Perceptual Distortions in the Adaptation of English Consonant Clusters: Syllable Structure or Consonantal Contact Constraints?

Barış Kabak[1], **William J. Idsardi**[2]

[1] *University of Konstanz*
[2] *University of Maryland*

Key words

L2 speech perception

perceptual epenthesis

phonotactics

syllable structure

Abstract

We present the results from an experiment that tests the perception of English consonantal sequences by Korean speakers and we confirm that perceptual epenthesis in a second languge (L2) arises from syllable structure restrictions of the first language (L1), rather than linear co-occurence restrictions. Our study replicates and extends Dupoux, Kakehi, Hirose, Pallier, & Mehler's (1999) results that suggested that listeners perceive epenthetic vowels within consonantal sequences that violate the phonotactics of their L1. Korean employs at least two kinds of phonotactic restrictions: (i) syllable structure restrictions that prohibit the occurence of certain consonants in coda position (e.g., *[c.], *[g.]), while allowing others (e.g., [k.], [l.]), and (ii) consonantal contact restrictions that ban the co-occurrence of certain heterosyllabic consonants (e.g., *[k.m]; *[l.n]) due to various phonological processes that repair such sequences on the surface (i.e., /k.m/→[ŋ.m]; /l.n/→[l.l]). The results suggest that Korean syllable structure restrictions, rather than consonantal contact restrictions, result in the perception of epenthetic vowels. Furthermore, the frequency of co-occurrence fails to explain the epenthesis effects in the percept of consonant clusters employed in the present study. We address questions regarding the interaction between speech perception and phonology and test the validity of Steriade's (2001 a,b) Perceptual-Mapping (P-Map) hypothesis for the Korean sonorant assimilation processes. Our results indicate that Steriade's hypothesis makes incorrect predictions about Korean phonology and that speech perception is not isomorphic to speech production.

Acknowledgments: This article is an extended version of an article that has appeared in the proceedings of the Berkeley Linguistics Society (Kabak & Idsardi, 2003), which was based on Kabak (2003), the doctoral dissertation of the first author. We would like to thank the audiences at the 29th Annual Meeting of the Berkeley Linguistics Society for comments and suggestions. We are also indebted to Sharon Peperkamp, John Kingston, Anne Cutler, Donca Steriade, Aditi Lahiri, Jaye Padget, Benjamin Bruening, Irene Vogel, Juliette Blevins, and Colin Phillips for important criticism and suggestions.

Address for correspondence. Barış Kabak, Department of Linguistics, University of Konstanz, Fach D 174, 78457 Konstanz, Germany; phone: +49-7531-884758; fax: +49-7531-884160; e-mail: <baris.kabak@uni-konstanz.de>.

1 Introduction

This article investigates the perceptual adaptation of English consonant clusters by Korean second language (L2) speakers of English. Our point of departure is recent findings on perceptual epenthesis by Dupoux, Kakehi, Hirose, Pallier, & Mehler (1999) who demonstrate that Japanese speakers hear epenthetic vowels in consonantal sequences that violate the phonotactic rules of their native language. Our primary aim here is to replicate and extend these findings by unconfounding the contribution of potential syllable structure constraints from purely sequential consonantal contact restrictions. Employing Korean, a language in which the two factors can be separated, we demonstrate that percepual epenthesis in L2 arises from syllable structure violations, rather than linear consonantal contact violations in the L1. Our study also shows that native phonological processes that would normally apply to resolve certain illicit sequences do not play a role in Korean listeners' perceptual adaptation of these sequences. Furthermore, explanations based on frequency fail to shed light on epenthesis effects in the percept of consonantal sequences. We evaluate the consequences of these findings in relation to the following questions: Can speech perception explain phonological processes? Does speech perception mirror speech production? In particular, we test our findings against Steriade's (2001 a,b) Perceptual-Mapping (P-Map) hypothesis and argue that the hypothesis makes incorrect predictions about Korean phonology. Accordingly, we suggest that speech perception is not always isomorphic to speech production. Our findings are consistent with the L2 literature in finding both commonalities and discrepencies between L1 and L2 phonology (e.g., Dulay, Burt, & Krashen, 1982), and also support recent psycholinguistic theories of loan adaptations that emphasize the role of perception (e.g., Peperkamp & Dupoux, 2003; Peperkamp, in press).

1.1 Phonotactics and speech perception

It has long been established that listeners' knowledge of L1 phoneme contrasts has an influence on the perception of non-native contrasts (e.g., Sapir, 1933/1949). L2 research has investigated similarities and differences between L1 and L2 sound systems as a potential predictor for the relative difficulty and ease in the acquisition of non-native sounds and sound contrasts (e.g., Flege, 1980, 1987; Major, 1987; Major, 2001, for a review). Models of cross-linguistic speech perception have mostly been along the lines of "equivalence classification" of L1 and L2 sounds based on their acoustic and phonetic proximity to one another (e.g., Best, 1995; Flege, 1995). For example, the Speech Learning Model (Flege, 1995) asserts that L1 and L2 sounds are perceptually related to one another at a position-sensitive phonetic level, and learners will have a hard time acquiring "similar" sounds since they perceive and classify them as equivalent to an already existing sound or sets of sounds in the L1. "Different" sounds, however, are expected to be easier to acquire since the perceptibility of the difference will enable the learner to build a new phonemic category. Likewise, Best's Perceptual Assimilation model claims that non-native sounds "tend to be perceived according to their similarities to, and discrepancies from, the native segmental constellations that are in closest proximity to them in native phonological space" (Best, 1995, p. 193). Accordingly, upon encountering non-native segments, whose gestural elements or intergestural phasing do not match precisely any native constellations, listeners are expected to detect gestural similarities to native phonemes as well as discrepancies from the gestural properties of native constellations. The degree to which a non-native contrast can be assimilated to native categories by the listener determines how well he/she will be able to perceive that non-native contrast. In this model, non-native sounds are either (1) assimilated to a native category, (2) assimilated to an uncategorizable (but yet within the native phonological space) speech sound, or (c) not assimilated to speech. For example, if two L2 categories are assimilated to two different L1 categories, discrimination is expected to be successful. If, on the other hand, two non-native categories are assimilated to a single L1 category, discrimination is expected to yield moderate or very poor discrimination depending on how the two L2 sounds differ from the single L1 category.

Languages, however, may not only differ with respect to their phoneme repertoires, but also with respect to the phonotactic rules and constraints that govern the distribution of sounds occurring in these inventories. For instance, while both Turkish and English contain /s/ and /t/, the combination of the two in the onset position is banned in Turkish (*/.st/), but not in English, where there are several words that begin with /st/ clusters (e.g., *street*, *stamp*, *stock*, etc.). Several researchers have focused on response biases in favor of perceiving ambiguous sound combinations as legal sequences, a pattern which is assumed to reflect the influence of phonotactic knowledge in perception (e.g., Hallé, Segui, Frauenfelder, & Meunier, 1998; Massaro & Cohen, 1983; Moreton, 2002; Pitt, 1998). Other studies have looked at the origin of phonotactic knowledge, such as frequency and lexical effects (e.g., Luce & Pissoni, 1998; McClelland & Elman, 1986; Norris, 1994; Pitt & McQueen, 1998; Vitevitch & Luce, 1998). Over all, these studies demonstrate that listeners' information about the legality and the probability of phonotactic patterns influence the processing of spoken stimuli. Phonotactic knowledge in infant language processing has also been extensively investigated. The studies suggest that the knowledge about L1 phonotactic regularities emerge very early on in the course of acquisition (e.g., Jusczyk, Luce, & Charles-Luce, 1994), and seem to bias infants' perception to prefer listening to words that accord with the phonotactics of their L1 (Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993).

Since L1 phonotactic regularities have been shown to influence infants' perceptual abilities, will they also play a role in L2? There are several instances in which we can observe the impact of L1 phonotactic regularities on the realization of non-native speech. For instance, speakers of Korean and Japanese are known to insert epenthetic vowels when they pronounce loan words involving sequences of segments that do not fit the syllable structure of their native languages:[1]

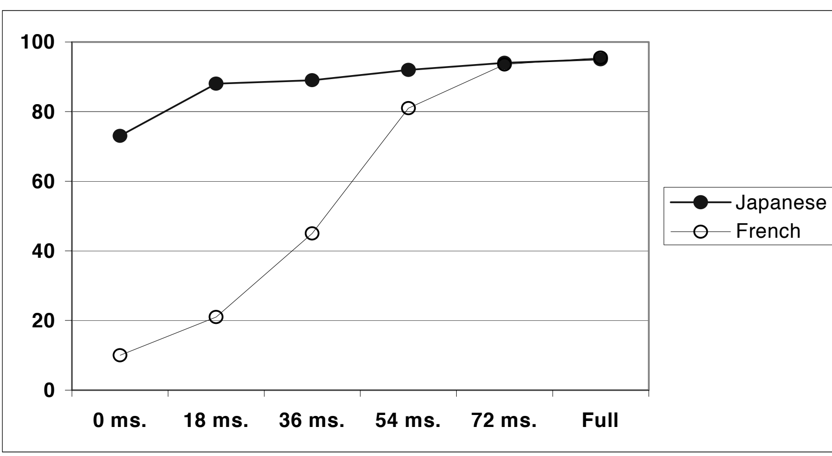
(1)	a.	[a.i.sɯ.k^{h}ɯ.ɾim]	'ice cream'	(Korean)
	b.	[k^{h}ɯ.ɾi.sɯ.ma.sɯ]	'Christmas'	(Korean)
	c.	[ma.kɯ.do.na.ɾɯ.do]	'Mac Donald'	(Japanese)
	d.	[sɯ.to.ɾa.i.kɯ]	'strike'	(Japanese)

[1] These transcriptions are given in the IPA except where noted; [ɯ] is a high back unrounded vowel, [ɾ] is an alveolar tap.

Are the epenthetic vowels inserted in production, or in perception? In a series of behavioral experiments, Dupoux and colleagues (Dehaene-Lambertz, Dupoux, & Gout, 2000; Dupoux, Kakehi, Hirose, Pallier, & Mehler, 1999; Dupoux, Pallier, Kakehi, & Mehler, 2001) compared Japanese listeners with French listeners in their perception of consonant clusters. For instance, Dupoux et al. (1999), in an off-line phoneme detection task (Experiment 1), used a series of six items created from naturally produced nonce words from a Japanese speaker (e.g., [abuno], [akumo], [ebuzo], [egudo], etc.) by gradually reducing the duration of the vowel [u] down to zero milliseconds. The participants were asked to respond whether each item they heard contained the sound [u]. Japanese listeners, unlike French listeners, overwhelmingly judged that the vowel was present at all levels of vowel length. Strikingly, this was the case 70% of the time even when the vowel had been completely removed (i.e., the zero ms condition). The French participants, on the other hand, judged that the vowel was absent in the no-vowel condition about 90% of the time and that a vowel was present only in 50% of the intermediate cases (Fig. 1).

Figure 1

Percept [u] vowel judgments as a function of vowel duration (adapted from Dupoux et al., 1999)



These results were confirmed in other experiments, including a replication using stimuli created from a French speaker, which have led Dupoux and colleagues to conclude that the influence of native language phonotactics can be so robust that listeners "invent" illusory vowels to accommodate illicit sequences of segments in their L1. Follow-up studies, again on Japanese but employing different experimental paradigms such as a lexical decision task (Dupoux et al., 2001) and event-related potentials (Dehaene-Lambertz et al., 2000) have further confirmed perceptual epenthesis with Japanese participants.

Dupoux and colleagues' findings have yielded crucial implications for the potential role of perception in explaining production errors such as epenthesis and consequently also for our understanding of the nature of phonological representations.

The researchers, however, have left the question open as to whether perceptual epenthesis stems from the fact that Japanese speakers never hear certain consonants in the coda position in their language or whether Japanese syllable structure is predominantly CV. That is, it is not clear from Japanese whether perceptual epenthesis arises due to restrictions on which consonants can co-occur in a sequence or due to syllable structure restrictions. Japanese syllables are predominantly CV, as Japanese licences very few coda consonants, and then only under greatly restricted circumstances, and therefore displays a paucity of coda-onset clusters generally. In particular, a consonant can only be licensed in the coda position if it is the first member of a geminate (2a,b), or if it is a nasal homorganic to the following consonant (2c,d) (Itô, 1986; 1989). Coda consonants (other than the mora nasal) not carrying either of these two properties cannot occur in Japanese, as can be seen in the hypothetical examples given in (3) (from Itô, 1989).

(2) a. kap.pa 'a legendary being'
 b. gak.koo 'school'
 c. tom.bo 'dragonfly'
 d. kaŋ.gae 'thought'

(3) a. * kap.ta
 b. * tog.ba
 c. * pa.kat

Closer inspection of the stimulus items used in Dupoux et al. (1999) shows that the items contained illicit coda consonants (e.g., [eb.zo], [iʃ.to], [eg.do], etc.). The study, thus, confounds whether the perceptual epenthesis induced in the percept of words such as [ebzo] is due to: (1) consonantal contact restrictions (i.e., the sequence [b.z] is impossible), or (2) syllable structure (coda) restrictions (i.e., [b] cannot be an independent coda). Since Dupoux and colleagues' primary aim was to document the perceptual vowel epenthesis per se, the restrictedness of Japanese phonology does not interfere with their findings and conclusions. It is important, however, to tease apart consonantal contact restrictions from coda restrictions since this has serious consequences for phonological theories explaining phonotactic patterns. In particular, while most theories refer to a prosodic domain, namely the syllable, to explain consonantal phonotactics, there are also those that employ syllable-independent, string-based, and linear statements (e.g., Blevins, 2002; Dziubalska-Kolaczyk, 1994; Steriade, 1999, 2001a). The theories that eschew references to the syllable or any other prosodic structure would explain the findings as a product of illegal consonant contact and would reject explanations in terms of syllable structure violations as being unnecessary. The present study sucessfully teases apart the two competing explanations by employing Korean, which offers a much more interesting array of consonant clusters. As noted above, vowel epenthesis in Korean arises to avoid consonant clusters in loan words and as well as in the context of L2 speech. As such, vowel epenthesis differs from other phonological processes because the latter processes are part of the native phonological grammar of Korean. Accordingly, our secondary aim is to test whether native phonological processes that would ordinarily resolve certain illicit consonantal sequences in Korean production grammar may also constitute as potential repair strategies in perception. Here, we particularly examine those cases where native phonological rules are most likely to apply as perceptual repair strategies. Based on our observations about Korean L2 English, as well as a pool of loan word evidence available to us, we restrict our focus only to sonorant assimilation processes in Korean, namely nasalization and lateralization.

1.2 English and Korean phonotactics

Phonotactics, in its most general sense, refers to the restrictions that govern sound sequences in a language. The coda or onset position of a syllable can be used to anchor particular phonological contexts where restrictions on consonants and their combinations may be observed. For instance, three-member consonantal onset clusters in English must start with /s/ (e.g., *spread, street, square, splash*, etc.), but codas do not obey the same requirement, nor the reverse requirement (e.g., *burst*). Likewise, English has some consonants that only appear in onset position (e.g., /h/) as well as consonants that only appear in coda position (e.g., /ŋ/). In Korean pronunciations, the alveo-palatal affricates (/c, c', c^h/[2]) can only occur in syllable onsets.

Korean and English have different restrictions on heterosyllabic consonantal contacts (i.e., [$C_1.C_2$]). Ewen and van der Hulst (2001) offer a clear discussion of English medial clusters, demonstrating the utility of the syllable as an important domain over which to explain phonotactic regularities. Ewen and van der Hulst propose that part of the phonological knowledge of a native speaker involves the specification of which consonant clusters are ill-formed in English, giving the examples in (4) (p. 123).

(4)	*initial*	*medial*	*final*
	*km-	*-pkm-	*-pk
	*mr-	*-kmr-	*-km
	*mw-	*-tnw-	*-tn

They note however that a fairly obvious redundancy arises in the above clusters: The constraints on medial clusters in English are not independent of those on initial and final clusters. Medial sequences must consist of a valid coda followed by a valid onset. Any other medial cluster is illicit. Other than this, English is quite profligate and seems to place no constraint on coda-onset contacts (see Lamontagne (1993) for a detailed discussion of the restrictions on English consonant clusters). It should be noted, however, that not all consonantal combinations are observed with equal frequency in English. For instance, while there are monomorphemic words with -kt- (e.g., *sector*, *doctor*, *October*, etc.), the reverse sequence -tk- is very difficult to find in

[2] Following the Korean linguistics tradition, we will use /c/ to refer to Korean voiceless coronal affricate consonant. Accordingly, no diacritic marker on a consonant indicates a plain consonant, as opposed to tense (e.g., /c'/) and aspirated (e.g., /c^h/). The symbol [j] represents the voiced allophone of the plain alveo-palatal affricate consonant (i.e., /c/). We use [y] to represent the palatal glide.

the language (except for *catkin* (a type of flower), or proper names such as *Atkins* or *Ratko*). Despite its rarity, English speakers readily accept and produce pronunciations of loan words containing this cluster (e.g., *Kamchatka*).

In Korean, inflected or derived words provide potential contexts for heterosyllabic consonantal contacts, as consonants belonging to different morphemes come into contact at morpheme junctures. Korean employs a number of phonological processes that affect such heterosyllabic consonant clusters. For instance, due to the process of nasalization, a sequence of an oral consonant followed by a nasal consonant (C+N) never occurs (e.g., /puək^h/ + /mun/ → [pu.əŋ.mun] 'kitchen door'; /os/ + /noŋ/ → [on.noŋ] 'clothes chest')[3] except where there is an intervening Intonational Phrase (IP) boundary.[4] It should be noted that no IP boundary occurs inside inflected or derived words. In English, however, although some word-internal consonant clusters such as /gt/ are nearly impossible to find, compounds with such clusters do exist without undergoing any phonological alterations (e.g., *pigtail, ragtime*). Moreover, many unusual heterosyllabic combinations of consonants in loan words, novel words, and acronyms can easily be incorporated into English without modification and are pronounced without any difficulty by English speakers (e.g., [g.d]: *Ma**gd**eburg*; [ŋ.y]: ***Pyongy**ang*; [t.k]: *do**t c**om*, etc.). Since English seems to place virtually no additional restrictions on the type of *coda-onset* clusters, in the following we will focus only on Korean, which has instructive phonotactic restrictions to separate consonantal contact restrictions from syllable structure restrictions.

We now turn to relevant properties of the phonological system of Korean that are crucial for our purposes in this study. Korean has 19 consonants and 10 vowels (Tables 1 and 2, adapted from Sohn, 1999, p. 153, 156).

Korean has a three-way contrast for oral stops and affricates: plain (C), aspirated (C^h) and tense (C'). The tense consonants are phonetically both long and pronounced with a constricted glottis. Korean syllables are maximally CGVC. As in Japanese, the onset and coda cannot branch in Korean. The glide (G) in CGV is part of the nucleus. Korean restricts the occurrence of consonants in certain syllabic positions. For instance, only seven among 19 consonantal phonemes (i.e., [p, t, k, m, n, ŋ, l]) can be pronounced in codas. The other consonants undergo various processes of neutralization when they end up in coda position. When coda consonants are in

3 Korean examples in this article are adapted from Sohn (1994; 1999) unless otherwise noted.

4 The rules that apply across morphemes (see the next section) such as nasalization and lateralization have also applied historically to what are now unanalyzable combinations of morphemes (see e.g. Martin (1992, p. 52), who cites a 15th century form /punnon/ <*/put^h-non/ 'igniting'). Since it is simply possible to lexically represent the now morpheme-internal sequence with two nasals, we cannot therefore find direct evidence for morpheme-internal applications of nasalization, as morphemes that involve CN sequences have been diachronically leveled out in favor of NN sequences. Similar considerations hold of borrowed words that have orthographic treatments in Hangul with 'nl' sequences, such as the Korean rendition of foreign names, for example *Marilyn Monroe*, orthographically 'monlo' but pronounced [monno] (Lee & Ramsey, 2000, p. 71). Since there is no evidence that nasalization does or does not apply morpheme-internally in Korean, in the absence of nonderived environment blocking effects, the simplest conclusion is that the rule applies both within morphemes as well as across morpheme boundaries.

Table 1

Consonantal inventory of Korean

		Labial	Alveolar	Alveo-palatal	Velar	Glottal
Plosives	Plain	p	t	c	k	
	Tense	p'	t'	c'	k'	
	Aspirated	p^h	t^h	c^h	k^h	
Fricative	Plain		s			h
	Tense		s'			
Nasal		m	n		ŋ	
Liquid			l			

Table 2

Vowel inventory of Korean

	-Back		+Back	
	-Round	+Round	-Round	+Round
High	i	ü	ɯ	u
Mid	e	ø	ə	o
Low	ɛ		a	

contact with other consonants at morpheme junctures, they undergo various assimilatory processes. Even consonants that cannot surface in coda position interact with adjacent consonants (e.g., /h/ + /t/→ [t^h]; /coh/ + /ta/→ [cot^ha] 'is good'). While assimilatory processes may alter the segmental realization of a given heterosyllabic consonant cluster on the surface, the consonants that form the contact can independently surface if they occur elsewhere. For instance, *[k.m] and *[l.n] are illicit but [k.t] and [l.t] are licit. Such illicit codas and consonantal contacts are repaired by various phonological rules. Thus Korean provides enough variety to distinguish clusters which are illicit because of the contact between the consonants versus those that are illicit because of the coda consonant alone. In the following, we will provide only those processes that are relevant for the present study.

First, strident consonants such as [c], [c^h], and [s] neutralize in coda position to the unreleased stop [t]. For instance, morphophonemic forms such as /nac/ 'daytime', /nac^h/ 'face' and /nas/ 'sickle' become homophonous when they are pronounced in isolation (i.e., [nat]). Second, Korean has a nasalization rule that turns stops into nasals before nasal segments (e.g., /k.m/→ [ŋ.m]: /hak+mun/→ [haŋ.mun] 'learning'; /p.m/ → [m.m]: /cip+mun/→ [cim.mun] 'house gate'). Third, a process of lateralization affects nasal sounds after lateral sounds (/l.n/→ [ll]: /tal+nala/ → [tal.la.ɾa] 'moon country'). Finally, voicing in Korean is predictable. Plain consonants become voiced between sonorants (e.g., /pa+po/→ [pa.bo] 'idiot'). It should be noted that these processes are very general processes of Korean *without* exceptions. They not only apply within words but also across word boundaries, although not across Intonational

Phrase boundaries (see Martin (1992, p. 30) and Kabak (2003) for an exhaustive list of possible [C_1.C_2] contacts and their surface realization). Finally, it should be noted that the Korean epenthetic vowels are [i] after palatal consonants and [ɯ] in other contexts (e.g., [a.i.sɯ.kɯ.ɾim] 'ice cream'; [sɯ.pʰən.ji] 'sponge').

1.3
The present study

A closer inspection of Korean phonotactic patterns reveals that there are at least three reasons why a word in the form of [VC_1.C_2V] can be illicit. First, [C_1.C_2] could induce a consonantal contact violation. A cluster presenting a bad syllable contact is defined as one in which C_1 is a licit coda, C_2 a licit onset, but the combination C_1.C_2 is illicit. For instance, as explained above, words containing *[k.m] or *[l.n] sequences would be required to undergo nasalization and lateralization by the Korean production grammar. Second, [C_1] could be an illicit coda (e.g., *[c.], *[kʰ.], *[h.], *[ɾ.], etc.). Third, [C_2] could be an illicit onset (e.g., *[.l], *[.ŋ], etc.). The first and the third reasons differ from the second in being sensitive to the nature of C_2. The present study is concerned only with the first two factors (see, however, Kabak (2003) for additional discussion of consonants disallowed in onsets). Of course, illicit codas or onsets necessarily preclude clusters with these elements in those positions, hence our definition of illicit contact as involving attested codas and onsets.

Given the two possible factors that induce an illicit sequence of consonants in Korean, namely illegal consonantal contact and illegal coda, we wish to investigate (1) whether all types of illicit sequences of consonants cause perceptual epenthesis (e.g., whether words with [k.m] sequences are confusable with words with [kɯm] instead), and (2) whether contact violations (e.g., *[k.m], *[l.n]) can be perceptually altered if there is a native phonological process available to fix the violation in the Korean production grammar. With respect to the first question, we separate those violations that are caused by illicit consonantal combinations (*C_1.C_2) (e.g., [k.t]) from those that are caused by an illicit coda (*C_1.) (e.g., [c.t]). As for the second question, we want to test whether words with [k.m] and [l.n] can be confused with those with [ŋ.m] and [l.l] sequences, respectively.[5] In fact, we have reasons to suspect that such confusions are possible given that Korean speakers may optionally employ lateralization and nasalization to alter English words that contain these sequences (e.g., Kang, 1996). For example, loans such as *walnut*, *Telnet*, *Big Mac*, and *Pacman* are sometimes produced as having undergone the lateralization and nasalization rules where applicable by Korean speakers (e.g., *wa*[l.l]*ut*, *te*[l.l]*et*, *bi*[ŋ.m]*ac*, *Pa*[ŋ.m]*an* etc.). It should be noted that to have a complete experimental design, it would be ideal to also have test pairs where nonce words with coda stridents are compared to other nonce words where the coda has been neutralized (by undergoing the coda neutralization rule; e.g., [pac.ta] vs. [pat.t'a]). We opt not to include such pairs, however, for the following reasons. First, unlike in the case of sonorant assimilation processes, we know of no loan word

[5] We wanted to be as exhaustive as possible to control for possible cases where nasal spreading may occur. Thus, we also included [n.n] as a possible confusion for [l.n]. In fact, [l.l] is a possible realization of [n.n] for at least older generation of speakers (Sohn, 1999, p. 168).

adaptation case in Korean where coda neutralization is observed. Furthermore, our observations about Korean L2 speech suggest that such an alternation is very unlikely. Second, coda neutralization feeds other phonological processes in consonantal contact situations under question, yielding several illicit intermediate representations that would need to be tested. For instance, an illicit sequence such as /cm/ involves at least two steps to arrive at the correct output [mm]. The coda /c/ is first neutralized to an alveolar obstruent after the removal of stridency (i.e., */c.m/→ *[t.m]). It then undergoes nasalization triggered by the following nasal (*[n.m]). Finally, labiality from the nasal spreads to the preceding nasal (i.e., [mm]). Therefore, we cannot exclude the possibility that a likely discrimination of [c.m] from [m.m] by Korean speakers is possible because the perceptual repair involves far too many steps. The inclusion of intermediate steps as test words, however, brings up another complexity to the problem since they still involve illicit sequences. A similar problem arises with /ct/, where the neutralized coda ordinarily triggers the tensification of the following onset obstruent, yielding [t.t']. Since the L2 targeted in this study is English, it is not clear to us how an English speaker would produce a tensed obstruent and how Korean speakers would interpret such a non-English obstruent.

Finally, we also want to know whether perceptual distortions can be also be induced by allophonic features violating the Korean phonological patters (e.g., [voice]) as opposed to phonologically contrastive features of Korean such as [strident], [spread glottis], [nasal], and so forth. That is, is the phonological status of the features relevant to perceptual epenthesis? Since Sapir's "The Psychological Reality to Phonemes," it has been often observed that speakers seem unaware of the subphonemic aspects of their speech; perhaps we would find a similar lack of attention to misplaced subphonemic features. That is, Korean speakers might treat pairs of segments differing only in a noncontrastive feature — such as [voice] — in the same way, responding in similar ways to [k.m] and [g.m] stimuli for example.

Since our departing point is to extend Dupoux et al.'s (1999) findings, we formulate our hypotheses primarily with reference to perceptual epenthesis. Accordingly, given a *[C_1.C_2] sequence, we foresee two phonological contexts in which perceptual epenthesis may arise.

> *Consonantal Contact hypothesis:* Korean listeners will apply perceptual epenthesis to all consonantal sequences that are illicit in Korean.

> *Coda Condition hypothesis:* Korean listeners will apply perceptual epenthesis only when there is a syllable structure violation concerning the coda consonant.

A number of statements explicating the above hypotheses are in order. First, the Consonantal Contact hypothesis relates perceptual epenthesis to illicit sequences, and is compatible with string-based approaches to sound distributions. In string-based phonotactics, the structure of speech input consists of linear strings of discrete abstract linguistic units (feature bundles or segments) that are ultimately bound by word or morpheme boundaries (cf. Blevins, 2002; Steriade, 1999). Steriade (1999), for instance, classifies the positions of segments not in syllabic terms (e.g., as onset vs. coda) or in linear terms (e.g., "before a vowel" vs. "after a consonant") but as positions where certain featural contrasts are more versus less perceptible. Accordingly, certain consonant clusters are illegal since the featural contrast on one or both of the segments in the cluster is difficult to perceive in that context. The Coda Condition hypothesis, on the other hand, is based on the view that phonological structure is hierarchical. In particular, the hypothesis crucially refers to syllable structure conditions in the L1 system, which specifically state that certain consonants can never surface as coda or as onset. That is, a given consonant may crucially lack a particular positional identity (e.g., coda identity or onset identity) in the inventory. Specifically, it predicts that Korean listeners hear epenthetic vowels when C_1 is an illicit coda consonant. Therefore, following a vowel, when the listener encounters a consonantal element that does not carry a coda identity (e.g., /c/), (s)he will automatically interpret the consonant as an onset.

Second, each hypothesis involves a different conception of perceptual epenthesis. The Consonantal Contact hypothesis motivates perceptual epenthesis to break up illicit sequences of consonants. In simple linear terms, *a là* Steriade, the percept of an epenthetic vowel ensures a buffer sound that breaks up the unwanted consonantal contact. Creating well-formed syllables, however, is the primary goal of the Coda Condition hypothesis. It should be noted that both hypotheses are unit-independent. That is, the unit of perception can be a segment or a bundle of features.

2 Experiment

2.1 Design

An AX discrimination paradigm was employed, comparing pairs of nonce words that contained a word with an illicit sequence of consonant and another word where the violation is repaired either through epenthesis (e.g., [pʰa**km**a] vs. [pʰa**kʰʊm**a]) or through native phonological rules of Korean (e.g., [pʰa**km**a] vs. [pʰa**ŋm**a]). The assumption was that if Korean listeners hear epenthetic vowels in consonant clusters, they are likely to interpret pairs such as [pʰa**km**a] versus [pʰa**kʰʊm**a] to be the same. If, on the other hand, native phonological processes apply to perception, they should hear pairs such as [pʰa**km**a] versus [pʰa**ŋm**a] to be the same.

2.2 Materials

We constructed nonsense words of the form of [pʰá**C_1(V)C_2**a] to create test pairs.[6] None of the test words corresponded to any existing word in Korean or English.

6 We employ the usual conventions for aspiration in English stops, marking onsets but not codas. English coda consonants are not generally aspirated, and thus coda stops were not pronounced with aspiration in this study. It should be noted that English voiceless stops are adapted into Korean onsets with aspiration, regardless of their original position (e.g., *camera* [kʰa.me.ɾa]; *guitar* [ki.tʰa], examples from Kang, 2003), even when the stops are clearly *not* aspirated in the English source (e.g., *spray* [sɯ.pʰɯ.ɾe.i]). The speaker employed in this study had very little rounding in [ʊ]. He pronounced [c] and [ɟ] as alveopalatal affricate, and had voicing into closure in the pronunciations of [g] and [ɟ].

Following the Korean epenthesis patterns, the vowel was [ɪ] after palatals and [ʊ], the closest approximation for Korean [ɯ] elsewhere. Although the usual transcription for the vowel of the English word 'put', [ʊ], indicates rounding for this vowel, there are several reasons to assume that the English [ʊ] will be mapped on to the Korean epenthetic [ɯ], rather than the rounded [u]. First, on acoustic/phonetic grounds, the English lax [ʊ] has very little rounding compared with English [u]. In fact, acoustic studies have shown that not only the American English lax [ʊ], but also its tense counterpart [u] overlaps with the Korean [ɯ] rather than the Korean [u] (cf. Yang, 1996, p. 259). Following these findings, we measured the first and the second formant frequency values (F_1 and F_2) of the [ʊ] and [ɪ] used in our stimuli to see whether the [ʊ] productions used in our study are comparable to those given for Korean [ɯ] in Yang (1996). Using Praat 4.2, a sample of 16 words containing [ʊ] in four different consonantal contexts contexts (i.e., *pakʊta*, *pagʊta*, *palʊna*, *palʊta*) and 16 words containing [ɪ] in four different consonantal contexts (*pacɪta*, *pacɪma*, *pajɪta*, *pajɪma*) used in our experiment were analyzed. Following Fant (1973), we converted the formant frequencies of the two vowels used in our study, as well as the Korean vowels produced by male speakers presented in Yang (1996, p. 251, Table 3), into a perceptual dimension, using the mel scale, in order to allow for a better approximation of the perceived distances among vowels. A comparison of formant frequences between the two languages reveals that the mean F_1 and F_2 (in mel) of our [ʊ] ($F_1 = 557$, $F_2 = 1142$) is closer to the Korean [ɯ] ($F_1 = 514$, $F_2 = 1284$) than to the Korean [u] ($F_1 = 477$, $F_2 = 987$) reported in Yang (1996). As for [ɪ], while both languages have very close values for F_2 ($F_{1(AE)} = 477$ vs. $F_{1(Kor)} = 447$), the mean of the F_2 values in our data stay far lower than that in the Korean [i] ($F_{2(AE)} = 1495$ vs. $F_{2(Kor)} = 1609$).[7] The superimposed mean F1 and F2 values of /ʊ/ and /ɪ/, used in our study, and F1 and F2 of all Korean vowels (produced by male speakers) given in Yang (1996) are presented in Appendix 1.

On perceptual grounds, we also have reasons to assume that the English [ʊ] will be mapped onto the Korean [ɯ] because Dupoux et al. (1999) demonstrated that the Japanese [ɯ], a phoneme that is exactly the same as the Korean vowel in question, was confused with the French [u]; they explicitly tested the perceptual equivalence of these two vowels in using both Japanese and French speakers in their Experiments 1 and 2 respectively. Furthermore, the high back lax vowel in our stimuli has so little rounding that the first author, a native speaker of Turkish which has both /ɯ/ and /u/ in its vowel inventory, hears it as [ɯ].

As for the consonants, the C_1 varied between a permissible coda, [k] and [l], and an impermissible one, [c]. The onset of the second syllable, that is C_2, was either a stop ([t]) or a nasal ([m] or [n]). The reason why [k] and [l] were specifically chosen as licit coda consonants (C_1) is that these consonants are least subject to assimilatory processes in Korean (see Kabak, 2003, p. 36–52 for details). Furthermore, employing both [k] and [l] allows us to investigate a variety of consonantal contact violation

[7] Anticipating our results from the perceptual study, the [ɪ] in our experimental items was indistinguishable from the (perceptually driven) Korean epenthetic [i] despite the difference in F2 between the two languages.

phenomena: nasalization in the case of [k.m], and lateralization in the case of [l.n], thus yields a broader generalization of the results to obstruents as well as sonorants.

The combination of C_1+C_2 produced either a permissible contact ([p^{h}á**kt^h**a], [p^{h}á**lt^h**a]) or an impermissible one (*[p^{h}a**km**a], *[p^{h}a**ln**a]). A number of observations with respect to the phonetic properties of the consonants used in the experiment are in order. First, all the instances of the coda [k] in the stimuli were unreleased, just like the way all coda stops would be in Korean. Second, the lateral liquid [l] was produced as dark everywhere (i.e., both in the onset and the coda position). In fact, the velarized production of /l/ intervocalically is one of the noticeable characteristics of General American English (the variety of the person who produced the test items) that distinguishes it from other varieties of English (e.g., Received Pronunciation) where /l/ would be clear intervocalically (e.g., Wells, 1982, p. 490). The allophonic variation in /l/ should not matter for the present purposes as Korean /l/ shows no such allophonic variation, alternating instead with a flapped pronunciation.

To answer our second question as to whether native phonological processes can explain perceptual distortions, likely surface interpretations of *[p^{h}akma] and *[p^{h}alna], that is [p^{h}aŋma] and [p^{h}alla]/[p^{h}anna], with nasalization and lateralization respectively, were also included in the experiment. This was specifically to test whether Korean listeners apply sonorant assimilation rules such as nasalization and lateralization that are a crucial part of their production grammar to perceptually repair the illicit contacts. To address our third question in relation to the nature of allophonic rules in perception, we also employed the voiced counterparts of [k] and [c], that is [g] and [ɟ], respectively, in the same cluster combinations (i.e., [p^{h}agt^{h}a], [p^{h}agma]; [p^{h}aɟt^{h}a], [p^{h}aɟma]).

Table 3 summarizes our test variables and their surface permissibility in the Korean production grammar and Table 4 lists all the test words used in the experiment with an example word in English that contains the contact.[8]

Table 3

Status of test clusters in Korean

			C_2	
			Oral Stop (i.e., [t^h])	***Nasal (i.e., [n]or[m])***
C_1	**Licit**	**[k]**	Licit	Illicit
		[l]	Licit	Illicit
	Illicit	**[c]**	Illicit	Illicit
		[ɟ]	Illicit	Illicit
		[g]	Illicit	Illicit

8 In an ideal experiment, the choice of C1 and C2 combinations should be systematically crossed. However, to avoid the duplication of the same phenomena that we wanted to test and to shorten the experiment, we omitted clusters such as [k.n], [gn], [c.n], [ɟ.n], [l.m].

Table 4

Test words used in the experiment

		Possible perceived nonwords		
Cluster	***Nonwords*** ([CVC_1C_2V])	***Epenthesis-nonwords*** ([CVC_1VC_2V])	***Sonorant Assimilation-nonwords***	***Example English words***
[kt^h]	[p^{h}ákt^{h}a]	[p^{h}ák^{h}ʊt^{h}a]	*Not applicable*	*doctor, factory*
[km]	[p^{h}ákma]	[p^{h}ák^{h}ʊma]	[ŋm]	*Acme, Ingmar, kingmaker*
[gt^h]	[p^{h}ágt^{h}a]	[p^{h}ágʊt^{h}a]	*Not applicable*	*pigtail, ragtime*
[gm]	[p^{h}ágma]	[p^{h}ágʊma]	[ŋm]	*dogma, segment*
[ct^h]	[p^{h}áct^{h}a]	[p^{h}ác^{h}ɪt^{h}a]	*Not applicable*	*pitch-tracker*
[c m]	[p^{h}ácma]	[p^{h}ác^{h}ɪma]	[nm] *not tested*	*Richmond, attachment*
[jt^h]	[p^{h}ájt^{h}a]	[p^{h}ájɪt^{h}a]	*Not applicable*	*vegetable (dialectal)*
[jm]	[p^{h}ájma]	[p^{h}ájɪma]	[nm] *not tested*	*arrangement*
[lt^h]	[p^{h}ált^{h}a]	[p^{h}álʊt^{h}a]	*Not applicable*	*saltines, shelter*
[ln]	[p^{h}álna]	[p^{h}álʊna]	[ll], [nn]	*walnut, vulnerable, mail-list; pinenut*[9]

Using a high impedance microphone (Audio-Technica, Omnidirectional, ATR35S), each nonce item was produced several times by a male native speaker of American English, who is a trained phonetician, in a sound attenuated room and recorded directly to compact disc with a JVC XL-R5020 compact disc recorder. The recorded tracks were then transferred onto a Mac OS 9.1 computer, where they were converted to 22kHz (16-bit resolution; 1-channel) and stored as ".aiff" files to be used in PsyScope 1.2.5 PPC (Cohen, MacWhinney, Flatt, & Provost, 1993). AX test pairs were created by pairing different exemplars of test words containing consonant clusters with words where the test clusters were separated by a vowel (e.g., [p^{h}acma] vs. [p^{h}ac^{h}ɪma]; [p^{h}akt^{h}a] vs. [p^{h}ak^{h}ʊt^{h}a]). Based on research by Werker and Logan (1985), who has shown that a longer ISI ensures perception to operate at the phonemic rather than at the phonetic/acoustic level, we used an interstimulus interval (ISI) of 1500 ms between each word (and a 2000 ms interval between trials) to enable the subjects to compare the stimuli in phonologically recoded form rather than in echoic memory, thereby ensuring phonological discrimination rather than very fine-grained acoustic discrimination.

Based on the principles of the Signal Detection Theory (Green & Swets, 1974; Macmillan & Creelman, 2005), we measured both the d-prime (d') and A-prime (A') scores, both of which measure the ability to discriminate between same and different pairs. Both methods have advantages and disadvantages (see Macmillan & Creelman, 2005, p. 100ff for some discussion and references), but in the present study the A' measure has three distinct advantages: (1) it yields values between 0 and 1, which can be compared with intuitive concepts such as "proportion correct," (2) it requires no corrections for all-correct or all-incorrect scores, and (3) it allows for situations in which the false-alarm rates are greater than the hit rates, all of which are relevant for the present experiment. As the Signal Detection Theory statistics requires each

9 Variation in geminate reduction can be observed in the pronunciation of words such as *pinenut* and *mail-list* in English

subject’s response to items that are the same to measure false alarms and correct rejections), we also included different renditions of the same experimental words (e.g., [p^{h}ak^{h}ʊma] vs. [p^{h}ak^{h}ʊma]; [p^{h}akma] vs. [p^{h}akma]).

Since Korean listeners were potentially at a disadvantage as most of the stimulus pairs were expected to yield a “same” response from them, we created filler pairs that were predicted not to pose any discrimination difficulty for Korean listeners (e.g., [palla-panna]; [pacma-panma]; [pak^{h}ɪma-pak^{h}ʊma], etc.). An experimental block was created with two random repetitions of each of the test and filler pairs. No two pairs contained identical recordings (e.g., [[p^{h}akma]$_1$ vs. [p^{h}akma]$_2$; [p^{h}akma]$_3$ vs. [p^{h}akma]$_4$; [p^{h}akt^{h}a]$_1$ vs. [p^{h}ak^{h}ʊt^{h}a]$_2$; [p^{h}ak^{h}ʊt^{h}a]$_3$ vs. [p^{h}akt^{h}a]$_4$]). We created an experimental block from these items, which contained 118 trials (39 test pairs plus 20 filler pairs presented in both possible orders). The experimental block was presented five times in a row, with four self-terminated rest periods in between, in order to get 10 total repetitions for each pair. The items in each block were automatically randomized by PsyScope each time a new block was presented. We employed a practice session for every subject before the actual experiment. Thus, another set of cluster-and epenthesis-words was created (e.g., [p^{h}abma], [p^{h}alt^{h}a], [p^{h}abʊma], [p^{h}alʊt^{h}a], etc.). The session contained five “same” and five “different” pairs, which were expected to cause no problem for both groups of subjects. None of the practice items were included in the actual experiment.

Twenty-five native speakers of Korean and 25 native speakers of English at the University of Delaware participated in the experiment. All the Korean speakers were residing in the U.S.A. for educational purposes and none of them had started learning English before the age of 12. None of the English subjects had any knowledge of Korean.

Each subject participated in the AX discrimination task in a single testing session. They were specifically told that they would hear an American man saying nonsense words of English in pairs, and their task was to determine whether the man repeated the same word the second time or said a different word, pressing <A> on the keyboard for same, and <L> for different. The subjects heard the stimuli over headphones at a comfortable intensity for the subject. No feedback was given regarding subjects’ responses during the practice session.

2.3
Predictions

Under the Consonantal Contact hypothesis, all of the cases that yield illicit consonantal contacts, namely [k.m], [l.n], [c.m] and [c.t^{h}], are expected to be misperceived by Korean participants. This is because they all contain contacts that are not permissible in Korean. If on the other hand, the Coda Condition hypothesis were true, all contacts that contain a permissible coda consonant such as [k.] and [l.] (i.e., [k.m], [k.t^{h}], [l.n], [l.t^{h}]) are *not* expected to be misperceived. With respect to voiced consonants, if Korean participants interpret voiced codas as illicit then both [g] and [j] should behave like [c] because they can also never occur as codas in Korean. To the contrary, if voicing, as an allophonic feature, does not play a role in perception then the same consonants, namely [g] and [j], should behave like [k] and [c], respectively. It should be noted that under all possible circumstances, clusters with strident C_1’s are expected to be misperceived while [k.t^h] and [l.t^h] should not be misperceived since they include both a permissible contact and a permissible coda consonant. Therefore, this experiment guarantees, regardless of our secondary questions, that there is both a good condition (i.e., [k.t^h], [l.t^h]), and a bad condition (i.e., [c.m], [c.t^h], etc.).

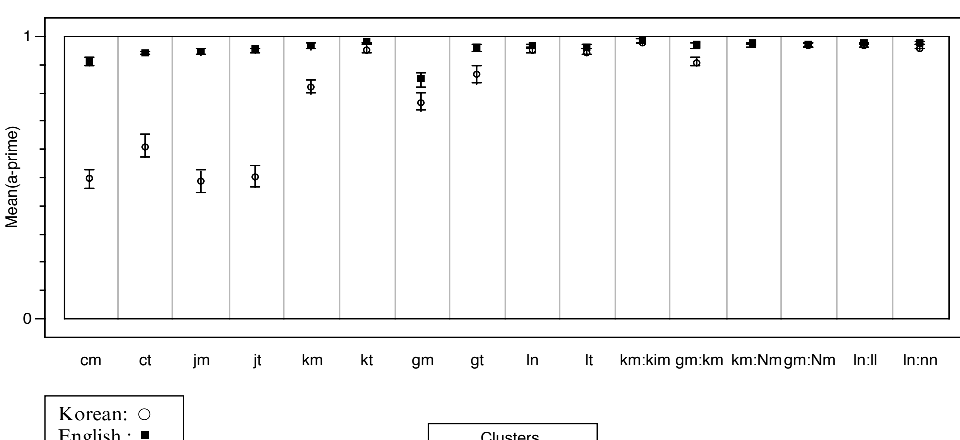
2.4 Results

In this article we report only the statistical analyses on A' scores but provide both the mean d' and A' scores in Appendix 2. When false alarms exceeded hits, d' was taken to be zero; in the case of zero hits, misses, false alarms, or correct rejections, the corrections suggested by Kadlec (1999) were applied to calculate the d' scores. The reader is referred to Kabak (2003) for further information and detailed statistical analyses based on d' scores.

Figure 2 plots the average A' scores for both the Korean and English groups. Overall, the English grouped successfully discriminated all test clusters from their epenthetic and assimilated counterparts, except for the anomalously poor performance on *gm-gʊm*, with an A' score of 0.846. An apparent three-way grouping, however, emerged from the Korean data: (1) the pairs containing a strident cluster (i.e., *cm-cɪm*, *jm-jɪm*, *jt-jɪt*), where the mean A' values were below or around 0.61; (2) *gm-gʊm*, *km-kʊm*, *gt-gʊt*, which formed an intermediate category with A' scores roughly between 0.74 and 0.85; and (3) and all other pairs (e.g., *ln-lʊn*, *lt-lʊt*, *km-kɪm*, etc.), where the A' scores were very close to those of the English group, ranging from 0.91 to 0.98. The first group with strident C_1 (which cannot occur in the Korean coda position), had an average A' of 0.527, indicating an abject failure to discriminate. Both the Korean and the English groups' performances on *gm-gʊm* were somewhat degraded. Except for this cluster, all the mean English A' scores exhibit similar scores above 0.90. While there is no immediate explanation for why both groups suffered some degree of difficulty with *gm-gʊm*, it should be noted that this performance was not nearly as bad as the Korean group's performance on strident clusters.

Figure 2

Korean and English mean A' scores with 95% confidence intervals



Given that the range of individual A' scores vary between 0.49 and 0.98, an important statistical question arises here with respect to which A' scores should be considered as "bad," "moderate," and "successful" performance. For instance, should A' scores between 0.70 and 0.80 indicate a "bad," "moderate," or "successful" discrimination ability? Likewise, should an A' value of 0.74, as in the Korean *gm-gɯm*, be grouped with the strident clusters whose scores were below 0.61 or with those above 0.90 (or with neither)? In the following, we explore answers to these questions by employing various statistical analysis methods.

2.4.1 Analysis of variance

To guard against the possible sphericity violations in repeated measure designs, we first ran Multiple Analysis of Varience (MANOVA) tests. The Mauchly sphericity test from the JMP statistical package revealed that both main effects and the interactions to be significant at a level of $p<.0001$, strongly suggesting that the data cannot be treated as a simple univariate problem without correction. Therefore, we ran a Mixed Effects Repeated Measures Model ANOVA to correct for the violations of sphericity. The resulting ANOVA table is provided in Table 5.

Table 5

JMP mixed effects analysis

Source	*Nparm*	*DF*	*DFDen*	*Sum of Squares*	*F Ratio*	*Prob>F*
Group	1	1	720	1.2989565	174.1028	<.0001
Cc	15	15	720	8.0925944	72.3116	<.0001
Group*cc	15	15	720	5.6616450	50.5898	<.0001
Subject & Random	50	48	720	0.5827476	.	. Shrunk

We found no differences between the two ANOVAs obtained from the two different statistical packages. All main factors were uniformly significant: Language Group: $F(1, 720)=174.103$, $p<.0001$; Consonant Cluster Type: $F(15, 720)=72.312$, $p<.0001$). Moreover, the two-way interaction between the Language Group and Consonant Cluster Type was also significant, $F(15, 720)=50.59$, $p<.0001$, which suggested that the Korean and English listeners performed significantly differently on some consonant clusters, as expected. Figure 2 given above provides the interaction line plot with the 95% confidence intervals for both groups, and Table 6 below provides the Least Square Means table based on the post-hoc Tukey HSD test with an alpha level of .05. Here, we choose to present only the results from the Turkey HSD post-hoc test due to space constraints (for the critical differences between means obtained by other post-hoc tests, see Appendix 3).

Table 6

LS Means Differences Tukey HSD (levels not connected by same letter are significantly different)

Level								Std. Dev	Std. Err.	Least Sq Mean
English, *km-kɯm*	A							0.021	0.004	0.98240839
Korean, *km-kɯm*	A							0.031	0.006	0.97978246
English, *kt-kʊt*	A							0.022	0.004	0.97539501
English, *ln-ll*	A							0.026	0.005	0.97442696
English, *ln-nn*	A							0.029	0.006	0.97442696
Korean, *km-ŋm*	A							0.032	0.006	0.97157007
English, *km-ŋm*	A							0.043	0.009	0.97091402
English, *gm-ŋm*	A							0.025	0.005	0.96919472
Korean, *gm-ŋm*	A							0.027	0.005	0.96684652
Korean, *ln-ll*	A							0.025	0.005	0.96583333
English, *gm-km*	A							0.034	0.007	0.96521854
English, *ln-lʊn*	A							0.031	0.006	0.96277784
English, *km-kʊm*	A	B						0.040	0.008	0.96184294
English, *lt-lʊt*	A	B						0.046	0.009	0.95887265
Korean, *ln-nn*	A	B						0.041	0.008	0.95780072
English, *gt-gʊt*	A	B						0.047	0.009	0.95694781
Korean, *kt-kʊt*	A	B						0.058	0.012	0.95438869
Korean, *ln-lʊn*	A	B						0.049	0.010	0.95259503
English, *jt-jɪt*	A	B						0.041	0.008	0.95136432
English, *jm-jɪm*	A	B						0.054	0.011	0.94427476
Korean, *lt-lʊt*	A	B	C					0.048	0.010	0.94290964
English, *ct-cɪt*	A	B	C					0.045	0.009	0.93854665
English, *cm-cɪm*	A	B	C	D				0.063	0.013	0.90903718
Korean, *gm-km*	A	B	C	D				0.060	0.012	0.90833077
Korean, *gt-gʊt*		B	C	D				0.163	0.033	0.86514065
English, *gm-gʊm*			C	D	E			0.115	0.023	0.84606056
Korean, *km-kʊm*				D	E			0.129	0.026	0.82243844
Korean, *gm-gʊm*					E			0.168	0.034	0.76622222
Korean, *ct-cɪt*						F		0.189	0.038	0.61088450
Korean, *jt-jɪt*							G	0.195	0.039	0.50475708
Korean, *cm-cɪm*							G	0.160	0.032	0.49644806
Korean, *jm-jɪm*							G	0.197	0.039	0.48573464

With respect to the English groups' A' scores, Figure 2 shows that the scores create a flat profile, except for *gm-gʊm*, whose confidence intervals overlap with those of *cm-cɪm*. The post-hoc Tukey Honestly Significant Differences test provided in Table 6 divides the English scores in different groups, which nevertheless overlap with one another. Such different renditions within the English group are an artifact of statistics whose goal is to vigorously detect differences that are often insensible. However, *gm-gʊm*, being significantly different from all other clusters (except *cm-cɪm*), does merit further consideration. A closer inspection reveals that there were vast individual differences in the English group for this item, suggesting that some of the participants experienced some trouble with the cluster. Given that English contains quite a number of simplex words with this cluster (e.g., *dogma, pigment, sigma, pragmatic*, etc.) compared to those with [cm] or [ct], which we cannot observe in simplex words, the difficulty experienced by the English group does not find an immediate explanation. Other than this small problem, the English group successfully discriminated all other clusters.

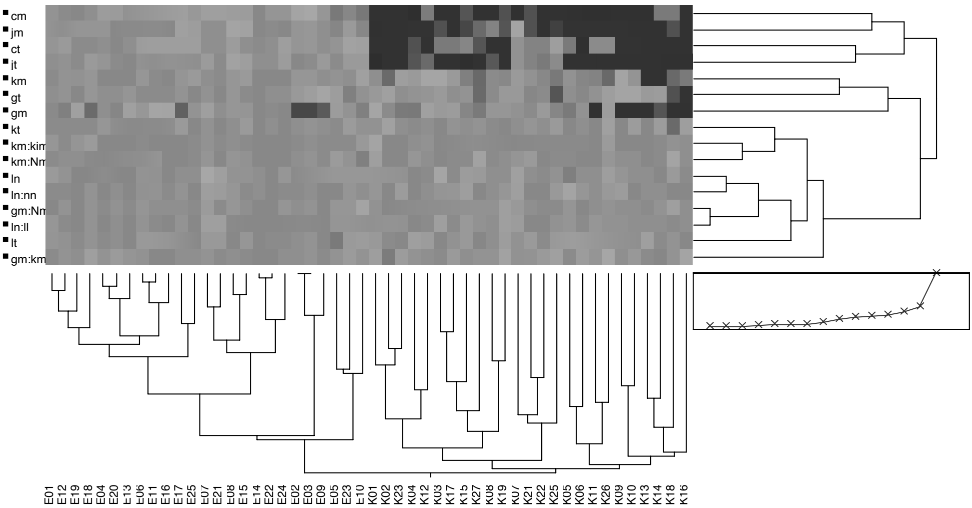
When we consider the Korean group's A' scores, we see in Figure 2 that all pairs comparing the illicit consonantal sequences with their legal counterparts in the Korean grammar (e.g., *km-ŋm, ln-ll*, etc.), as well as other pairs such as *km-kim* and *gm-km* have scores that are not significantly different from those of the English group. Concerning the pairs in the epenthesis condition, which are crucial for our hypotheses, the Tukey HSD post hoc test divides the consonant clusters into three relatively clear groups: (1) *cm, jm, jt, ct* — Groups F and G, (2) *gm, km, gt* — Groups D and E, and (3) *lt, ln, kt* — Group A, ordered from the lowest to the highest A'. Korean *gt* is the most difficult to classify, as it is nondistinct by the post hoc tests from both *km* (below) and *lt* (above). With the data divided into three distinct groups, however, our hypotheses cannot be evaluated in a simple and linguistically meaningful way. First, it cannot be said that the intermediate performance on *km-kʊm*, *gm-gʊm*, and *gt-gʊt* provides sufficient evidence for the Consonantal Contact hypothesis because *ln-lʊn*, which also induces a consonantal contact violation in Korean, is among the top three successfully discriminated clusters in the Korean group. Furthermore, the Korean mean score on *ln-lʊn* overlaps with that of the English group's performance on the same cluster. Indeed, this finding, by itself, disputes the Consonantal Contact hypothesis. Second, the Korean group's performance on these clusters is not as bad as the performance on the strident clusters, which yielded near complete indiscriminability. Therefore, it is necessary to distinguish relative degrees of discriminability in order to better evaluate the hypotheses. How can the discriminability indices on the consonant clusters be grouped if only two different groups of performance were to be made?

To answer this question, we ran a multivariate hierarchical Cluster Analysis (CA) on the Korean data using the JMP Statistical Package. CA is a multivariate analysis technique that organizes information about variables and sorts cases into clusters in such a way that the degree of association is strong between the members of the same cluster and weak between the members of different clusters. CA in Figure 3 simultaneously provides two dendograms, which are tree structures for clusters of cases (in the case of the present study, the test consonant clusters and individual subjects) by calculating multivariate correlation between cases and then clustering them from strongest mutual correlation to weakest. The degree of association between the clusters in the subjects-dendogram and those in the consonant cluster pair-dendogram are indicated by the density of the colors on the rectangular area.

The subjects are arranged on one axis of the dendogram, and we clearly see the two language groups — English and Korean — emerge from the similarity metric, breaking exactly between the groups to form the two clusters for the final merge into the single cluster for the whole dataset. The AX pairs are listed on the other axis, and we see the same three groups emerge as did in the post hoc tests: (1) *cm*, *jm*, *ct*,

Figure 3

Korean and English *A*' cluster analysis dendogram



jt; (2) *km*, *gt*, *gm*; and (3) *kt*, *ln*, *lt* and the various other conditions. Furthermore, the dendogram also assesses the similarity between these three groups, and rates Groups (2) and (3) as more similar than (1) and (2). That is, the "intermediate" group is classified with the "good performance" group, not with the "bad performance" group. Thus, the degradation in performance for *km*, *gt*, and *gm* is not that significant, and future work should try to ascertain why the performance was degraded slightly for these cases. With the cluster analysis in hand, we conclude that there is only one group with truly bad performance: Korean listeners on *cm*, *jm*, *ct*, and *jt*.

It should be remembered that to guard the conclusions on perceptual epenthesis against other types of phonological adjustments such as lateralization and nasalization, we also included a number of doublets which compared [k.m] and [l.n] with their likely output forms in the Korean production grammar, [ŋ.m] and [l.l] or [n.n], respectively. The mean A' scores for these pairs revealed that they are almost identical and very high in both the English (mean A' = 0.972) and the Korean (mean A' = 0.967) groups, indicating that the Korean listeners did not confuse the illicit consonant clusters with their likely output forms in Korean. In summary, the difference between [k.m] and [ŋ.m], and [l.n] and [l.l] was very noticeable to Korean listeners; nevertheless, the alternation from these input forms to output forms is made in Korean production. In the following, we will take this finding to suggest that not all phonological processes are relevant for the Korean listeners' perception.

3 Discussion

In this study, we have motivated two fundamentally different hypotheses based on two different approaches to phonotactics, which make crucially different predictions about the way L1 phonotactic knowledge influences the perception of L2 consonant clusters. The results of our experiment showed that the Korean listeners can distinguish a given consonantal sequence $C_{1.}C_2$ containing a permissible coda consonant (C_1) from its epenthetically adjusted counterpart (i.e., C_1VC_2) regardless of whether the following consonant (C_2) is a plosive or a nasal sound. Consequently, clusters (e.g., [k.m]) that never occur in Korean can nonetheless be successfully perceived by Korean listeners.

Note that the [l]-clusters, namely, [l.n] and [l.t^h], form a replication of the same question that [k.m] and [k.t^h] aimed to test in the study. According to the Consonantal Contact hypothesis, [ln] is expected to be problematic while both clusters should be successfully discriminated under the Coda Condition hypothesis. This is because [l] is a permissible coda consonant. An analysis of the lower and upper bounds for the clusters [l.n] and [l.t^h] within 95% confidence rate reveal that these clusters are not significantly different from each other. In addition, both the English and the Korean listeners do not differ from one another on these clusters.

The most crucial finding of the present study is that not all bad contact cases behave the same in Korean perception, which is summarized in Table 7, where line (= =) indicates the separation into two clusters determined by the Cluster Analysis.

Table 7

Summary of Korean results

Sequence	Type	mean A'
ln	Bad contact	0.95
kt^h	OK	0.94
lt^h	OK	0.94
gt^h	Bad contact	0.85
km	Bad contact	0.82
gm	Bad contact	0.74
==========	==========	==========
ct^h	Bad coda	0.61
jt^h	Bad coda	0.51
cm	Bad coda	0.50
jm	Bad coda	0.49

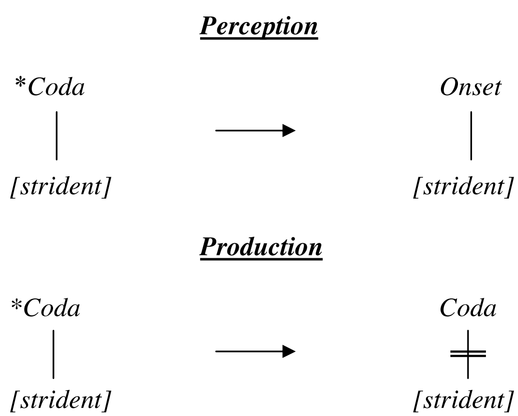
The response patterns can only be explained if we assume that the L1 syllable structure constraints, rather than contact violations, influence the perception of consonant clusters. This is in support of the Coda Condition hypothesis, which is syllabically motivated, and in support of those psycholinguistic models that recognizes the role of syllables (e.g., Cutler, Demuth & McQueen, 2002; Cutler & Norris, 1988). These results also constitute important evidence against views that attribute perceptual preference for certain consonant clusters to the frequency with which those clusters occur in the language thus to the listeners' differing experience with them (e.g., Vitevitch & Luce, 1998, 1999). If perceptual epenthesis were a means by which the perceptual system biases processing of clusters that presumably have zero frequency, then all the illicit consonant clusters in the present study would be more susceptible to epenthesis. Our findings, however, demonstrate that Korean listeners' exhibit poor performance on only a certain set of consonant clusters (i.e., the strident cases) although all the illicit consonantal sequences have zero frequency of occurrence in Korean production. It is instructive to quote Kim-Renaud (1995, p. 223) here: "a sequence of *l* and *n* in either order is absolutely not permissible on the surface in Korean, and a complete assimilation of *n* to *l* occurs whenever such a sequence arises." The same also goes for *k* and *m*. Consequently, we suggest that a phonological influence of L1 phonotactic knowledge, rather than an effect of frequency, plays a primary role in explaining Korean groups' performance. The role of frequency cannot certainly be ruled out here since, after all, such sequences can be attested in Korean beyond Intonational Phrase boundaries. While we fail to see how our words can be perceived as containing IP boundaries, we believe that our results show a secondary effect of frequency. It should be noted that these findings are consistent with those of Moreton (2002), who argues for structural (featural) differences, rather than frequency differences, influencing English listeners' perceptual biases against certain illicit consonant clusters such as [dl] compared to [bw], both of which have zero frequency in English.

We would like to briefly touch upon a secondary question we aimed to investigate in this study, namely the status of voicing, an allophonic rule in Korean, in perceptual adaptations. Our results indicate that [voice] does not create a coda violation in the perception of [g]-clusters although [g] never surfaces in the coda position in Korean. We take this to suggest that the place and manner information retrieved from the stimuli was mapped onto the underlying phoneme, /k/, of which [g] is a predictable variant in Korean. Since Korean allows /k/ in the coda position, the patterning of [g]-clusters with [k]- and [l]-clusters finds an explanation. With respect to [ʤ]-clusters, however, we do not have any evidence to show that voicing is also suppressed here since, after all, the phoneme underlying [ʤ] is also impermissible in the coda. While further experimental work is necessary, we believe that our results indicate that listeners can differentially respond to sounds that are impermissible on the surface. In particular, while both strident codas and voiced codas never surface in Korean, stridency, as opposed to voicing, seems to be the relevant feature for building phonological representations, predictable information being suppressed. This is indeed in line with previous work that supports the hypothesis that representations computed from speech are phonemic in nature (e.g., Dehaene-Lambertz & Gliga, 2004; Kazanina, Phillips, & Idsardi, 2005; Whalen & Liberman, 1987).

Finally, we have demonstrated that the processes of nasalization and lateralization do not play the same role as epenthesis in perception such that the illicit consonant clusters are not misperceived as their likely output forms, as we have seen in Table 4 above. In the case of comparably impossible clusters that contained [strident] onsets, our results suggested that Korean listeners did perceptually change the input to fit into the L1 sound patterns. It is instructive here to note that the perceptual strategy that was employed by Korean listeners with [strident] codas is presumably not the same as in their production system. This is because all [strident] codas are neutralized in Korean. It should be noted that although we did not include any test pair to compare [p^{h}act^{h}a] or [p^{h}acma] with their neutralized counterparts: [p^{h}at't^{h}a] and [p^{h}atma] / [panma], respectively, we can confidently say that the Korean listeners should be able to distinguish them given that the discriminability scores between pairs containing illicit sequences of consonants and their likely surface realizations (e.g., [pakma] vs. [paŋma]) were very high. Indeed, one reasonable interpretation of these results is that stridency is so salient that Korean listeners could not suppress it. Hearing stridency can be said to make Korean listeners place the strident segment in an onset position, which automatically evokes perceptual epenthesis. The different mechanisms employed by the perceptual and production systems are graphically illustrated in Figure 4.[10]

Figure 4

Korean perception versus production grammar



Having demonstrated that perceptual phenomena are not simple inversions of phonological phenomena in the production system, the findings of the present study are difficult to interpret in models that incorporate listeners' knowledge of perceptibility of sound contrasts to predict phonological alternation, the centerpiece of Steriade's model (e.g., Steriade, 1999, 2001 a,b), which we will discuss in the following section.

4 The P-map hypothesis

The P-map hypothesis (Steriade, 2001 a,b) has been proposed to account for directional asymmetries in phonological processes based on an assumption that less perceptible changes are cross-linguistically preferred since they involve repair strategies that

[10] Naturally, the assumed intermediate representation is different in production and perception. While in production, the feature [strident] is associated with a coda position and then undergoes change, the same feature is never in a coda position in perception. Rather, the original stimulus, that is, the English production, has the feature [strident] linked to a coda.

mediate between a maximally similar underlying representation and a surface form. According to Steriade, speakers have the knowledge of the listener's ability to perceive differences and such knowledge of perceptibility contrasts is encoded in the phonology as similarity rates. That is, speakers choose less noticeable repairs since they know they are less noticeable by listeners. The P-map hypothesis has its roots in Steriade's (1999, p. 4) Licensing by Cue hypothesis: "the likelihood that distinctive values of the F-contrast will occur in a given context is a function of the *relative perceptibility* of the F-contrast in that context" (emphasis added). Discriminability scores obtained from Signal Detection Theory (SDT) analysis (i.e., d' scores or A' scores) give a direct measurement of relative perceptibility, allowing a direct test of the Licensing by Cue hypothesis. Accordingly, the probability that a consonant undergoes or feeds a phonological process (such as deletion, triggering epenthesis, or blocking vowel deletion) correlates with the quality and quantity of the auditory/phonetic cues associated with the contrast in a given context. For instance, voiceless, and voiced obstruents are neutralized before obstruents and word-finally as opposed to intervocalic contexts because the contrast between the two is more perceptible in intervocalic contexts as more of the acoustic cues to voicing such as formant values of adjacent vowels, closure duration, VOT value, are available. For Steriade, voicing is likely to be maintained or lost depending on the speakers' precise knowledge of the perceptibility of voicing contrasts in particular phonological contexts that make the contrast more or less salient for the listener. This knowledge is the P-map, "the repository of speakers' knowledge, rooted in observation and inference, that certain contrasts are more discriminable than others, and that the same contrast is more salient in some positions than others" (Steriade, 2001b, p. 236).

4.1
Testing the P-map

According to Steriade's (2001b, p. 222) "Perceptual Similarity to Input" idea, "the likelihood that a lexical representation R will be realized as R' is a function of the perceived similarity between R and R'." Since Steriade explicitly claims that perceptual similarity is psychologically real and rooted in observation and inference, we maintain the position that it is empirically testable. Indeed, the degree of similarity between any two features, by definition, should be inversely related to their discriminability. The methodology employed in the present study can, therefore, constitute an empirical test for the validity of Steriade's statement. Essentially, the discriminability indices that were obtained from the present study are the inverse of "perceived similarity." The SDT analysis computes a discriminability value by comparing hit rates with false alarm rates. Thus, "perceived similarity" can be calculated with reference to the index of discriminability, which must be the inverse of similarity. The mathematical model in (5) illustrates the conversion of discriminability scores (A') on [k.m] versus [ŋ.m] into "perceived similarity" scores:

(5) p ([…ŋ.m…]| /…k.m…/) = 2 (1 – A'([…ŋ.m…], […k.m…]))

Accordingly, the probability of producing [ŋ.m] given /k.m/ is an inverse of the discriminability score obtained on [ŋ.m] versus [k.m]. That is, when A' is 1—perfect discrimination—then *p* is zero—and we should have no phonological rule, but when A' is 0.5 — discrimination equal to chance — then *p* is 1 and we should have a phonological rule. This model assumes a linear relationship, which is almost certainly incorrect, and yields absurd probabilities when A' is less than 0.5, but the model is very simple and suffices to illustrate our point.[11] Our findings showed that Korean group's A' scores on [pʰakma] versus [pʰaŋma] (or vice versa) is .97. According to (5), then p([pʰaŋma]| /pʰakma/) =.06, and there should be only a 6% chance of a phonological rule relating the items. Thus, the /k.m/ → [ŋ.m] process in Korean cannot be explained by Steriade's notion of *Perceptual Similarity to Input*.

The same equation can also be applied to [l.n] versus [l.l]. Our results show that the Korean group's A' scores on this pair are again very high (0.97). Such a high score results in probability scores that are very close to 0, which again contradicts Steriade's hypothesis. In summary, if assimilatory processes were predicted by an index of perceived similarity between the assimilated variant and the underlying representation, Korean listeners should confuse clusters such as [k.m] with its likely output form [ŋ.m] in Korean. The way loan words containing [k.m] and [l.n] are adapted in Korean (e.g., [*Pa*[ŋ.m]*an*] from *Packman*; see Kang, 1996) gives us every reason to expect that native phonological rules may affect perception. It turns out however that neither epenthesis nor native phonological rules affect the perceptual processing of such strings. Thus, the best working hypothesis is that violations involving syllable structure instead of consonantal contact affect perception, and neither nasalization nor lateralization have any basis in perception.

One reviewer suggests that the status of *km* and *cm* sequences is different in English and therefore the Koreans are just displaying a difference inherent in English (with *km* being licit in English, but *cm* being illicit). This putative explanation does not find any support for two very strong reasons. First, the *English* listeners do not differentiate between these two cases (they are both in Group A by the Tukey-Kramer Honestly Significant Differences post hoc test). Second, there is no support for this hypothesis from electronic dictionary searches. A search of the phonetic transcription field of the Oxford Advanced Learner's Dictionary of Current English from the Oxford Text Archives yielded 38 headwords containing a [km] pronunciation sequence and 33 headwords containing a [cm] pronunciation sequence. The vast majority of these forms were polymorphemic (mostly in -*man*, -*men*, -*make*, etc.). The only plausibly monomorphemic words were *acme*, *drachma*, and *parchment*. The best first approximation for the situation, then, would be that both [km] and [cm] are not allowed within morphemes in English, but both sequences are licit across morpheme boundaries. There is no evidence whatsoever of a difference in status between these two cases in English, and thus no possible explanation for the Korean behavior in this manner.

Finally, we observed that strident codas are perceptually altered via epenthesis by Korean listeners, which suggests that they are indistinguishable from onset stridents. However, no such phonological alteration is attested in the synchronic phonology of Korean. It should be noted that the phonetic properties of stridency and perceptual epenthesis are confounded in our study since the only cases that invoked perceptual

[11] See Kabak (2003) for the conversion of d' scores into similarity scores, which does not yield absurd probabilities since A' values below 0.5 correspond to 0 in d' statistics.

epenthesis involved a strident segment and the segments we employed in the experiment were strident. Thus, one could suggest that the distinction between [c] and [cɪ] is minimal, thus increasing the likelihood of perceiving an epenthetic vowel according to the P-map hypothesis. However, we know of no cross-linguistic tendency that particularly requires stridents to be prevocalic. More importantly, Korean phonology alone does not employ epenthesis to salvage stridency in the coda position although it is phonetically very salient but rather it removes it from the surface representation through neutralization. However minimal the perceptual similarity between [c] and [cɪ] may be, our perceptual results do not suggest that the neutralization of strident segments in the coda position is based in perception. Nor do they explain lateralization and nasalization processes in Korean. It is also true that stridency is particularly salient acoustically; we leave for future research tests of other languages with coda restrictions on other features.

5 Summary and Conclusions

In this article, based on a perceptual experiment, we have demonstrated that language-specific co-occurrence restrictions do not explain the perceptual epenthesis phenomenon in words containing sequentially illicit consonantal sequences. The results obtained from the study suggest that perceptual epenthesis is evoked when the members of the illicit sequences incur a syllable structure violation. The interpretation of our results, therefore, constitutes evidence against theories and analyses that use syllable-independent and linear statements to explain the consonantal distributions (e.g., Blevins, 2002; Dziubalska-Kolaczyk, 1994; Steriade, 1999; 2001a). Another important finding regards the underspecified nature of L2 percepts. We have claimed that L2 representations suspend featural information in the stimuli if the detected features correspond to those that are underspecified in the L1. Accordingly, the present study demonstrated that [voice] was suspended in the representations of Korean listeners.

In addition, we have demonstrated that not all phonological phenomena have relevance to speech perception, as processes such as neutralization, lateralization, and nasalization were not employed by Korean listeners to repair misplaced sequences of consonants which would be altered by the application of these processes according to Korean phonology. This finding provides evidence against models that prioritize the role of speech perception in explaining synchronic phonological processes (Steriade, 1999; 2001a,b). Equally importantly, we have shown that the effects we found with respect to Korean speakers' perception of epenthetic vowels cannot be explained by the frequency of the sequences that are involved since certain consonant clusters ([k.m] and [l.n]) were successfully discriminated by Korean listeners although they have zero probability of occurrence in Korean.

manuscript received: 12. 21. 2004
manuscript accepted: 03. 14. 2006

References

BEST, C. T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange (Ed.) *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 167–200). Timonium, MD: York Press.

BLEVINS, J. (2002). The independent nature of phonotactic constraints: An alternative to syllable-based approaches. In C. Féry & G. van de Weijer (Eds.), *The syllable in optimality theory* (pp. 375–403). Cambridge: Cambridge University Press.

COHEN, J. D., MacWHINNEY, B., FLATT, M., & PROVOST, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, **25**, 257–271.

CUTLER, A., & NORRIS, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, **14**, 113–121.

CUTLER, A., DEMUTH, K., & McQUEEN, J. M. (2002). Universality versus language-specificity in listening to running speech. *Psychological Science*, **13**, 258–262.

DEHAENE-LAMBERTZ, G., DUPOUX, E., & GOUT, A. (2000). Electrophysiological correlates of phonological processing: A cross-linguistic study. *Journal of Cognitive Neuroscience*, **12**, 635–647.

DEHAENE-LAMBERTZ, G., & GLIGA, T. (2004). Common neural basis for phoneme processing in infants and adaults. *Journal of Cognitive Neuroscience*, **16**(8), 1375–1387.

DULAY, H., BURT, M., & KRASHEN, S. (1982). *Language two*. New York: Cambridge University Press.

DUPOUX, E., KAKEHI, K., HIROSE, Y., PALLIER, C., & MEHLER, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, **25**, 1568–1578.

DUPOUX, E., PALLIER, C., KAKEHI, K., & MEHLER, J. (2001). New evidence for prelexical phonological processing in word recognition. *Language and Cognitive Processes*, **16**, 491–505.

DZIUBALSKA-KOLACZYK, K. (1994). Acquisitional evidence against the phonological syllable. *Papers and studies in contrastive linguistics*, **29**, 17–30.

EWEN, C., & van der HULST, H. (2001). *The phonological structure of words: An introduction*. Cambridge: Cambridge University Press.

FLEGE, J. (1980). Phonetic approximation in second language acquisition. *Language Learning*, **30**, 117–134.

FLEGE, J. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, **15**, 47–65.

FLEGE, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 233–273). Timonium, MD: York Press.

FANT, G. (1973). Speech sounds and features. Cambridge, MA: The MIT Press.

GREEN, D. M., & SWETS, J. A. (1974). *Signal detection theory and psychophysics*. Huntington, New York: Krieger.

HALLÉ, P., SEGUI, J., FRAUENFELDER, U., & MEUNIER, C. (1998). Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance*, **24**, 592–608.

ITÔ, J. (1986). *Syllable theory in prosodic phonology*. Unpublished PhD Dissertation. University of Massachusetts, Amherst.

ITÔ, J. (1989). A prosodic theory of epenthesis. *Natural Language and Linguistic Theory*, **7**, 217–259.

JUSCZYK, P., FRIEDERICI, A., WESSELS, J., SVENKERUD, V., & JUSCZYK, A. (1993). Infants' sensitivity to the sound pattern of native language words. *Journal of Memory and Language*, **32**, 402–420.

JUSCZYK, P., LUCE, P., & CHARLES-LUCE, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, **33**, 630–645.

KABAK, B. (2003). *The perceptual processing of second language consonant clusters*. Unpublished PhD Dissertation. University of Delaware.

KABAK, B., & IDSARDI, W. (2003). Syllabically conditioned perceptual epenthesis. *Proceedings of the Berkeley Linguistic Society*, **29**, 233–245.

KADLEC, H. (1999). Statistical properties of d' and [beta] estimates of signal detection theory. *Psychological Methods*, **4**, 22–43.

KANG, H. (1996). English loanwords in Korean. *Studies in phonetics, phonology, and morphology*, **2**, 21–48. Seoul: The Phonology-Morphology Circle of Korea.

KANG, Y.–J. (2003). Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. *Phonology*, **20**, 219–274

KAZANINA, N., PHILLIPS, C., & IDSARDI, W. (2005). *The influence of meaning on perception of speech sounds*. Ms.: University of Maryland / University of Delaware.

KIM-RENAUD, Y.-K. (1995). *Korean consonantal phonology*. Seoul: Hanshin.

LAMONTAGNE, G. (1993). *Syllabification and consonant cooccurrence restrictions*. Unpublished PhD Dissertation, University of Massachusetts, Amherst.

LEE, I., & RAMSEY, S. R. (2000). *The Korean language*. Albany: State University of New York Press.

LUCE, P. A., & PISSONI, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, **19**, 1–36.

MacMILLAN, N. A., & CREELMAN, C. D. (2005). *Detection theory: A user's guide.* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

MAJOR, R. C. (1987). Foreign accent: Recent research and theory. *International Review of Applied Linguistics in Language Teaching (IRAL)*, **15**, 185–202.

MAJOR, R. C. (2001). *Foreign accent: The ontogeny and phylogeny of second language phonology*. Mahwah, NJ: Lawrence Erlbaum Associates.

MARTIN, S. (1992). *A reference grammar of Korean*. Rutland: Charles E. Tuttle Company.

MASSARO, D. W., & COHEN, M. M. (1983). Phonological constrainst in speech perception. *Perception and Psychophysics*, **34**, 338–348.

McCLELLAND, J. L., & ELMAN, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, **18**, 1–86.

MORETON, E. (2002). Structural constraints in the perception of English stop-sonorant clusters. *Cognition*, **84**, 55–71.

NORRIS, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, **52**, 189–234.

PEPERKAMP, S., & DUPOUX, E. (2003). Reinterpreting loanword adaptations: The role of perception. *Proceedings of the 15th International Congress of Phonetic Sciences*, 367–370.

PEPERKAMP, S. (in press). A psycholinguistic theory of loanword adaptations. *Proceedings of the 30th Annual Meeting of the Berkeley Linguistics Society*.

PITT, M. A. (1998). Phonological processes and the perception of phonetically illegal consonant clusters. *Perception and Pschycophysics*, **60**, 941–951.

PITT, M. A., & McQUEEN, J. M. (1998). Is compensation for coarticulation mediated between monosyllabic spoken words. *Journal of Experimental Psychology: Learning, Memory & Cognition*, **15**, 378–387.

SAPIR, E. (1933 / 1949). The psychological reality of phonemes. D. G. Mandelbaum (Ed.), *Selected writings of Edward Sapir in language, culture and personality*. Berkeley, CA: University of California Press.

SOHN, H.-M. (1994). *Korean*. New York: Routledge.

SOHN, H.-M. (1999). *The Korean language*. Cambridge: Cambridge University Press.

STERIADE, D. (1999). Alternatives to syllabic analysis of consonantal phonotactics. In O. Fujimura & B. Palek (Eds.), *Proceedings of the 2nd Linguistics and Phonetics Conference* (pp. 205–245). Prague: Charles University Press.

STERIADE, D. (2001a). *The phonology of perceptibility effects: The P-map and its consequences for constraint organization*. Unpublished Manuscript. University of California, Los Angeles.

STERIADE, D. (2001b). Directional asymmetries in place assimilation: A perceptual account; In E. Hume & K. Johnson (Eds.), *The role of speech perception in phonology* (pp. 219–250). New York: Academic Press.

VITEVITCH, M. S., & LUCE, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science*, **9**, 325–329.

VITEVITCH, M. S., & LUCE, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, **40**, 374–408.

WELLS, J. C. (1982). *Accents of English*. Cambridge: Cambridge University Press.

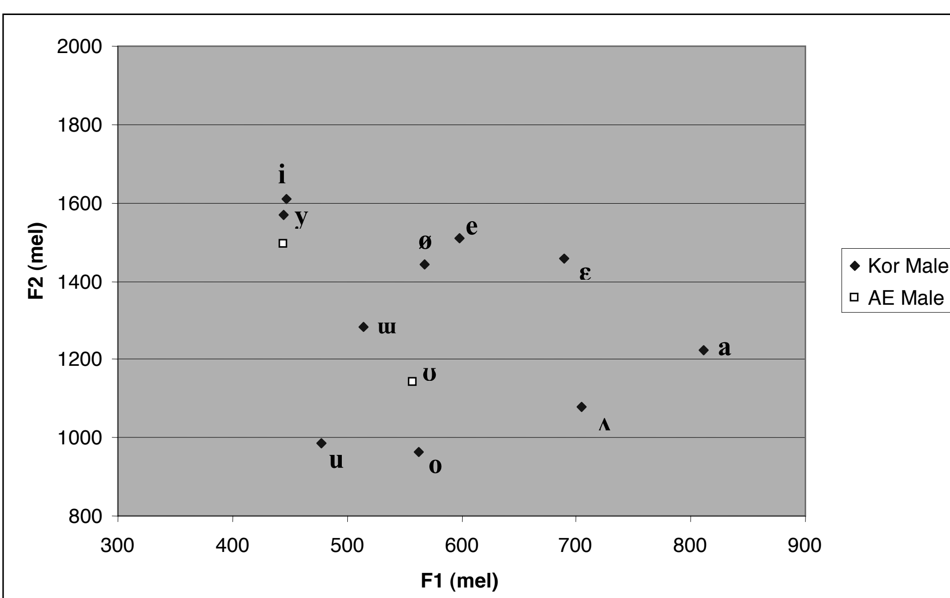
WERKER, J. F., & LOGAN, J. S. (1985). Cross-language evidence for three factors in speech perception. *Perception and Psychophysics*, **37**, 35–44.

WHALEN, D. H., & LIBERMAN, A. M. (1987). Speech perception takes precedence over nonspeech perception. *Science*, **23**, 169–171.

YANG, B. (1996). A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics*, **24**, 245–261.

Appendix 1

Superimposed F1/F2 (in mel) of /ɪ/ and /ʊ/ from the experimental materials and Korean male speakers based on formant values reported in Yang (1996, p. 251, Table III)



Appendix 2

Mean d' and A' scores

	d'	d'	A'	A'
pairs	English	Korean	English	Korean
cm-cɪm	3.032	0.308	0.909	0.495
ct-cɪt	3.338	0.887	0.939	0.613
jm-jɪm	3.416	0.478	0.944	0.487
jt-jɪt	3.436	0.469	0.951	0.514
km-kʊm	3.575	2.282	0.962	0.823
kt-kʊt	3.764	3.580	0.975	0.944
gm-gʊm	2.526	1.807	0.846	0.741
gt-gʊt	3.554	2.724	0.957	0.849
ln-lʊn	3.587	3.421	0.963	0.950
lt-lʊt	3.552	3.280	0.959	0.941
km-kɪm	3.849	3.827	0.982	0.980
gm-km	3.607	2.946	0.965	0.905
km-ŋm	3.689	3.734	0.971	0.973
gm-ŋm	3.657	3.637	0.969	0.967
ln-ll	3.741	3.560	0.974	0.967
ln-nn	3.752	3.483	0.974	0.959

Appendix 3

Critical differences between means according to different post-hoc tests

	alpha level of 0.05		*alpha level of 0.01*	
	English	Korean	English	Korean
Fisher's PLSD	0.024	0.064	0.031	0.084
Scheffe	0.061	0.163	0.068	0.181
Bonferroni / Dunn	0.043	0.115	0.048	0.129
Tukey/Kramer HSD	0.042	0.112	0.047	0.127