, then we cannot be sure that experiment III is a valid test of the results in experiments I and II.

Two types of acoustic distance estimates were made for the comparisons in experiment III based on stimulus measurements. The first measure quantifies the differences in VC formant transition values between the compared stimuli; the second quantifies the differences in overall stimulus length. These acoustic measures were then correlated with subject response data to test for strength of fit.

The difference in F1, F2 and F3 transitions between two stimuli was measured (separately for each formant) in terms of the difference between those stimuli’s formant values collected through spectral slices (shown in figures 2.2 and 2.3 in §2.1.2) and integrated over time. The difference between, e.g., F2 transitions in two stimuli amounts to an approximation of the integral of the frequency divergence in their F2 values over time, and is measured in units of ΔHertz-seconds (ΔHz*s). The measure is only approximate though, in that the formant differences were evaluated only at the (evenly-spaced) discrete times corresponding to spectral slices. The resulting difference measurements were then summed as a middle Riemann sum over the intervals between slices (7 intervals of 6.25 ms each). As for stimulus duration, differences between stimuli were calculated from the data reported in table 4.1.

These two measures of stimulus distance were computed for each of the comparison types listed in table 4.2, and the results are listed in appendix D. The measurements were then correlated with mean sensitivity over all comparisons using Spearman’s correlation coefficient for ranked data. The results of this correlation analysis are listed in table 4.3.

| Table 4.3. Correlations ($r^2$) between acoustic measures of comparison distance and subject sensitivity. Significance levels are all $p < 0.001$***. |
|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| F1 distance | F2 distance | F3 distance | Total formant distance | Duration distance |
| 0.47 | 0.69 | 0.64 | 0.79 | 0.27 |

Correlations were relatively high between subject sensitivity and F2 and F3 differences ($r^2 > 0.64$), though F1 differences were correlated with sensitivity to a
lesser degree \( (r^2 = 0.47) \). The best correlation was between sensitivity and the sum of all formant differences \( (r^2 = 0.79) \), indicating that subjects’ ability to discriminate differences between stimuli was highly dependent on overall differences in formant values. Differences in stimulus duration were correlated positively with subject performance, though only very weakly \( (r^2 = 0.27) \).

Taken all together, these results indicate that stimulus duration contributed only a very small perceptual effect at best, relative to formant values. Experiment III thus provides a valid corroboration of the conclusion that the results of experiments I and II reflect perceptual changes in VC formant transitions, and not native-language categorization biases.

4.3 Wrap-up of experiments I, II and III

The three experiments reported in this paper have given us reason to soundly reject hypothesis H1.1, and shed light on the open issues raised in §1.3.2. The major results can be summarized as follows:

- Dorsal codas undergo significant perceptual assimilation when overlapped with a following labial, but labials are not significantly affected by overlap with a following dorsal.
- Some labial PK codas at the very highest levels of overlap were misperceived slightly more often than dorsal KP codas at low levels of overlap.
- Perceptual assimilation in overlapped dorsal KP codas has its locus in changes in VC formant transitions.
- Perceptual assimilation in overlapped labial PK codas is chiefly related to shortened consonantal closure durations.
- There is a critical threshold level of overlap at which underlyingly dorsal KP codas abruptly become no longer identifiable as dorsal, but no such threshold exists among labial PK codas.

Interestingly, the perceptual interactions between labials and dorsals outlined above cannot be reduced to simple differences in magnitude or spatial duration of gestures, as has been hypothesized for perceptual assimilation involving coronals. As discussed in §2.3.2, labials are consistently shorter than dorsals within the stimuli used in these experiments, and yet they are still by far the more perceptually influential of the two. This insensitivity to low-level gestural parameters suggests that the results we have found are robust, and not an artifact of TADA’s gestural model.

5 Implications and further research goals

The results of the three experiments presented in this paper argue strongly for three conclusions: (i) minimally overlapped, unreleased dorsal KP codas can be more identifiable than highly overlapped, unreleased labial PK codas; (ii) Korean labial-to-dorsal place assimilation appears phonetically natural only if language-
specific levels of gestural overlap are taken into consideration; and (iii) labial-to-dorsal assimilation patterns against the prevailing typology of place assimilation systems predicted from phonetic principles in §1.2.2.

Firstly, our present results provide support for Son et al.’s (2007) hypothesis that high levels of gestural overlap in PK clusters are responsible for Korean place assimilation targeting labials but not dorsals. This hypothesis, though, has somewhat contradictory consequences for Korean when we consider the relationship between typological generalizations and phonetic facts. Language-Internally, the differential overlap settings in Korean PK and KP clusters make assimilation appear to indeed target the most “natural” codas, namely the highly-overlapped, perceptually-masked labials in PK clusters. But cross-linguistically, we expect assimilation systems to more readily target those codas that undergo greater perceptual changes under high gestural overlap, such as coronals. Our perceptual findings point to dorsals in KP clusters as being such codas: not only are they perceptually weak independent of overlap, but they also undergo drastic assimilatory perceptual changes when overlap levels are increased. Thus, despite a tight fit between phonetic and phonological facts within Korean, labial-to-dorsal place assimilation still appears to be typologically and phonetically unnatural.

Does Korean represent a fatal counterexample to the generalization that place assimilation systems can be predicted from the effects of gestural overlap on coda POA perception? I would argue that, to the contrary, Korean sheds a great deal of light on the true relationship between (mis)perception and place assimilation, which we have found to be somewhat less straightforward than previously thought (see Winters (2001) and Kochetov & So (2007) for similar sentiments). A crucial element in Korean place assimilation is its cluster-specific gestural overlap settings. Our perceptual explanation therefore has to involve a theory of how gestural overlap settings are determined in clusters.

5.1 Gestural overlap optimization

5.1.1 Parallel transmission and language specific overlap settings

Just as perception and place assimilation have long been linked in the minds of phonologists, so too has a long research tradition proposed a link between perception and gestural overlap settings (cf. Liberman et al. 1967, Mattingly 1981, Silverman 1995, Byrd 1996, Wright 1996, Chitoran et al. 2002). Liberman et al. (1967) pointed out that gestural overlap can have both advantages and disadvantages for the efficacy of speech communication. Overlapping articulations have the potential to quickly transmit information about multiple linguistic categories at once (the state of so-called “parallel transmission”), but for certain segment types high overlap leads to an obscuring of category information, as we have found in this study. Mattingly (1981) argued that parallel transmission is an organizing principle of linguistic systems, which tend to use the most overlapped articulations possible that are still able to achieve a reliable rate of information recovery. His example of this principle in action is the sonority hierarchy in syllable structure, which favors clusters at syllable margins that increase in articulatory aperture towards the nucleus. This ordering of apertures
allows multiple gestural movements to be recoverable simultaneously, leading to high efficiency in information flow.

More recently, Chitoran et al. (2002) have investigated a case of the parallel transmission principle with significant parallels to Korean, namely consonant cluster organization in Georgian. Noting the asymmetry in C1 burst production among C1:C2 stop clusters in front-to-back vs. back-to-front order, they find that front-to-back stop clusters are far more overlapped word-initially in Georgian than back-to-front clusters are. In word-initial position, the primary cue to C1’s presence in a two-stop cluster is its release burst, which can only be produced under very low overlap in back-to-front clusters. However, low overlap is unnecessary and inefficient in front-to-back clusters, which can be highly overlapped and still recoverable. Hence, overlap levels are balanced separately to achieve the best combination of efficiency and recoverability in the two cluster types.

5.1.2 Systemic non-optimality in Korean
Son et al. (2007) and Son (2008) note a striking similarity between the gestural overlap settings of Korean and Georgian: in both, PK clusters are far more overlapped than KP clusters. Is it possible that a drive for parallel transmission is at work in determining Korean overlap settings? I argue that it cannot be, at least in a strict synchronic sense. The reason is quite simply that dorsal KP codas are not overlapped enough to be efficiently transmitted, and labial PK codas are too overlapped to be recoverably transmitted – that is, efficiency and recoverability are not optimally balanced on the systemic level. To see why, it is helpful to consider what competitor overlap/assimilation systems might look like. For instance, imagine a language with a gradient assimilation system targeting KP instead of PK, which could do much better than Korean in terms of overlap. High overlap levels could be achieved in both clusters (though not quite as high as Korean’s current PK settings), but only KP would suffer perceptually for it. System-wide efficiency would exceed that of Korean, and non-assimilating PK could be more recoverable than Korean’s current non-assimilating KP. Synchronically, this overlap/assimilation system outperforms Korean in terms of both recoverability and efficiency at the same time. To borrow a poignant term from Optimality Theory (Prince & Smolensky 1993/2004), such a system harmonically bounds the Korean system relative to recoverability and efficiency.

Thus, Korean’s gestural overlap settings themselves are functionally unnatural, failing to take advantage of the great perceptual stability of PK clusters. Note too that PK clusters have been shown here to be doubly stable: not only does high gestural overlap not affect PK coda release bursts, as Chitoran et al. (2002) found, but our study has shown overlap to not greatly affect labial VC formant transitions either. What could account for Korean’s apparent misuse of articulatory and perceptual resources? I propose that, as in other cases of unnatural, idiosyncratic, or more generally language-specific grammatical patterns, the answer can be found in diachrony. That is, while Korean appears to mismanage phonetic difficulty at the present stage of its history, I claim it got to where it is now through a series of gradual, phonetically motivated and usage-based changes that led PK
and KP clusters to evolve as separate systems, without the benefit of teleological systemic oversight. I now briefly flesh out my proposal.

5.2 Paths of sound change for overlap evolution

One major difference found in this study between how PK and KP clusters respond to changes in gestural overlap is the source of assimilatory misperceptions in the two. Dorsal codas in KP clusters undergo drastic changes in the two primary coda POA cues, namely release bursts and VC formant transitions. Labial PK codas, on the other hand, only seem to undergo perceptual assimilation once closure durations are short enough to emulate those of singleton stops; their VC formant transitions and release bursts are relatively stable across all levels of overlap. These distinct types of overlap-induced perceptual change suggest the possibility of distinct paths of sound change.

5.2.1 Exemplar-based phonology

To cash out the idea that differences in how articulatory variation is perceived can lead to diverging paths of sound change for PK and KP clusters, I propose to take advantage of recent insights from the domain of the evolutionary exemplar theory of phonology (Pierrehumbert 2001, Hume & Johnson 2001, Blevins 2004, Wedel 2004, 2007, Silverman 2006, Bybee 2007). In this connectionist framework, language learners construct phonological representations out of a lexicon of richly phonetically-detailed exemplars. When a sound or word is encountered, language users attempt to categorize it by comparing it to stored perceptual representations. Based on this comparison, the sound can either be incorporated as a new category exemplar if it activates a single category strongly enough, or it can be rejected if it is judged to be a bad exemplar of the category that the listener expects it to encode. Thus, perceptual categories directly encode the variation a user encounters in their speech environment. Speakers’ own productions are drawn from this pool of variation, which drives category entrenchment. Since successfully categorized exemplars are incorporated into a category’s representation, any low-level biases that are present in the experienced variation of those exemplars have the potential to shift the category over time, in the direction of the bias. Such biases in variation are hypothesized by the authors cited above to be the basis of sound change, and can have their root in perceptual, articulatory, or functional facts of language use.

While a goodly amount of perceptual variation is tolerated by learners in this model, exemplars that stray too far from the category mean (or too close to another category’s boundaries) are dismissed and go uncategorized. Uncategorized variation ultimately exerts no influence on the category representation, and hence cannot be the basis of sound changes. It is this difference between variation that is successfully categorized and variation that goes uncategorized that I propose to be responsible for the divergent paths of sound change in Korean stop clusters.

---

40 Wedel (2004, 2007) takes advantage of categorical perception effects (specifically, acquired similarity) to explain why some variation is tolerated in categorization: exemplars that are sufficiently close to a perceptual category’s mean value are perceived as being identical to that mean.
5.2.2 Perceptual consequences of overlap variation for categorization in Korean

According to the model just outlined, the identification of production biases is a necessary first step in determining potential avenues of sound change. In the diachrony of place assimilation, the standardly assumed direction of sound change is that of gradually increasing inter-consonantal overlap. A plausible production bias that could be responsible for this trend is speakers’ articulatory accommodation to the demands of fast speech, the consequences of which become re-appropriated by new learners as part of their own normal speech.

Biased production variation is not by itself sufficient for sound change though, since the variation must also be successfully categorized based on perception. Articulatory variation that results in small perceptual changes in a category has the potential to be fed forward and telescoped among successive generations of learners. However, large perceptual changes are hypothesized to be rejected by learners, along with the articulatory changes that bring them about. Such variation is not fed forward to subsequent generations, and is not the basis for telescoping sound change.

Assuming for the sake of argument that a variation bias towards increased gestural overlap has been at work in the history of Korean, we must pose the question of how its effects are perceived at various stages in the language. The answer is going to vary from cluster to cluster, since different consonant clusters are perceptually altered by small overlap variations in different ways, as we have found in the present study. Starting with the state of the language immediately after the loss of regularly released coda stops, we can track the evolution of PK and KP separately. My proposals for the evolution of each cluster are as follows.

Among PK clusters, I submit that the high perceptual uniformity in labial POA cues found at all levels of overlap allows fast-speech tokens, produced with somewhat higher degrees of overlap than the mean, to be regularly categorized and stored as standard PK exemplars. The fast-speech exemplars thus work to pull the overall category mean in the direction of higher overlap. New generations of speakers whose own standard PK categories incorporate this high overlap will themselves produce even higher-overlap tokens when faced with the demands of fast speech, further driving the trend. All the while, the main perceptual difference between normal and fast-speech PK tokens will be in closure durations – burst and formant transition cues will remain largely uniform. Eventually a language state will be reached where the standard closure duration of the PK category mean is close to that of a singleton stop. Some of the new fast speech tokens at this state will have literally singleton-like closure durations, but will still be able to correctly activate the PK category, given their similarity to the already singleton-[k]-like PK category mean. Crucially, at this stage, even some articulatorily singleton-[k] tokens will successfully activate the PK category, despite differences in VC formant transitions. Taking a result from Dorman et al. (1979), I claim that those VC formant differences, even if perceived, will tend to be thrown out or disregarded by listeners in favor of matching the highly similar duration cues between the PK mean and [k]. Listeners who indeed categorize the articulatorily singleton-[k] exemplars as PK tokens will now be prone to produce
them, resulting in stochastically-distributed, categorical, labial-to-dorsal place assimilation, just as we see in the Korean of today.\textsuperscript{41}

When we turn to KP clusters, the story looks quite different, though not at first. Fast-speech KP tokens will be produced with slightly higher gestural overlap, as always. In the early stages of our imagined history this greater degree of overlap will mainly affect closure durations, just as was the case with PK tokens. Given the similarity in VC formant transitions between the KP category mean and the higher-overlap KP tokens, the new fast-speech forms will be correctly categorized and incorporated into the KP category. As with PK, the overall trend will be towards gradually higher overlap levels. At a certain stage, though, the KP category will reach a standard mean level of overlap that is close to, but not at, that of the steep perceptual drop-off found in our three experiments. At this stage, new fast-speech tokens with slightly higher levels of overlap will sound radically different from the KP category mean. The perceptual divergence of these fast-speech KP tokens is rather different than for any of the PK tokens: KP closures are not nearly short enough to be singleton-like, so VC formant transitions are fully perceived and taken into consideration in KP categorization. I claim that the radical formant differences in the new fast-speech KP tokens will lead them to be rejected by listeners as speech errors and to go uncategorized. With no new fast-speech exemplars to pull the KP category mean in the direction of higher overlap, sound change grinds to a halt. The KP category stabilizes at a moderately low overlap level, and no assimilation system develops, despite the weakness of dorsal coda POA cues.

The final outcome of the two courses of articulatory evolution, each proceeding independently with its own internal motivation, is the counter-functional assimilation of labial codas, but not dorsal codas, in modern Korean. Crucially, perceptual considerations entered into the determination of place assimilation not through reference to a single systemic criterion of recoverability. Rather, perceptual “faithfulness” was judged separately for each category on the basis of the surface variation that the category displayed and the cues used to activate it. It is therefore faithfulness of allophones to other allophones that matters to this model of place assimilation, not of allophones to abstract, Platonic-ideal phonemes. Despite the opacity of phonetic motivation that this hypothesis implies for Korean, it is still plain to see that place assimilation is causally linked to facts of

\textsuperscript{41} It might be objected that, in my system, PK clusters in all unreleased-coda languages (such as English) should quickly evolve to a similar state of degenerate closure durations, leading labial assimilation processes to appear much more common than they actually do. I would like to speculate that in fact the course of evolution I describe for PK may be special to Korean, for a language-specific functional reason. In many languages, English included, closure duration is a primary cue to the singleton-stop/cluster distinction, and as such it resists synchronic variation, stifling language change towards higher overlap. Korean, however, has another highly salient cue for the presence of non-singleton stop complexes, namely post-obstruent tensification (Kim-Renaud 1991). If the application of this tensification rule is sufficient to cue the lexical presence of a non-singleton cluster in Korean, then closure duration may be diminished in importance as a cue in its own right. Korean speakers could allow it to vary more widely by speech rate, without the worry that stop clusters would be confused for singleton stops in fast speech. This extra variability in closure duration would then drive the language change that I propose.
perception and articulation as always, just as desired. Even though labial-to-dorsal place assimilation defies typological generalization, it does not defy phonetic causation.

5.2.3 Challenges for an evolutionary account of place assimilation

A theory of phonology with no systemic oversight, such as the one I propose here, has several challenges to meet before it can claim to be a viable alternative to synchronic-optimization theories of phonology. Some of these challenges are listed as follows.

- How are “standard” cases of place assimilation treated, such as those involving coronal or nasal codas? Can overlap evolution and exemplar-based category abstraction of the kind claimed here lead to normal, well-behaved systems without having to stipulate anything extra?
- Why are so many place assimilation systems not phonetically opaque? What makes the exceptional cases like Korean so exceptional?
- In the same vein, if each cluster is treated as its own evolutionary system, why is it that consonant clusters tend to look as uniform as they do within languages? Why have we only recently discovered systematic non-uniformities in the implementation of consonant clusters? Where does the appearance of systematic order come from?
- Can the non-optimality-driven theory of language change advocated here predict why parallel transmission appears to emerge as a principle of language structure?
- If language learners and users are as perceptive and attuned to fine-grained phonetic detail as we claim they are here, then how does categorical place assimilation ever arise? Why aren’t all place assimilation systems gradient?

Challenges of the above types are not unique to my proposal, but rather come as necessary baggage with any functionalist theory where order and structure are emergent instead of given. Since evolutionary accounts are by their nature very difficult to falsify, we must look to new ways of empirically supporting hypotheses about language change. It is becoming ever clearer, as proposals and phonetic description become more nuanced and numerically concrete, that predictive computational implementation is a realistic goal of functionalist research. Proof-of-concept models of evolutionary exemplar approaches to phonology have already been implemented, largely by the authors cited above in §5.2.1. But to my knowledge, none of them goes so far as to model phonetic cues and perceptual facts directly. The content of my proposals and their ability to address the challenges mentioned above, however, demand just such a model. Furnishing this is the next goal of my research.
6 Conclusion

In this paper I have presented the results of three perceptual experiments, whose goal is to determine how clusters of labial and dorsal stops are perceptually altered under increasing gestural overlap. Other studies of how gestural overlap affects consonant clusters have not investigated the interaction between labials and dorsals in great depth. The results show that labial stops have an overall greater ability to influence VC formant transitions under high levels of overlap than dorsals do, and consequently are more effective at inducing perceptual assimilation and resisting the same than dorsals. This perceptual asymmetry between labials and dorsals has interesting consequences for phonological systems in which labial codas are targeted for place assimilation, such as Korean. Following previous work by Son et al. (2007), I argue that language-specific gestural overlap settings are crucial in explaining Korean place assimilation, but that Korean’s overlap/assimilation system as a whole cannot be explained using synchronic phonetic grounding. That is, Korean remains phonetically “unnatural” once overlap settings and assimilation are considered together.

As an alternative, I propose that Korean’s overlap system evolved to be what it is today through phonetically optimizing changes that treated each cluster as a separate system. Using an evolutionary exemplar theory framework, I argue that perceptual differences in how clusters with dorsal vs. labial codas respond to variations in gestural overlap determined their separate paths in sound change. Clusters with labial codas undergo very little perceptual change as overlap increases, leading small overlap changes to be tolerated and to accrue into larger changes over time, and eventually resulting in widespread assimilatory misperceptions. On the other hand, clusters with dorsal codas experience large perceptual perturbations from small changes in overlap, leading such changes to be rejected by listeners and not fed forward to subsequent generations. In this way, dorsal codas remain non-overlapped and unassimilated while the perceptually superior labial codas evolve to be highly overlapped and assimilated.

While the evolutionary account argued for here is still only speculative, it has the advantage of preserving an explanatory connection between phonetic processes like misperception and typologies of phonological processes like place assimilation. In a purely optimality-driven, synchronic model of that connection, Korean represents a potentially fatal counterexample: a phonetically opaque place assimilation system. By abandoning the oversight of system-wide optimization, we are able to re-ground Korean labial-to-dorsal place assimilation as a phonetically natural process.
APPENDIX A.  SUBJECT RESULTS FOR EXPERIMENT I

Response accuracy for heterorganic sequences.

[Graphs showing accuracy for different subjects]
APPENDIX B. SUBJECT RESULTS FOR EXPERIMENT II

Response accuracy for truncated heterorganic sequences.

- Subject FDA
- Degree of overlap: 3 5 7 9 11 13
- Accuracy: 0.0 0.2 0.4 0.6 0.8 1.0
- Chance response rate: AP(KA), AK(PA)

- Subject HRM
- Degree of overlap: 3 5 7 9 11 13
- Accuracy: 0.0 0.2 0.4 0.6 0.8 1.0
- Chance response rate: AP(KA), AK(PA)

- Subject JAM
- Degree of overlap: 3 5 7 9 11 13
- Accuracy: 0.0 0.2 0.4 0.6 0.8 1.0
- Chance response rate: AP(KA), AK(PA)

- Subject JLB
- Degree of overlap: 3 5 7 9 11 13
- Accuracy: 0.0 0.2 0.4 0.6 0.8 1.0
- Chance response rate: AP(KA), AK(PA)

- Subject MAC
- Degree of overlap: 3 5 7 9 11 13
- Accuracy: 0.0 0.2 0.4 0.6 0.8 1.0
- Chance response rate: AP(KA), AK(PA)

- Subject NMW
- Degree of overlap: 3 5 7 9 11 13
- Accuracy: 0.0 0.2 0.4 0.6 0.8 1.0
- Chance response rate: AP(KA), AK(PA)

- Subject RRD
- Degree of overlap: 3 5 7 9 11 13
- Accuracy: 0.0 0.2 0.4 0.6 0.8 1.0
- Chance response rate: AP(KA), AK(PA)

- Subject SEN
- Degree of overlap: 3 5 7 9 11 13
- Accuracy: 0.0 0.2 0.4 0.6 0.8 1.0
- Chance response rate: AP(KA), AK(PA)

- Subject SOS
- Degree of overlap: 3 5 7 9 11 13
- Accuracy: 0.0 0.2 0.4 0.6 0.8 1.0
- Chance response rate: AP(KA), AK(PA)

- Subject TMK
- Degree of overlap: 3 5 7 9 11 13
- Accuracy: 0.0 0.2 0.4 0.6 0.8 1.0
- Chance response rate: AP(KA), AK(PA)

- Subject ZL1
- Degree of overlap: 3 5 7 9 11 13
- Accuracy: 0.0 0.2 0.4 0.6 0.8 1.0
- Chance response rate: AP(KA), AK(PA)
APPENDIX C. SUBJECT RESULTS FOR EXPERIMENT III

d’ for within-sequence comparisons: 2 frames of overlap distance.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Degrees of overlap compared</th>
<th>Sensitivity (d’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC</td>
<td>3-5 5-7 7-9 9-11 11-13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>AML</td>
<td>3-5 5-7 7-9 9-11 11-13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>DRS</td>
<td>3-5 5-7 7-9 9-11 11-13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>GG</td>
<td>3-5 5-7 7-9 9-11 11-13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>JLR</td>
<td>3-5 5-7 7-9 9-11 11-13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>MGC</td>
<td>3-5 5-7 7-9 9-11 11-13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>MNT</td>
<td>3-5 5-7 7-9 9-11 11-13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>NAK</td>
<td>3-5 5-7 7-9 9-11 11-13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>NCE</td>
<td>3-5 5-7 7-9 9-11 11-13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>RR</td>
<td>3-5 5-7 7-9 9-11 11-13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>SDC</td>
<td>3-5 5-7 7-9 9-11 11-13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>THH</td>
<td>3-5 5-7 7-9 9-11 11-13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>TLN</td>
<td>3-5 5-7 7-9 9-11 11-13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>ZL</td>
<td>3-5 5-7 7-9 9-11 11-13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
</tbody>
</table>
APPENDIX C. (CONT’D)

d’ for within-sequence comparisons: 3 frames of overlap distance

Subject AC

Subject AML

Subject DRS

Subject GG

Subject JLR

Subject MGC

Subject MNT

Subject NAK

Subject NCE

Subject RR

Subject SDC

Subject THH

Subject TLN

Subject ZL
APPENDIX C. (CONT’D)

d’ for between-sequence comparisons: comparisons to AP(PA)
**APPENDIX C. (CONT’D)**

\(d'\) for between-sequence comparisons: comparisons to AK(KA)

<table>
<thead>
<tr>
<th>Subject</th>
<th>Degree of overlap</th>
<th>Sensitivity (d')</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC</td>
<td>3 5 7 9 11 13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>AML</td>
<td>3 5 7 9 11 13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>DRS</td>
<td>3 5 7 9 11 13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>GG</td>
<td>3 5 7 9 11 13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>JLR</td>
<td>3 5 7 9 11 13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>MGC</td>
<td>3 5 7 9 11 13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>MNT</td>
<td>3 5 7 9 11 13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>NAK</td>
<td>3 5 7 9 11 13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>NCE</td>
<td>3 5 7 9 11 13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>RR</td>
<td>3 5 7 9 11 13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>SDC</td>
<td>3 5 7 9 11 13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>THH</td>
<td>3 5 7 9 11 13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>TLN</td>
<td>3 5 7 9 11 13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td>ZL</td>
<td>3 5 7 9 11 13</td>
<td>0 1 2 3 4 5 6</td>
</tr>
</tbody>
</table>

UCSC Linguistics Research Center
APPENDIX D. ACOUSTIC DISTANCE MEASURES IN EXPERIMENT III STIMULI

Time-integrated frequency differences in F1, by stimulus comparison type

2 frames of overlap distance

![Graph showing time-integrated frequency differences in F1 for 2 frames of overlap distance.]

Degree of overlap for X-A in AXA comparison

3 frames of overlap distance

![Graph showing time-integrated frequency differences in F1 for 3 frames of overlap distance.]

Degree of overlap for X-A in AXA comparison

Comparisons with AP(PA)

Degree of overlap for X in AXA comparison

Comparisons with AK(KA)

Degree of overlap for X in AXA comparison

Time-integrated frequency differences in F2, by stimulus comparison type

2 frames of overlap distance

![Graph showing time-integrated frequency differences in F2 for 2 frames of overlap distance.]

Degree of overlap for X-A in AXA comparison

3 frames of overlap distance

![Graph showing time-integrated frequency differences in F2 for 3 frames of overlap distance.]

Degree of overlap for X-A in AXA comparison

Comparisons with AP(PA)

Degree of overlap for X in AXA comparison

Comparisons with AK(KA)

Degree of overlap for X in AXA comparison
APPENDIX D. (CONT’D)

Time-integrated frequency differences in F3, by stimulus comparison type

Differences in overall duration between compared stimuli, by comparison type
REFERENCES


