

Syllabically Conditioned Perceptual Epenthesis

BARIS KABAK and WILLIAM IDSARDI
University of Delaware

0. Introduction

This paper focuses on perceptual epenthesis; a phenomenon where listeners perceive illusory vowels within consonant clusters which deviate from the phonotactic norms of their native language (see Dupoux et al., 1999). We present results from an experiment on Korean listeners' perception of English consonant clusters which replicates and extends previous studies on Japanese. Our primary aim is to tease apart two explanations for perceptual epenthesis which are confounded in the Japanese studies: consonantal contact violations and syllable structure violations. In light of our results, we suggest here that perceptual epenthesis is caused by syllable structure violations rather than illicit consonantal contact. In addition, we show that speech perception is not always governed by the same system of rules and restrictions that govern speech production. We discuss the consequences of the non-isomorphism between speech production and perception observed in our experiment in the context of the P-map hypothesis (Steriade, 2001a, b). Furthermore, we show that frequency-based analyses fail to account for our results. Finally, we propose a syllabically conditioned perceptual algorithm that accounts for perceptual epenthesis in a simple and straightforward fashion.

1. Background and Problem

The ability to discriminate non-native phonological contrasts deteriorates during the first year of life (see Jusczyk, Houston and Goodman (1998) for a review). Furthermore, it has long been observed that the perception of non-native stimuli strongly suggests that listeners have the expectation that the stimuli they hear are utterances in their L1. Such an L1 biased expectation seems to arise due to our internalized knowledge about the L1 sounds and their patterning, that is the phonology of our L1, which operates as a "filter" both in production and perception of non-native sounds.

One such kind of phonological knowledge pertains to phonotactics. Phonotactics broadly refers to our knowledge about which sound sequences are legitimate and which are not. Languages not only differ in the number and types of sounds they possess in their inventories but also in terms of the various possibilities that exist for the combination of the sounds they have. For instance, while German, English and Dutch allow consonantal sequences such as [dr] and [pl] to begin words, Turkish lacks such complex clusters in the same position although the language contains all of the relevant consonants in its phonemic inventory. The influence of L1 phonotactic restrictions is often reflected in the production of loan words, which are typically altered to fit the L1 norms. For instance, Japanese speakers produce English words that contain consonant clusters by epenthesizing into illicit consonantal sequences (e.g., [makʌdonaRʌdo] 'Mac Donald',

[sutoRaikʌ] 'strike'). Japanese speakers' modification of L2 forms is perhaps not surprising given the very poor consonant cluster inventory of Japanese, which predominantly consists of CV syllables.

Research by Dupoux and colleagues (Dupoux et al., 1999) investigated an interesting question: are the epenthetic vowels added only in production, or are they in fact perceptually derived. They compared Japanese listeners with French listeners in their perception of consonant clusters in a series of behavioral experiments. An off-line phoneme detection task in their study employed nonce words such as [ebuzo] and [egudo] and they created series of six items from each by gradually reducing the duration of the vowel [u] down to zero milliseconds. The subjects were instructed to respond whether each item they heard contained the sound [u]. Unlike French listeners, Japanese listeners overwhelmingly judged that the vowel was present at all levels of vowel length. Interestingly, this was the case 70% of the time even when the vowel had been completely removed (i.e. the 0 ms condition).

Based on these and other similar results, Dupoux and colleagues concluded that the way in which a continuous speech stream is processed in the mind can be heavily influenced by the typical patterning of sounds in the L1 of the listener. Crucially, the influence can be so robust that listeners "invent" illusory vowels to accommodate illicit sequences of segments in their L1. Follow-up studies on Japanese employing a lexical decision task (Dupoux et al., 2001) and a neurolinguistic study using event-related potentials (Dehaene-Lambertz et al., 2000) have further confirmed the perceptual epenthesis phenomena in Japanese subjects.

There are several reasons, however, why a comparison between Japanese and French is not ideal. Japanese contains very few coda-onset clusters as it is predominantly a CV language, and only licenses coda consonants under a very restricted set of circumstances. Specifically, a consonant can only occur in the coda position if it forms the first member of a geminate construction (1a, b), or if it is a homorganic nasal (1c, 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 (2).

- | | | | | |
|--------|---------|---------------------|--------|----------|
| (1) a. | kap.pa | 'a legendary being' | (2) a. | * kap.ta |
| b. | gak.koo | 'school' | b. | * tob.ba |
| c. | tom.bo | 'dragonfly' | c. | * pa.kat |
| d. | kaN.gae | 'thought' | | |

Closer examination on the stimulus items used in Dupoux et al. (1999) reveals that all the test words contained illicit coda consonants (shown in bold): e.g., [i**S**.to], [**eb**.zo], [**ek**.to], etc. Therefore, we cannot tell whether the perceptual epenthesis induced in the percept of words such as [ebzo] is due to (1) contact restrictions (i.e., the sequence [b.z] is impossible), or (2) coda restrictions (i.e., [b] cannot be an independent coda). Since Dupoux and colleagues' primary aim in the study was to document the perceptual vowel epenthesis itself, they did not consider that two phonological explanations for their findings are confounded by the restrictedness of Japanese phonology. Consequently, they chose a more general explanation for Japanese speakers' perception of epenthetic vowels as being the result of their specific knowledge about which consonantal *sequences* are licit and which are not in the L1. That is, their explanation was in terms of contact.

In the present study, we aim to tease apart consonantal contact restrictions from coda restrictions, the confounded factors in the previous studies on Japanese, by employing Korean, which offers a much more interesting array of consonant clusters.

2. Korean Phonotactics

Korean exhibits a three-way contrast for oral stops and affricates: plain (C), aspirated (C^h) and tense (C'). Like Japanese, Korean does not allow complex coda and onsets. However, unlike Japanese, Korean offers a number of consonantal contact situations that help us to separate coda and contact restrictions. First, Korean puts an absolute ban on certain consonants in coda position (e.g., *[c.], *[s.], *[R.], *[h])¹. Second, it additionally disallows certain consonant clusters to surface. For instance, while the sequences *[k.m] and *[l.n] are illicit, [k.t] and [l.t] are licit. As in Japanese, Korean speakers produce foreign words that contain consonantal sequences deviating from Korean phonotactics with epenthetic vowels; [i] after palatal consonants and [μ] in other contexts (e.g., [a.i.sμ.kμ.Rim] 'ice cream'; [sμ.p^hn.ji] 'sponge').

The illicit codas and consonantal contacts are repaired by various phonological rules. In the following, we present only those processes that are relevant for the present study. First, strident consonants such as [c], [c^h], and [s] neutralize in codas to the unreleased stop [t]. For instance, morphophonemic forms such as /nac/ 'daytime', /nac^h/ 'face' and /nas/ 'sickle' become homophonous when they pronounced in isolation (i.e., [nat]). Second, Korean has a nasalization rule that turns stops into nasals in pre-nasal contexts (e.g., /k.m/ → [N.m]: /hak+mun/ → [haN.mun] 'learning'; /p.m/ → [m.m]: /cip+mun/ → [cim.mun] 'house gate'). Third, a process known as lateralization affects nasal sounds after lateral sounds (/l.n/ → [ll]: /tal+naRa/ → [tal.la.Ra] 'moon country'). Finally, voicing in Korean is predictable. Plain sounds become voiced intersonorantly (e.g., /pa+po/ → [pa.bo] 'idiot'). It should be noted that all of the processes illustrated above apply across words (though not across intonational phrases) without any restrictions in Korean.

3. Research Questions

In summary, there are at least three reasons why a word in the form of *[VC₁.C₂V] can be impermissible in Korean. First, [C₁.C₂] could be an illicit contact (e.g., *[k.m], *[l.n]), in the present cases triggering either nasalization or lateralization in the Korean production grammar, as explained above. Second, [C₁] could be an illicit coda (e.g., *[c.]). Third, [C₂] could be an illicit onset (e.g., *[.l], *[N]). In this paper we will only be concerned with the first two factors, see Kabak (2003) for additional discussion of consonants disallowed in onsets.

We investigated whether coda or contact violations caused perceptual epenthesis. We also investigated whether contact violations (e.g., *[k.m], *[l.n]) were perceptually altered to fit their likely surface forms according to the norms of the Korean production grammar (that is, for example, whether words with [k.m] were confused with words with [N.m]). We also investigated the phonological status of the violation; namely whether perceptual epenthesis was induced only by contrastive features of Korean (e.g. [strident], [nasal]) or whether it could also be induced by allophonic features (e.g. [voice]).

3.1. Hypotheses

Given the ways in which *[C₁.C₂] sequences can be illicit in Korean, we propose two main hypotheses to be tested. First, the Consonantal Contact (CC) Hypothesis relates perceptual epenthesis to illicit sequences. Specifically, it asserts that Korean listeners

¹ In this paper, we adhere to the phonetic symbols commonly used in the Korean linguistics literature. Specifically, [c] and [j] stand for voiceless and voiced alveo-palatal affricate, respectively. The symbol [μ] represents the high back unrounded vowel.

hear epenthetic vowels when a given sequence is not valid in Korean. Second, the Coda/Onset Identity (COI) Hypothesis states that perceptual epenthesis arises when there is a syllable structure violation. Specifically, it predicts that Korean listeners hear epenthetic vowels when *[C₁.] is an illicit coda consonant. It should be noted that each hypothesis involves a different conception of perceptual epenthesis. While the CC hypothesis motivates perceptual epenthesis to break up illicit sequences of consonants, creating well-formed syllables is the primary goal of the COI hypothesis.

With regard to our secondary question as to the nature of L2 representations, we propose two sub-hypotheses. According to the Phonetic Processing Hypothesis, all phonetic features in the L1, including allophonic ones, such as voicing in Korean, are represented upon the perception of L2 forms. The Phonological Processing Hypothesis, on the other hand, claims that only contrastive (non-redundant) features in the L1 are represented in the L2 whereby predictable phonological information is suspended. Thus, the Phonetic Hypothesis claims that misplaced instances of [voiced] segments should also induce perceptual epenthesis.

3.2. Design

We employed an AX discrimination paradigm where pairs of nonce words with and without consonant clusters (e.g., [p^hakma] vs. [p^hak^hUma]) were compared with the assumption that if Korean listeners hear epenthetic vowels, they are likely to interpret the two words in the pair to be the same. Naturally produced words in the form of [p^haC₁(V)C₂a] were used to construct test pairs. Following the Korean epenthesis patterns, the vowel was [i] after palatals and [U], the closest approximation of [μ] elsewhere. C₁ varied between a permissible coda, [k] and [l], and an impermissible one, [c]. The onset of the second syllable, that is C₂, was either a stop ([t]) or a nasal ([m] or [n]). Consequently, the combination of C₁+C₂ produced either a permissible contact ([p^hakt^ha], [p^halt^ha]) or an impermissible one (*[p^hakma], *[p^halna]). It should be remembered that in the illicit contact cases, the coda consonants (i.e., [k] and [l]) are permissible coda consonants. The combinations of these consonants, however, with a following nasal can never surface in Korean due to (1) nasalization process in the case of [k], and (2) the lateralization process in the case of [l]. Rather, in the Korean production grammar, [k.m] and [l.n] surface as [N.m] and [l.l], respectively. On the other hand, when C₁ is [c], neither a following stop (i.e., *[p^hact^ha]) nor a following nasal (*[p^hacma]) yield permissible sequences of consonants since [c] can never surface as a coda consonant in Korean.

We also employed the voiced counterparts of [k] and [c], that is [g] and [j], respectively, in the same cluster combinations (i.e., [p^hagt^ha], [p^hagma]; [p^hajt^ha], [p^hajma]) to investigate specifically whether voicing, an allophonic feature predictable by the intervocalic voicing process in Korean, would be represented by Korean listeners, and whether voicing in an inappropriate context would cause perceptual epenthesis. Table 1 summarizes our test variables and their surface permissibility in the Korean production grammar.

Furthermore, the likely surface interpretations of *[p^hakma] and *[p^halna], that is [p^haNma] and [p^halla]/[p^hanna], respectively, were also included in the experiment. This was specifically to test whether Korean listeners apply phonological rules such as nasalization and lateralization that are active in their production grammar to perceptually repair the illicit contacts *[k.m] and *[l.n].

Table 1: Status of test clusters in Korean

		<i>C₂</i>		
			<i>Oral Stop (i.e., [t^h])</i>	<i>Nasal (i.e., [n/or/m])</i>
<i>C₁</i>	Licit	[k]	Licit	Illicit
		[l]	Licit	Illicit
	Illicit	[c]	Illicit	Illicit
		[j]	Illicit	Illicit
		[g]	Illicit	Illicit

AX test pairs were created by putting different exemplars of test words containing consonant clusters with words where the test clusters were separated by a vowel (e.g., [p^hacma] vs. [p^hac^hima]; [p^hakt^ha] vs. [p^hak^hUta]), with a 1500 ms interval between each word. Based on the principles of the Signal Detection Theory (Green and Swets, 1974), we measured d', the ability to discriminate between same and different pairs. Thus, our test pairs also included different exemplars of the same words (e.g., [p^hak^hUma] vs. [p^hak^hUma]; [p^hakma] vs. [p^hakma]). An experimental block was created with 2 random repetitions of each of the test and filler pairs, yielding 118 trials in one block (39 test pairs plus 20 filler pairs presented in both possible orders). The blocks was presented 5 times each, yielding 10 total repetitions for each pair.

Twenty-five native speakers of Korean and 25 native speakers of English were recruited at the University of Delaware. All the Korean speakers were residing in the USA for educational purposes and none of them had started learning English before the age of 12. None of the English subjects knew Korean. Further details on the subjects are given in Kabak (2003).

3.3. Predictions

The predictions of the four combinations of hypotheses considered are summarized in Table 2, in which X indicates a situation where the subjects are predicted to be unable to discriminate the contrast between [paCCa] and [paCVCa].

Table 2: Discrimination predictions. X indicates a prediction of non-discrimination.

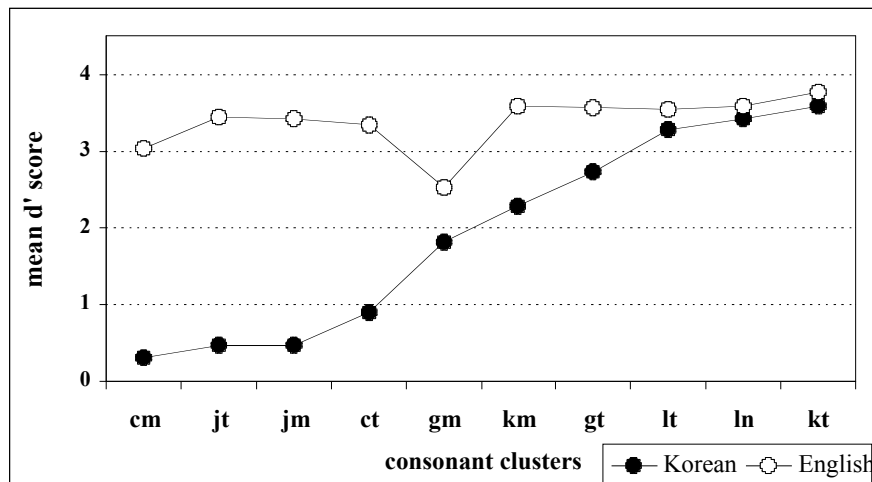
	CC Hypothesis		COI Hypothesis	
	Phonetic	Phonological	Phonetic	Phonological
k.t ^h				
g.t ^h	X		X	
k.m	X	X		
g.m	X	X	X	
l.n	X	X		
l.t ^h				
c.t ^h	X	X	X	X
c.m	X	X	X	X
j.t ^h	X	X	X	X
j.m	X	X	X	X
	A	B	C	D

Under the Consonantal Contact (CC) hypothesis, all of the cases that yield illicit consonantal contacts, namely [k.m], [l.n], [c.m] and [c.t^h], should be misperceived by Koreans. If on the other hand, the Coda/Onset Identity (COI) hypothesis were true, all those with permissible coda consonants (i.e., [km], [kt^h], [ln], [lt^h]) should *not* be misperceived. When it comes to interpreting voiced consonants, if Korean listeners employ a phonetic processing approach and thus represent voicing then both [g] and [j] should behave like [c] since these consonants also never occur in coda position. On the other hand, under a phonological processing hypothesis, if voicing, being predictable, is perceptually suppressed, then [g] and [j] should behave like [k] and [c], respectively.

4. Results

The experimental conditions yielded *d'* values ranging from 4.21 (best) to 0 (worst). The average *d'* scores for both groups for the various clusters are shown in Figure 1.

Figure 1: Mean *d'* scores



The English group successfully discriminated all test clusters, with an average *d'* of 3.38. Observationally, the Korean listeners' *d'* scores exhibited a three-way grouping: (1) [cm], [jm], [jt^h], [ct^h], where the mean *d'* values were below 1; (2) [gm], [km], [gt^h], which formed an intermediate category with *d'* scores roughly between 2 and 3; and (3) [ln], [lt^h], [kt^h], where the *d'* scores were very close to the English group's mean *d'* scores. The first group with strident C1 (which cannot occur in the Korean coda position), had an average *d'* of 0.53, indicating an inability to discriminate. The Korean group has a *median* *d'* value of 0 these clusters, whereas for the other clusters the median *d'* scores were similar to those for the English group. Interestingly, both groups' performance on [gm] was somewhat degraded. While there is no immediate explanation for why both groups suffered some degree of difficulty with [gm], this performance was not nearly as bad as the Korean group's performance on strident clusters.

THESE F STATEMENTS CAN'T BE RIGHT AS THEY STAND.

A repeated measures Analysis of Variance (ANOVA) of *d'* revealed significant statistically significant effects throughout the model. All main factors were significant: language, $F(1, 48)=184.627, p<.0001$, first consonant (C1), $F(4,15)$ THIS CAN'T BE $F(4,15)!! =96.131, p<.0001$ and second consonant (C2), $F(1,48)=55.555, p<.0001$. All

two-way interactions were also significant, though C_2 by language was marginal: C_1 by language, $F(4,45)=57.871$, $p<.0001$, C_2 by language, $F(1,48)=4.159$, $p<.047$, and C_1 by C_2 , $F(4, 45)=18.492$, $p<.0001$. Ordinarily, these two-way interactions could be taken as the first step evidence for an effect of consonantal contact. However, the three-way interaction for C_1 by C_2 by language is also significant, $F(4,45)=6.08$, $p<.001$, and therefore we need to look more carefully at the post-hoc tests. The appendix provides a mean d' scores, standard errors, lower and upper bounds within 95% confidence rate for CVCCV vs. CVCVCV discrimination. Additional statistical analysis can be found in Kabak (2003).

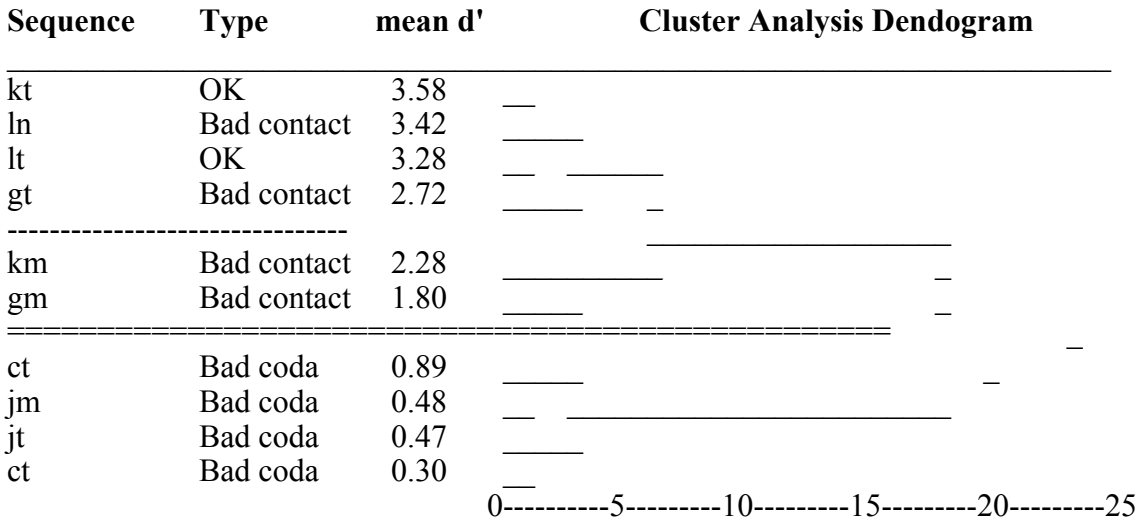
Scheffé and Tukey-Kramer post-hoc tests on the Korean group's d' scores consistently divided the consonant clusters into three groups (1) [cm], [jt^h], [jm], [ct^h], (2) [gm], [km], [gt^h], (3) [ln], [lt^h], [kt^h], ordered from the lowest to the highest. With the data divided into three distinct groups, however, the main hypotheses of the study cannot be evaluated in a straightforward way. First, it cannot be said that the intermediate performance on [km], [gm] and [gt^h] provides evidence for the CC hypothesis because the cluster [ln], which also induces a consonantal contact violation in Korean, is among the top three successfully discriminated clusters in the Korean group and overlaps with the English group's performance. Indeed, this finding, by itself, disputes the CC 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. Third, we constructed our hypotheses in such a way that they allow for only two distinct levels of performance on the test consonant clusters: (1) those that yield poor performance, and (2) those that yield good performance. Therefore, relative degrees of badness needed to be distinguished in order to better evaluate the hypotheses. The question, therefore, remained as to how the discriminability indices on the consonant clusters could be grouped if only two different groups of performance were to be made.

Using SPSS, we ran a multivariate hierarchical Cluster Analysis (CA) on the Korean data to answer this question. The CA produces a dendrogram that graphically clusters cases (in this case, the test consonant clusters) by starting with single member clusters, which are then gradually fused until finally one large cluster is formed (see Figure 2). In a dendrogram, the degree of association is strong between the members of the same cluster and weak between members of different clusters. The ruler at the bottom of the dendrogram gives an index of an arbitrary measure of dissimilarity. Accordingly, from left to right, the CA divided the consonant clusters into 8, 6, 3, and 2 groups, then finally into one big group. While the 8 and 6-way clustering cannot be interpreted in any phonologically meaningful sense, the 3-way clustering, (1) [cm], [jt^h], [jm], [ct^h], (2) [gm], [km], (3) [gt^h], [ln], [lt^h], [kt^h], cannot sufficiently provide evidence against the COI hypothesis and the Phonological processing hypothesis because there is still a difference between the middle group (i.e., [gm] and [km]), and the lower group (the strident clusters). Furthermore, not all clusters with voiced segments are in the middle group (i.e., [gt^h] is grouped with the clusters that induced high d' scores), suggesting that the voicing information on C_1 does not matter for the Korean listeners, which rules out the Phonetic Processing hypothesis. As mentioned before, another cluster that induces a contact violation has been discriminated by the Korean listeners as well as the English group thus the middle group does not contain this cluster, either.

The two-way grouping that emerges out of the CA, on the other hand, groups the test clusters in a more homogenous way, all the strident clusters on one side and the rest of the test clusters on the other. Such a grouping matches the predictions of Column D of the evaluation table provided in Table 2 above, where only strident codas are expected to

be bad. This confirms only the Coda/Onset Identity (COI) hypothesis in combination with the Phonological processing hypothesis.

Figure 2: Korean d' cluster analysis



It should be noted that the [l]-clusters, namely, [ln] and [lt^h], form a replication of the same question that [km] and [kt^h] aimed to test in the study. We used these clusters as an independent test for the evaluation of the main hypotheses. According to the CC hypothesis, [ln] should be bad while both clusters should be successfully discriminated under the COI hypothesis since [l] is a legitimate coda consonant. An analysis of the lower and upper bounds for the clusters [ln] and [lt^h] within 95% confidence rate (see Appendix 1) reveal that these clusters are not significantly different from each other. In addition, both groups do not differ from one another on these clusters. A power analysis reveals that any real difference in d' scores for these clusters must be less than 0.8 units of d' with 99% confidence if such a difference existed. With such small probable difference, the statistics on the [l]-clusters provide no support for the CC hypothesis, yet again supporting the COI hypothesis.

The experiment also included a number of doublets which compared [km] and [ln] with their likely output forms in the Korean production grammar, [Nm], [ll]/[nn], respectively. This was done to guard the conclusions on perceptual epenthesis from the effects of other types of phonological adjustments such as nasalization and lateralization on consonant clusters in Korean. The mean d' scores for pairs revealed that they are all relatively similar and very high in both the English and the Korean groups, indicating that the Korean listeners did not confuse the illicit consonant clusters with their likely output forms in Korean (Table 4). This suggested that not all phonological processes are relevant for the Korean listeners' perception, an important finding that provides evidence against the P-map hypothesis (Steriade, 2001a), which predicts perceptual similarity between underlying forms and output forms that are derived from them. Our study suggested that the difference between [km] and [Nm], and [ln] and [ll] was very noticeable to Korean listeners; nevertheless, the alternation from these input forms to output forms is made in Korean production.

IS THIS THE RIGHT TABLE? KM VS NGM?

Table 4: Discrimination of palla/palna???

Language	C ₁ C ₂	Mean	Std. Error	95% Confidence Level	
				Lower Bound	Higher Bound
English	ln	3.587	0.108	3.370	3.804
	lt ^h	3.552	0.119	3.314	3.790
Korean	ln	3.421	0.108	3.204	3.638
	lt ^h	3.280	0.119	3.042	3.518

In summary, the results indicated that a given consonantal sequence C₁C₂ containing permissible coda consonants as C₁ is distinguishable from its epenthetically adjusted counterpart (i.e., C₁VC₂) for the Korean group regardless of whether C₂ is a plosive or a nasal consonant. As a consequence, some clusters (e.g. [km]) that do not occur in Korean can nevertheless be accurately perceived by Korean listeners. This supports the syllabically conditioned COI hypothesis in that the response patterns can only be explained if we assume that the L1 syllable structure violations, rather than contact violations, play an important role in the perception of consonant clusters. Furthermore, the voiced coda consonants do not cause perceptual epenthesis in the Korean group, suggesting that the voicing information is suppressed in speech perception. This supports phonological processing of features based on their abstract underspecified representations in the L1 of the listener.

5. Discussion

In this paper, we investigated two competing hypotheses based on two different views on phonotactics. The Consonantal Contact hypothesis followed a string-based approach towards phonotactics, predicting perceptual epenthesis as a consequence of consonantal contact restrictions in the L1 of the listener. The Coda/Onset Identity hypothesis assumed that perceptual epenthesis arises if the illicit consonantal sequence violates the L1 syllable structure conditions. Three different levels of performance on the test clusters were observed. Various statistical analyses suggest that the discrimination of the consonant clusters with strident consonants, namely [c] and [j], was significantly worse than other clusters with [g], [k], and [l] as C₁. Based on empirical evidence, we claim that having a sequentially illicit consonantal sequence is not sufficient to induce perceptual epenthesis. Rather, the cluster must induce a syllable structure violation in order to evoke perceptual epenthesis effects. Our results constitute substantial evidence against any analyses that employ syllable-independent, string-based and linear statements to explain consonantal phonotactics (e.g., Steriade, 1999; 2001a; Blevins, 2002).

The present study also revealed that voicing information was suppressed by the Korean group. Specifically, the clusters [gt^h] and [gm] were successfully discriminated while [jt^h] and [jm] were treated just like [c]-clusters. We suggest that [j] was mapped to /c/, therefore, it induced perceptual epenthesis because the segment is still a strident, whereas [g] was mapped onto /k/, which is a licit coda consonant. We take this finding to suggest that L2 representations are constructed on the basis of the abstract phonological properties of the L1 system of contrast and that the perceptual system suspends featural information in the speech signal if the detected values correspond to those that are underspecified in the L1 of the listener.

We obtained another interesting finding suggesting that not all phonological processes are relevant for speech perception. The experiment showed that processes such

as nasalization and lateralization do not constitute possible perceptual strategies for Korean listeners. Specifically, although the successive occurrences of [k] plus [m], and [l] plus [n] are banned on the surface and /km/ → [Nm] and /ln/ → [ll] are all mandatory in Korean production, the processes of assimilation do not play the same role as epenthesis in production. In the case of [strident] codas, which are similarly impossible on the surface, Korean listeners employed perceptual epenthesis to alter the input. The perceptual strategy that they employed is not in line with the way the same feature is accommodated in the production grammar. While in perception coda [strident] sounds were heard in the onset position causing perceptual epenthesis, in production Korean delinks the same feature from the coda position (i.e., /c/ → [t]). In short, the perceptual system does not appear to use repair strategies that are the simple inversions of the assimilatory processes that are active in the production grammar. This finding provides evidence against models that incorporate listeners' knowledge of perceptibility of sound contrasts to predict phonological alternations.

One such view is employed by Steriade (2001b) in her P-map hypothesis, which attempts to explain phonological alterations based on listeners' observations that certain contrasts are more discriminable than others, and that the same contrast is more salient in some positions than others (p. 236). Steriade claims that "the likelihood that a lexical representation R will be realized as R' is a function of the perceived similarity between R and R'" (p. 222). Indeed, the degree of similarity between any two features, by definition, should be inversely related to their discriminability. Thus, Steriade's statement is empirically testable using the present methodology. The discriminability indices we obtained are essentially the inverse of her "perceived similarity". Accordingly, the probability of saying [ll] given /ln/ must be proportional to a suitable function of the reciprocal of the discriminability score on [ll] vs. [ln]. The same equation can also be applied to [km] vs. [Nm]. Since the discriminability is the inverse of similarity, as $d' \rightarrow 0$, $p(\text{similarity}) \rightarrow 1$. Our results show that the Korean group's d' scores on these pairs are very high (3.56 in the case of [ll] vs. [ln], and 3.73 in the case of [km] vs. [Nm]). Such high scores result in probability scores that are very close to 0, which contradicts Steriade's statement regarding perceptual similarity to input. Specifically, 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 [ln] with its likely output form [ll] in Korean. Furthermore, 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.

Our 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. If perceptual epenthesis were a means by which the perceptual system biases processing of clusters that have zero frequency, then all the illicit consonant clusters in the present study would be more susceptible to epenthesis. Our findings, however, show 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. This suggests that a phonological influence of L1 phonotactic knowledge, rather than an effect of frequency best explain the Korean groups' performance and goes along with the findings of Moreton (2002), who argues for structural 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.

Having discussed that neither consonantal contact violations nor frequency can explain the present findings, we propose a syllable-based perceptual algorithm that explains our results in a straightforward fashion. The algorithm crucially employs syllabically-conditioned detectors that check for the identity of features in relation to a specific prosodic position where they can only be licensed. In the case of Korean, we suggest that the feature [strident] is associated to [Onset]; therefore, it must be linked to an onset position in the perceptual system. When such an [Onset]-linked feature is detected, the perceptual system builds a CV syllable to accommodate the feature. Since onsets must be non-branching in Korean, a vowel must follow to fill the nucleus position in the CV syllable that is newly postulated. Accordingly, perceptual epenthesis is motivated from a necessity to build well-formed syllables. The implications of syllabically conditioned detectors to other features in Korean are extensively discussed in Kabak (2003).

6. Conclusions

Our findings from Korean disentangle the phonological confounds in the previous experiments on Japanese, and strongly indicate that perceptual epenthesis is induced by the violation of syllable structure conditions of the L1, rather than string-based contact restrictions. Furthermore, our findings suggest that L2 forms are perceived based on the abstract underspecified properties of features in the L1 system of contrast. Neither views that rely on listeners' knowledge of perceptability of sound contrasts nor frequency-based accounts can explain Korean listeners' speech perception.

Appendix

Language	C ₁ +C ₂	Mean	Std. Error	95% Confidence Level	
				Lower Bound	Higher Bound
English	cm	3.032	0.125	2.781	3.284
	jm	3.416	0.141	3.132	3.700
	ct ^h	3.338	0.163	3.010	3.665
	jt	3.436	0.134	3.167	3.706
	km	3.575	0.171	3.231	3.918
	gm	2.526	0.214	2.096	2.957
	kt	3.764	0.103	3.556	3.972
	gt ^h	3.554	0.165	3.222	3.886
	ln	3.587	0.108	3.370	3.804
	lt ^h	3.552	0.119	3.314	3.790
Korean	cm	0.308	0.125	0.056	0.559
	jm	0.478	0.141	0.195	0.762
	ct ^h	0.887	0.163	0.559	1.215
	jt	0.469	0.134	0.199	0.738
	km	2.282	0.171	1.938	2.625
	gm	1.807	0.214	1.377	2.238
	kt ^h	3.580	0.103	3.372	3.788
	gt ^h	2.724	0.165	2.393	3.056
	ln	3.421	0.108	3.204	3.638
	lt ^h	3.280	0.119	3.042	3.518

References

- Blevins, J. (2002). The independent nature of phonotactic constraints: an alternative to syllable-based approaches. In C. Féry and R. van de Weijer (eds.). *The Syllable in Optimality Theory*, 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, 6, 1568-1578.
- Dehaene-Lambertz, G., Dupoux, E., & Gout, A. (2000). Electrophysiological correlates of phonological processing: a cross-linguistic study. *Journal of Cognitive Neuroscience*, 12, 635-647.
- Green, D. M., and Swets, J. A. (1974). *Signal Detection Theory and Psychophysics*. Huntington, NY: Krieger.
- Itô, J. (1986). *Syllable theory in prosodic phonology*. PhD Dissertation, University of Massachusetts, Amherst.
- Itô, J. (1989). A prosodic theory of epenthesis. *Natural Language and Linguistic Theory*, 7, 217-259.
- Jusczyk, P., Houston, D. & Goodman, M. (1998). Speech perception during the first year. In A. Slater (ed.), *Perceptual development: Visual, Auditory, and Speech Perception in Infancy*. Psychology Press.
- Kabak, B. (2003). *The Perceptual Processing of Second Language Consonant Clusters*. Ph.D. dissertation. University of Delaware.
- Moreton, E. (2002). Structural constraints in the perception of English stop-sonorant clusters. *Cognition*, 84, 55-71.
- Steriade, D. (1999). Alternatives to syllabic analysis of consonantal phonotactics. In O. Fujimura, B. Joseph, & B. Palek (eds.), *Proceedings of the 2nd Linguistics and Phonetics Conference*, 1998 (pp. 205-245). Prague: Charles University, Press.
- Steriade, D. (2001a). The Phonology of Perceptibility Effects: The P-map and its consequences for constraint organization. Ms. UCLA.
- 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.

ADD AUTHORS' ADDRESSES