

## **Perception of onset clusters by English and Korean listeners: Universal markedness and L2 phonotactic knowledge<sup>\*</sup>**

Eunkyung Sung

(Cyber Hankuk University of Foreign Studies)

**Eunkyung Sung. 2016. Perception of onset clusters by English and Korean listeners: Universal markedness and L2 phonotactic knowledge.** *Studies in Phonetics, Phonology and Morphology* 22.3. 477-497. This study investigates the role of universal markedness based on sonority distance and L2 phonotactic knowledge in the perception of onset consonant clusters. Ten native English listeners (NE), ten native Korean listeners with higher English proficiency (NKH), and ten native Korean listeners with lower English proficiency (NKL) participated in a syllable count task. The participants listened to four types of onset clusters based on sonority distance in monosyllables (i.e., large sonority rises, small rises, plateaus and falls) and matched disyllabic counterparts with vowel epenthesis. The results showed that for response accuracy, the NE and NKH groups showed more sensitivity to sonority distance than the NKL group. In monosyllables, both the NE and NKH groups were more accurate as the sonority distance increased, but they did not distinguish between sonority plateaus and falls. For disyllables, all the listeners differentiated consonant clusters with a large sonority rise from the other types of clusters. For response time, the NE group was faster than the two Korean groups. Sonority-based markedness was not able to explain all the results of the present study. Different misperception patterns shown in the three listener groups are partly due to linguistic experience in English. Universal markedness could be restricted by L2 phonotactic knowledge in Korean listeners' interlanguage. (Cyber Hankuk University of Foreign Studies)

**Keywords:** universal markedness, sonority distance, phonotactic knowledge, constraints, L2 proficiency, onset clusters, accuracy, response time

---

\* This work was supported by the Cyber Hankuk University of Foreign Studies Research Fund of 2016. I am grateful to three anonymous reviewers for their insightful comments and suggestions on this paper.

## 1. Introduction

Phonotactic constraints have been shown to play a crucial role in first and second language acquisition. In order to probe the effect of phonotactic constraints in perception, a number of investigations have focused on consonant clusters. Listeners seemed to use their knowledge of native language (L1) phonotactic constraints in their perception of nonnative sequences (Hallé et al. 1998, Pitt 1998, Dupoux et al. 1999, Moreton 2002, Kabak and Idsardi 2007, Mikhaylova 2009, Sperbeck and Strange 2010, Lee 2011). Moreover, it has been shown that universal markedness constraints of onset clusters affected speech perception (Altenberg 2005, Albright 2007, Berent et al. 2007, 2008, 2009, Lennertz and Berent 2015).

Sensitivity to L1 phonotactic constraints has been reported in previous studies of nonnative consonant clusters. Pitt (1998) showed that English speakers had difficulty in distinguishing between illegal clusters (e.g., *tla*) and their epenthetic counterparts (e.g., *talə*). Also, when they listened to legal or illegal onset clusters, their responses were strongly biased toward the favor of legal sequences in comparison with illegal sequences. For example, when the illegal cluster /tʌ/ was presented, the participants tended to respond as if they heard /tr/. Similar results were found by Hallé et al. (1998). When native French speakers were presented with illegal /dl/ and /tl/ sequences, they often misperceived the sequences as close legal clusters in their L1 such as /gl/ and /kl/. Hallé et al. (1998) claimed that nonnative listeners assimilated illegal consonant clusters to legal sound sequences in their L1.

Dupoux et al. (1999) conducted experiments to examine perceptually-driven vowel epenthesis in L2 consonant clusters. They tested Japanese and French speakers' perception of word-medial consonant clusters by using two types of ABX discrimination task: an epenthesis contrast (e.g., *abge* – *abuge*) and a vowel length contrast (e.g., *abuge* – *abuuge*). The results showed that Japanese speakers had more difficulty with epenthesis contrast than with the vowel length differences. That is, Japanese speakers, in contrast to French speakers, tended to perceive illusory epenthetic vowels within consonant clusters. Dupoux et al. (1999) argued that perception of illusory epenthetic vowels was due to the influence of native phonotactic constraints on speech perception. They used the term “phonetic illusion” to characterize perceptual adjustments because Japanese speakers heard epenthetic vowels that were not included in the stimuli. In addition, Kabak and Idsardi (2007) found that Korean listeners' perception of consonant clusters was influenced by a

syllable structure violation (e.g., \*[c], \*[j]), rather than a consonantal contact violation (e.g., \*[k.m], \*[l.n]) in their L1.

Moreton (2002) investigated whether various stop-sonorant onset clusters were equally difficult for native English speakers to perceive. The results revealed that there was a bias against the onset /dl-/, but not against /bw-/ even though both onset clusters are illegal in English. He stated that the listeners' linguistic experience was not able to explain the difference in performance because the probability of both sequences were close to zero. He attributed this asymmetry to structural constraints such as the ban on [coronal][coronal] sequences, which did not apply to two adjacent labial consonants.

Sperbeck and Strange (2010) compared Japanese learners of English and native English speakers in a categorial ABX discrimination test of onset consonant clusters and matched counterparts with vowel insertion (e.g., *spani* vs. *sepani*). The results indicated that Japanese speakers' overall accuracy was significantly lower than that of English speakers (72% vs. 98%). Sperbeck and Strange (2010) mentioned that although certain clusters were more difficult for Japanese speakers, sonority-based markedness was not able to reliably predict relative perceptual difficulty of clusters. They claimed that acoustics characteristics of stimuli and native phonological structure seemed to affect the perceptual patterns.

Lee (2011) examined the perceptual interaction between schwa deletion and English phonotactic knowledge by employing a syllable-count task with three different groups of L1 speakers (i.e., native English, Korean and Japanese speakers). The results of both response accuracy and response time showed that only native English speakers were able to distinguish between legal and illegal onset clusters. In addition, Korean speakers displayed a better performance than Japanese speakers although they were not sensitive to English phonotactic constraints. She maintained that language-specific phonotactic restrictions played an important role in perception of onset clusters.

Markedness is related to frequency of a linguistic structure in world languages and is often employed to explain relative difficulty in acquisition. More marked structures are less frequent across languages and harder to acquire than less marked ones. Markedness of onset clusters is systematically constrained by the sonority profile<sup>1</sup>. If a language permits an onset of a small sonority distance (e.g., *lp*), it also

---

<sup>1</sup> Clements (1990) stated the sonority hierarchy in the following order: Stops < Fricatives < Nasals < Liquids < Glides < Vowels (from the least to the most sonorous segments).

allows larger sonority distances (e.g., *pt*, *pn*, *pl*) (Greenberg 1978, Steriade 1982, Selkirk 1984, Clements 1990, Berent et al. 2007, Cardoso 2008, Eckman, 2008). Prior research on the perception of onset clusters suggests that inputs with more marked structures (e.g., *lp*) are more frequently misperceived than less marked structures (e.g., *pl*). The misperception often takes the form of inserting a vowel between two consonants (e.g., *lpa* → *ləpa*; Berent et al. 2007).

Altenberg (2005) examined the acquisition of English onset consonant clusters by native speakers of Spanish in a judgment task, a perception task, and a production task. The results indicated that all L2 learners regardless of their proficiency levels had accurate knowledge of English onset consonant clusters in a judgment task, and that individuals' L1 and L2 could function independently. However, L2 learners were not able to use that information as effectively as native English speakers in perception and production tasks. Altenberg (2005) pointed out that L2 learners' misperception was not due to transfer from L1 as perception accuracy did not differ between the legal and illegal onset clusters. She concluded that L2 perception might be constrained by universal factors such as markedness and sonority principles.

Berent et al. (2007, 2008, 2009), and Lennertz and Berent (2015) attempted to demonstrate universal markedness regarding the sonority profile of onset clusters in the series of experimental studies. Berent et al. (2007) examined native English speakers' perception of onset clusters with small sonority rises (e.g., *bnif*), plateaus (e.g., *bdif*) and falls (e.g., *lbif*). In terms of markedness of sonority profiles of onset clusters, small sonority rises are more marked than large sonority rises, and sonority plateaus are more marked than sonority rises. Also, sonority falls are more marked than plateaus. The results indicated that falling-sonority onset clusters were more likely to derive misperception (e.g., *lba* → *ləba*) than rising-sonority onset clusters (e.g., *bnā* → *bənā*). They suggested that the participants were sensitive to the markedness of onset clusters that were unattested in their L1. However, Berent et al. (2007) also noted that the perceptual illusions of English speakers were partly due to their L1 because Russian speakers whose language allowed these consonant clusters

---

Berent et al. (2007, 2008) examined four types of onset clusters based on the property of sonority. Stops and fricatives are least sonorous ( $s=1$ ). Next are nasals ( $s=2$ ), and then come liquids ( $s=3$ ). Glides are most sonorous ( $s=4$ ). Thus, the consonant cluster in *blif* displays a large rise in sonority ( $s(l) - s(b) = 2$ ), and the cluster in *lbif* shows a sonority fall ( $s(b) - s(l) = -2$ ).

perceived these clusters accurately on most cases. In addition, the preference for obstruent-obstruent (e.g., *bdif*) over liquid-obstruent (e.g., *lbif*) did not extend to all profiles of sonority plateaus and falls.

Berent et al. (2008) examined the role of sonority distance in initial consonant clusters, and found that Korean speakers more frequently misperceived universally more marked onset clusters than less marked one. They presented the participants with four types of syllable-initial consonant clusters, *blif* (large sonority rise), *bnif* (small sonority rise), *bdif* (sonority plateau), and *lbif* (sonority fall), and the matched disyllabic words (e.g., *bəlif*, *bənif*, *bədif*, *ləbif*). When presented with consonant sequences with low sonority distance in monosyllables, Korean speakers tended to misperceive them as disyllables by inserting a schwa-like vowel. They pointed out that listeners' auditory failure to detect the relevant phonetic cues, L2 proficiency, or Korean phonetic and phonological properties were not able to account for these results. They argued that Korean speakers showed the universal phonological knowledge based on sonority distance although their language lacks any type of consonant clusters in both onsets and codas. Yet, in Berent et al. (2008) the participants were not able to discern the differences among all four types of consonant clusters. For example, the Korean speakers did not distinguish between small sonority-rising onsets and large sonority-rising onsets.

Berent et al. (2009) also investigated English speakers' perception of nasal-initial clusters which are unattested in English, and disyllabic counterparts with a schwa between the onset clusters (e.g., *mlif* - *məlif*, *mdif* - *mədif*). They found greater accuracy for rising-sonority and contended that the participants had knowledge of markedness constraints. Moreover, Lennertz and Berent (2015) compared stop-nasal and fricative-nasal onsets (e.g., *pnik* vs. *fnik*) to matched obstruent-obstruent controls (e.g., *ptik* vs. *fsik*), and found that fricative-nasal onsets (e.g., *fnik*) are more often misperceived than stop-nasal onsets (e.g., *pnik*). They argued that linguistic preferences shown in the study reflected the distribution of onset clusters across languages.

Previous research showed that listeners possess universal preferences for the sonority profile of onset clusters. However, the prediction based on sonority-driven onset markedness has not always been verified in previous experiments. As mentioned above, Berent et al. (2007) noticed that the results that Russian speakers whose phonology allowed all the clusters used for the experiment (e.g., *bnif*, *bdif*, *lbif*), correctly perceived these clusters on most trials were partly due to L1

phonotactic constraints. Furthermore, Berent et al. (2008) did not consider L2 proficiency because the participants' levels of English were not modulated in the experiments.

The purpose of the present study is to investigate the role of sonority distance and L2 phonotactic knowledge in perception of onset consonant clusters. This study addresses the questions about whether universal markedness based on sonority distance or L2 phonotactic knowledge affects the perception of onset consonant clusters by three groups, native English listeners, native Korean listeners with higher English proficiency, and native Korean listeners with lower English proficiency. Korean is different from English in terms of consonant sequences. Unlike English, Korean does not permit any consonant clusters in syllable structure. In other words, Korean allows only one consonant in the onset position, while English permits up to three consonants (e.g., *spl-*, *str-*, *skr-*). English allows two-segment onset clusters with a large sonority rise, but not onsets with a small sonority rise, sonority plateau, or sonority fall. Accordingly, English onset clusters are most likely to be the combination of obstruent + sonorant (e.g., *pl-*, *br-*, *tr-*, *dr-*, *fl-*, *gr-*, *sl-*).

More specifically, this study endeavors to answer the following research questions: (1) Are there any differences in perceptual accuracy among the three groups, native English listeners (NE), native Korean listeners with higher English proficiency (NKH), and native Korean listeners with lower English proficiency (NKL)? (2) Are there any differences in perceptual accuracy between monosyllables including consonant clusters and disyllable counterparts for each listener group? (3) Are there any differences in perceptual accuracy among four cluster types based on sonority distance for monosyllables and disyllables, respectively? (4) Are there any differences in response times among the three listener groups, between monosyllables and disyllables, and among the four cluster types? In order to address these questions, response accuracy and response time were measured in a syllable count task.

If universal markedness based on sonority distance operates, all three groups will show the similar perceptual patterns regarding different types of onset clusters. In addition, if listeners are sensitive to universal markedness, their response accuracy will be lowered when marked clusters in monosyllables are presented (e.g., *lpeet*). Conversely, when the marked clusters are repaired by epenthetic vowels in disyllables (e.g., *lupéet*), the accuracy will be higher than the other types of disyllables. On the other hand, if L2 phonotactic knowledge or L2 proficiency has an

effect, the performance of the NKH group will differ from that of the NKL group. Moreover, if native phonotactic knowledge strongly affects listeners' perception, the NE group will be different from the two Korean listener groups.

## 2. Method

### 2.1 Participants

There were three groups of participants, ten native English listeners (NE), ten native Korean listeners with higher English proficiency (NKH), and ten native Korean listeners with lower English proficiency (NKL). The NE group consisted of five males and five females, who were English instructors at a university in Seoul. In the NKH group, there were two males and eight females, who were graduate students majoring in English Education. They were considered to have an advanced level of English based on their self-reported English proficiency, the researcher's observation in class, and their TOEIC (Test of English International Communication) scores (mean = 950). The NKL group consisted of one male and nine females, who were graduate or undergraduate students majoring in Business, Japanese, or Korean Education. Only two of them had taken the TOEIC, and their TOEIC scores were 595 and 675. None of them had lived or studied in an English-speaking country for more than six months. They were considered to have a low level of English proficiency based on self-reporting and the researcher's evaluations during short interviews. Ages of native English listeners range from 27 to 44 while Koreans from 26 to 33.

### 2.2 Stimuli

The experimental materials consisted of 336 nonce words (112 nonce words  $\times$  3 repetitions), and the 112 test items were composed of 56 CCVC monosyllabic and 56 CəCVC disyllabic words. The 56 monosyllabic words had four syllable types based on the sonority distance between two onset clusters: 14 including legal onset clusters (obstruent-liquid sequences: *plieve*), 14 including illegal onset clusters with a small sonority rising (obstruent-nasal sequences: *pnese*), 14 including illegal onset clusters with a sonority plateau (obstruent-obstruent sequences: *ptack*), and 14 including illegal onset clusters with a sonority fall (liquid-obstruent sequences: *lpeet*). These nonce words conformed to English phonotactic rules except the onset clusters. These

words were matched to disyllabic counterparts, and the only difference between monosyllabic and disyllabic words was whether or not the schwa was inserted between two onset consonants. The disyllabic words were stressed on the second syllable (e.g., *pulieve*, *punése*, *putáck*, *lupéet*). All the test items were recorded by a phonetically trained native speaker of English. The recorded items were checked by both eye and ear using the waveform and spectrogram of *Praat* (version 5.4.04) regarding the vowel presence or absence between two onset clusters. A total of 10,080 responses (336 tokens x 30 participants) were obtained. Table 1 shows some examples of stimuli based on the sonority profile of onset clusters. The complete list of stimuli used in the present study is presented in the Appendix.

**Table 1. Sonority profile of onset clusters in some example items**

Cluster type	Legal syllables	Illegal syllables		
	Large sonority rise	Small sonority rise	Sonority plateau	Sonority Fall
Monosyllables	<i>plieve</i> , <i>troon</i>	<i>pnese</i> , <i>tmote</i>	<i>ptack</i> , <i>kpel</i>	<i>lpeet</i> , <i>rtel</i>
Disyllables	<i>pulieve</i> , <i>turóon</i>	<i>punése</i> , <i>tumóte</i>	<i>putáck</i> , <i>kupél</i>	<i>lupéet</i> , <i>rutél</i>

### 2.3 Procedure

A syllable count task was used to investigate a perceptual epenthesis effect. The experiment was carried out individually in a quiet room using a laptop computer. The stimuli were presented to each participant through a syllable count task scripted in *Praat* with a low-noise headset (Plantronics DSP-500). Participants started the trial by pressing any keys, which resulted in the presentation of an auditory stimulus. The participants were asked to listen to each stimulus and determine whether the presented stimulus had one syllable or two syllables by clicking on ‘1’ or ‘2’ on the screen, respectively. The stimuli were presented in individually randomized order. Before the actual experiment, each participant completed a practice session which included 10 items that were real English words to ensure familiarity with task. The researcher explained about the task in detail in order to make sure that the participants knew how to count syllables. Feedback on response accuracy was provided during the practice session only. Both the practice and test sessions were implemented using *Praat* (version 5.4.04). The mean time to complete the task was 25 minutes. Response times were reported from the onset of the stimulus.



The following figure shows the computer screen used during the experiment.

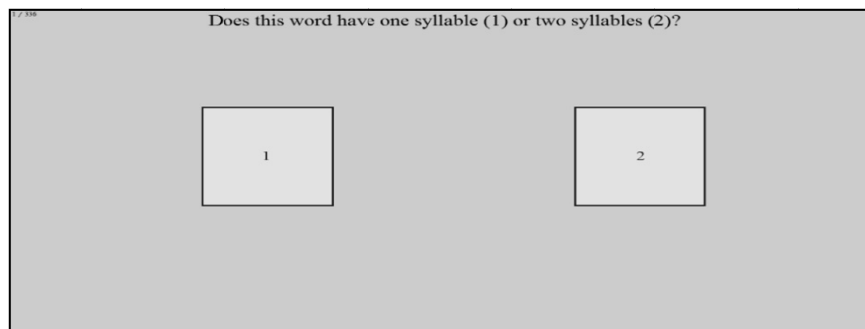


Figure 1. Computer screen prompted by the *Praat* software

### 3. Results

#### 3.1 Response accuracy

Mean response accuracy for three listener groups is provided in Figure 2. The average correct percentage obtained from all participants were 77.1% for the onset clusters with a large sonority rise, 66.9% for clusters with a small sonority rise, 59.7% for clusters with a sonority plateau, and 58.2% for clusters with a sonority fall.

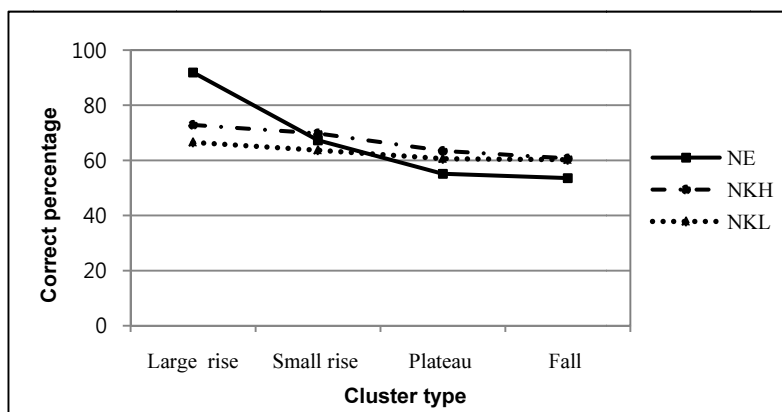
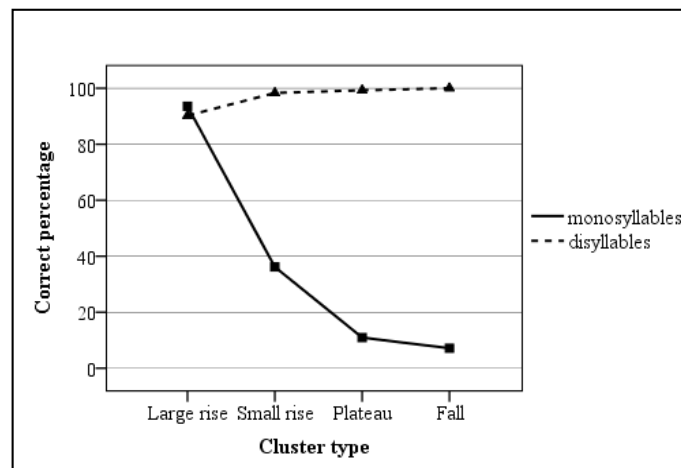


Figure 2. Mean correct percentage by three listener groups

In order to investigate the effects of cluster type and listener group, a repeated measures ANOVA (analysis of variance) was conducted with cluster type and listener group as factors (4 cluster types x 3 listener groups) and accuracy percentage as the dependent variable. All the analyses were performed using SPSS Statistics 21. There were significant effects of cluster type [ $F(3, 26) = 65.29, p < .0001$ ] and the interaction between cluster type and listener group [ $F(6, 26) = 19.01, p < .0001$ ]. The native English speakers performed much better than the other two Korean groups as can be seen in Figure 2. Post-hoc pairwise comparisons (LSD test) of listener groups indicated that the response accuracy of native English listeners (NE) was significantly different from that of native Korean listeners with lower proficiency (NKL) ( $p < .05$ ). However, the response accuracy of native English listeners was not significantly different from that of Korean listeners with higher proficiency (NKH) ( $p = .893$ ).

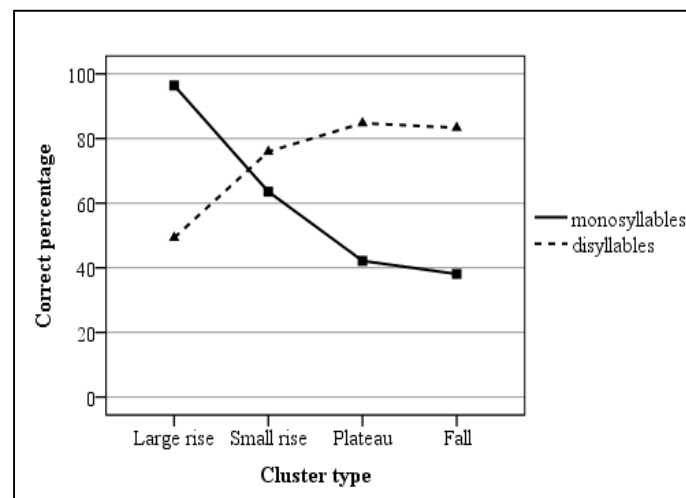
In addition, a repeated measures ANOVA was conducted for the three listener groups separately in order to determine whether there were syllable number and cluster type effects (2 syllables x 4 cluster types). For the NE group, there was a significant effect of syllable number [ $F(1, 36) = 1460.12, p < .0001$ ] and the interaction between syllable number and cluster type [ $F(3, 36) = 199.35, p < .0001$ ]. The English listeners perceived disyllabic inputs more accurately than monosyllabic input as shown in Figure 3.



**Figure 3. Mean response accuracy of monosyllables and disyllables by the NE group**

Furthermore, the effect of cluster type was investigated separately for monosyllabic and disyllabic inputs. With monosyllabic inputs, English listeners yielded more accurate responses as the sonority distance of onset clusters increased. They distinguished between all three cluster type ( $p < .0001$  for all pairwise comparisons among large rises, small rises, and plateaus), whereas there was no difference between plateaus and falls. However, for disyllabic inputs, onsets with large sonority rises were different from the other cluster types ( $p < .0001$  for all pairwise comparisons between large rises and the other three cluster types), whereas there was no difference among the other three cluster types (i.e. small rises, plateaus, and falls).

For the NKH group, there was a significant effect of syllable number [ $F(1, 36) = 6.14, p < .02$ ] and a significant effect of the interaction between syllable number and cluster type [ $F(3, 36) = 16.08, p < .0001$ ]. Similar to English listeners, Korean listeners with higher proficiency perceived disyllabic inputs more accurately than monosyllabic inputs for all onsets except those with a large rise as shown in Figure 4.

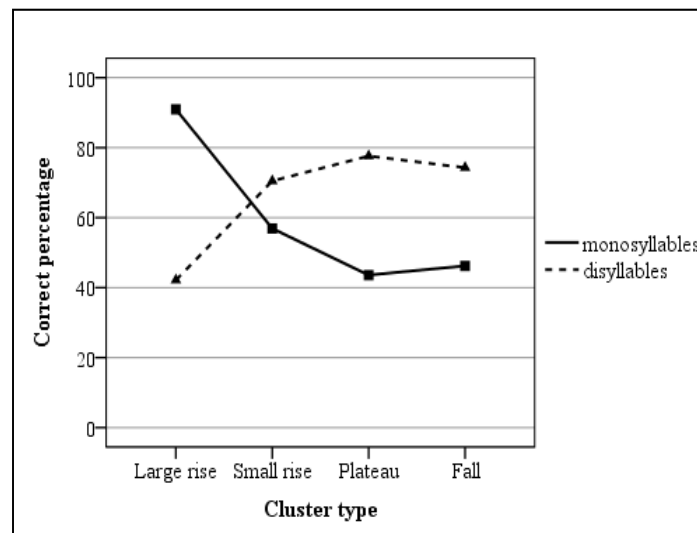


**Figure 4. Mean response accuracy of monosyllables and disyllables by the NKH group**

In addition, the effect of onset cluster type was investigated separately for monosyllabic and disyllabic inputs. With monosyllabic inputs, the NKH group yielded less accurate responses as the markedness of onset clusters increased. They

distinguished between large rises and the other three cluster types ( $p < .005$  between large rises and small rises;  $p < .0001$  for both between large rises and plateaus, and between large rises and falls). Their perceptions between small rises and the other two cluster types also showed a significant effect ( $p < .05$  for pairwise comparisons between small rises and plateaus, and between small rises and falls). However, there was no difference between plateaus and falls. With disyllabic inputs, they showed the opposite patterns. Their correct responses were lower with onsets with large sonority rises than with the other three cluster types ( $p < .0001$  for all pairwise comparisons between large rises and the other three cluster types), whereas there were no differences among the other three cluster types (i.e. small rises, plateaus, and falls).

For the NKL group, there was no significant effect of syllable number [ $F(1, 36) = 1.14$ ,  $p > .05$ ], but there was a significant effect of the interaction between syllable number and cluster type [ $F(3, 36) = 9.12$ ,  $p < .0001$ ]. Figure 5 displays that the NKL group perceived disyllabic inputs more accurately than monosyllabic inputs including onsets with small rises and plateaus and falls, but statistical differences were not revealed in terms of syllable number.



**Figure 5. Mean response accuracy of monosyllables and disyllables by the NKL group**

Furthermore, the effect of cluster type was investigated separately for monosyllabic and disyllabic inputs. With monosyllabic inputs, the NKL group yielded less accurate responses as the sonority distance of onset clusters decreased. They distinguished between large rises and the other three cluster types ( $p < .005$  between large rises and small rises;  $p < .0001$  for both between large rises and plateaus, and between large rises and falls), whereas there were no differences among the other three cluster types (i.e. small rises, plateaus, and falls). With disyllabic inputs, their response accuracy was lower with onsets with large sonority rises than with the other three cluster types ( $p < .01$  for all pairwise comparisons between large rises and the other three cluster types), while response accuracy among the other three cluster types did not differ with one another. Table 2 summarizes the statistical results of participants' mean accuracy in terms of listener group, syllable number and cluster type.

**Table 2. Summary of statistical results of mean accuracy by three listener groups**

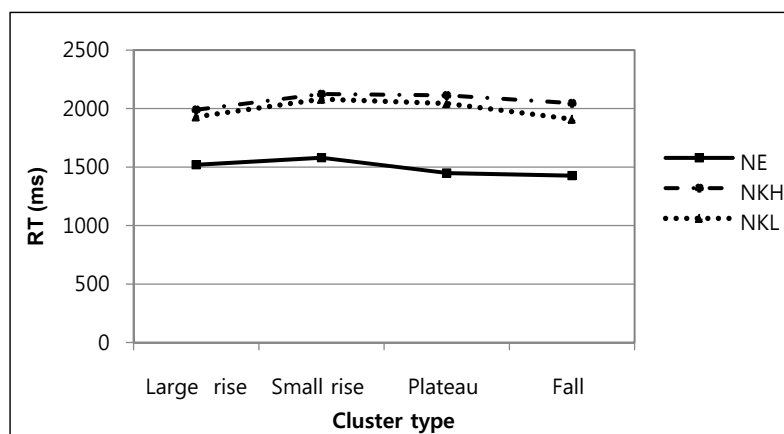
		NE	NKH	NKL
Syllable number		Monosyllables $\neq$ Disyllables ( $p < .0001$ )	Monosyllables $\neq$ Disyllables ( $p < .0001$ )	Monosyllables = Disyllables ( $p > .05$ )
Cluster type	Mono-syllables	LR $\neq$ SR ( $p < .0001$ )	LR $\neq$ SR ( $p < .005$ )	LR $\neq$ SR ( $p < .005$ )
		LR $\neq$ P ( $p < .0001$ )	LR $\neq$ P ( $p < .0001$ )	LR $\neq$ P ( $p < .0001$ )
		LR $\neq$ F ( $p < .0001$ )	LR $\neq$ F ( $p < .0001$ )	LR $\neq$ F ( $p < .0001$ )
		SR $\neq$ P ( $p < .0001$ )	SR $\neq$ P ( $p < .05$ )	SR = P ( $p > .05$ )
		SR $\neq$ F ( $p < .0001$ )	SR $\neq$ F ( $p < .05$ )	SR = F ( $p > .05$ )
		P = F ( $p > .05$ )	P = F ( $p > .05$ )	P = F ( $p > .05$ )
	Di-syllables	LR $\neq$ SR ( $p < .0001$ )	LR $\neq$ SR ( $p < .0001$ )	LR $\neq$ SR ( $p < .01$ )
		LR $\neq$ P ( $p < .0001$ )	LR $\neq$ P ( $p < .0001$ )	LR $\neq$ P ( $p < .01$ )
		LR $\neq$ F ( $p < .0001$ )	LR $\neq$ F ( $p < .0001$ )	LR $\neq$ F ( $p < .01$ )
		SR = P ( $p > .05$ )	SR = P ( $p > .05$ )	SR = P ( $p > .05$ )
		SR = F ( $p > .05$ )	SR = F ( $p > .05$ )	SR = F ( $p > .05$ )
		P = F ( $p > .05$ )	P = F ( $p > .05$ )	P = F ( $p > .05$ )

Note: LR= large rises, SR=small rises, P=plateaus, F=falls

As shown in Table 2, with the syllable number, the NE and NKH groups showed a similar pattern. Both groups performed better with disyllables than monosyllables, whereas NKL did not show a statistical difference in terms of syllable number. Furthermore, in monosyllabic inputs it was shown that onset sonority distance modulated the perception of both NE and NKH groups except sonority plateaus and falls. Pairwise comparisons showed that as the sonority distance decreased, the listeners tended to misperceive the onset clusters as two syllables. However, the perception of the NKL group did not reveal this tendency. In addition, all the groups did not show this tendency based on sonority distance in disyllabic items. With disyllabic inputs, the differences in terms of sonority distance were only revealed between large rises and the other three cluster types for all three groups. They perceived the disyllabic inputs with a large sonority rise less accurately than all the other types of disyllabic inputs, but there were not differences among the other types of disyllabic inputs.

### 3.2 Response time (RT)

In order to analyze the mean response time a repeated measures ANOVA (4 cluster types x 3 listener groups) was conducted. There was a significant effect of cluster type [ $F(3, 26) = 4.42, p < .05$ ], but the interaction between cluster type and listener group did not yield a significant effect. Post-hoc pairwise comparisons (LSD test) of listener groups indicated that the response time of the NE group was significantly different from that of both NKH and NKL groups. The NE group responded much faster than the other two Korean groups. However, the two Korean groups were not different from each other in terms of RT as shown in Figure 6.



**Figure 6. Mean response time by three listener groups**

In addition, a repeated measures ANOVA was conducted for the three listener groups separately in order to determine whether there were syllable number and cluster type effects (2 syllables  $\times$  4 cluster types). For the NE group, there was a significant effect of syllable number [ $F(1, 36) = 26.38, p < .0001$ ], but there was no significant effect of the interaction between syllable number and cluster type. The NE group was much slower to perceive monosyllabic words than disyllabic ones. Furthermore, the effect of cluster type was investigated separately for monosyllabic and disyllabic inputs. No significant effects were found for either monosyllabic or disyllabic inputs.

For the NKH group, a repeated measures ANOVA was conducted in order to determine whether there were syllable number and cluster type effects (2 syllables  $\times$  4 cluster types). An effect of syllable number was not found [ $F(1, 36) = 2.41, p > .05$ ], and there was a marginally significant effect of the interaction between syllable number and cluster type [ $F(3, 36) = 4.06, p < .05$ ]. In addition, the effect of cluster type was investigated separately for monosyllabic and disyllabic inputs. No significant effects were found for either monosyllabic or disyllabic words.

For the NKL group, a repeated measures ANOVA was conducted in order to determine whether there were syllable number and cluster type effects (2 syllables  $\times$  4 cluster types). An effect of syllable number was not found [ $F(1, 36) = .138, p > .05$ ], nor was any effect of the interaction between syllable number and cluster type found [ $F(3, 36) = 2.12, p > .05$ ]. Additionally, the effect of cluster type was investigated separately for monosyllabic and disyllabic inputs. No significant effects were found for

either monosyllabic or disyllabic words. Table 3 displays the average response time of four cluster types in monosyllables and disyllables.

**Table 3. Mean response time (ms) for monosyllabic and disyllabic inputs by three listener groups (standard deviation in parentheses)**

Cluster type	NE		NKH		NKL	
	Mono-syllables	Di-syllables	Mono-syllables	Di-syllables	Mono-syllables	Di-syllables
Large rise	1587 (519)	1453 (438)	1807 (348)	2173 (407)	1883 (485)	1976 (457)
Small rise	1778 (703)	1519 (585)	2082 (489)	2085 (336)	2147 (665)	2031 (558)
Plateau	1567 (726)	1331 (384)	2128 (379)	2098 (456)	2106 (615)	1980. (506)
Fall	1491 (596)	1371 (624)	2067 (367)	2027 (427)	1862 (537)	1949 (478)

#### 4. Discussion and conclusion

The present study used a syllable count task to investigate whether three groups, native English listeners (NE), native Korean listeners with higher English proficiency (NKH), and native Korean listeners with lower English proficiency (NKL), show the same perceptual patterns in terms of sonority distance in onset clusters. If listeners are sensitive to universal markedness, all three groups of listeners will show the same patterns. That is, their response accuracy will be lowered if clusters are marked in monosyllables (e.g., *rpect*). Conversely, if the marked clusters are repaired by epenthetic vowels in disyllables (e.g., *rupéct*), their response accuracy will be higher than other types of disyllables.

For response accuracy, when three groups are compared, NE and NKH groups showed the similar perceptual patterns in terms of syllable number. Both NE and NKH groups showed the sensitivity to syllable number, but the NKL group was not sensitive to syllable number. When monosyllables and disyllables are compared for each listener group, the NE and NKH groups performed better on disyllabic counterparts than monosyllabic inputs, whereas the NKL group did not differentiate monosyllables from disyllables.

In the analyses of sonority distance in onset clusters, both NE and NKH groups were



more sensitive to the sonority distance in monosyllabic inputs than disyllabic counterparts. For monosyllabic inputs, they were more accurate as the sonority distance increased, but they were not able to distinguish between sonority plateaus and falls. On the other hand, the NKL group only distinguished between large sonority rises and the other three cluster types (i.e., small rises, plateaus, and falls), whereas there were no differences among the other cluster types. It seems that the NKL group's performance on monosyllables including small sonority rises, plateaus and falls was at a chance level. For disyllabic counterparts, all three groups displayed similar patterns. All the listeners differentiated disyllabic items involving a large sonority rise from the ones involving the other cluster types (i.e., small rises, plateaus, and falls), but there were no differences among the other cluster types. Specifically, two Korean groups were much less accurate responding to the disyllabic counterparts of monosyllables with large sonority rises (e.g., *pulieve*) than the other types of disyllables (e.g., *punése*, *putáck*, *lupéet*). This result indicates that the Korean listeners differentiated legal English clusters from illegal ones although they did not discern an epenthetic vowel between two consonants. The distinction between legal clusters and illegal ones cannot be explained by L1 phonotactic constraints because in Korean no consonant clusters are allowed. Since Korean listeners were familiar with English phonotactic constraints to some degree, they tended to perceive disyllabic inputs involving a large sonority rise as monosyllables. Also, the prediction that disyllabic counterparts of monosyllables with more marked clusters would be perceived more accurately was not born out.

Sonority distance of onset clusters is not able to explain all the results of the present study. For monosyllabic items, as sonority distance decreased, the responses were less accurate, indicating that the listeners frequently misperceived monosyllabic inputs as two syllables. However, the effect of sonority distance was not shown between two cluster types, plateaus and falls, for all three listener groups. Further, the NKL group did not differentiate between small rises and plateaus, or between plateaus and falls. Moreover, the perceptual patterns based on the sonority distance shown in monosyllabic inputs were not shown in disyllabic inputs.

The present results are partially consistent with those of previous studies (Berent et al. 2008, Lee 2011). Berent et al. (2008) argued that Korean speakers whose language lacks initial consonant clusters demonstrated preference for the large sonority distance over the small sonority distance among onset clusters. However, in their study the Korean speakers' responses did not differ between large sonority rises and small rises or between small rises and plateaus in both monosyllables and disyllables. Thus, in

their study, the Korean speakers showed a tendency of less accurate responses as the sonority distance decreased, yet they were not able to discern the differences among some cluster types. Lee (2011) found that English speakers' perceptions of legal English onset clusters were significantly more accurate and faster than those of illegal clusters. Korean and Japanese speakers showed partial effects of onset markedness, but they were not sensitive to English phonotactic constraints. She mentioned that Korean speakers' better performance than Japanese speakers may be due to the Korean speakers' overall higher English proficiency in comparison to the Japanese speakers.<sup>2</sup>

For response time (RT), the NE group was faster than both NKH and NKL groups. Also, only the NE group showed an effect of syllable number. Their response time was significantly slower for monosyllables than for disyllabic counterparts. However, Korean listeners' response time did not reveal a sensitivity to syllable number. Furthermore, for all three groups a sensitivity to the sonority distance in onset clusters was not found either in monosyllables or disyllables.

The participants in this study seemed to partially reflect abstract phonological knowledge concerning the markedness based on sonority distance of onset clusters. In terms of response accuracy, both NE and NKH groups displayed more sensitivity to sonority distance in onset clusters than the NKL group, but neither group discriminated between sonority plateaus and falls. Further, the NKL group was not able to perceive the differences among small sonority rises, plateaus and falls, although they discriminated between legal onset clusters and illegal ones. Considering that the NKH group has more experience with English and higher English proficiency than the NKL group, different misperception patterns shown in the three groups are partly due to their linguistic experience in English. Universal

---

<sup>2</sup> One reviewer raised a question about whether native Korean listeners' performance would be as bad as Japanese listeners if Korean listeners took an ABX discrimination test involving monosyllables and disyllables (e.g., *plieve* vs. *pulieve*). In the previous research (Dupoux et al. 1999, Sperbeck and Strange 2010) native Japanese listeners' perceptual accuracy in an ABX task was significantly lower than that of native French or English listeners. Further, Durvasula and Kahng (2015) conducted an ABX task with native Korean and English listeners to investigate a perceptual epenthesis effect (e.g., *esma* vs. *esima*). They found that English listeners had higher accuracy than Korean listeners. However, no previous research compared native Japanese and Korean listeners in an ABX task. Thus, this comparison seems to be worth pursuing in the future study.

markedness could be restricted by English phonotactic knowledge in Korean listeners' interlanguage and L2 proficiency. Thus, the present results suggest that Korean L2 listeners' ability to distinguish onset clusters does not correspond to sonority-based markedness. Other factors such as L2 phonotactic knowledge or acoustic characteristics of the stimuli may have played a role. The question of what acoustic properties of stimuli affects L2 listeners' perception remains unanswered in the current study.

#### Appendix. Stimuli used in the syllable count task

Legal syllables in English		Illegal syllables in English					
Large sonority rise		Small sonority rise		Sonority plateau		Sonority fall	
(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
plieve	pulieve	pnese	punése	pkome	pukóme	rpect	rupéct
proot	puróot	pmong	pumóng	ptack	putáck	lpeet	lupéet
troon	turóon	tnese	tunése	tkome	tukóme	ltack	lutáck
troov	turóov	tmote	tumóte	tpeel	tupéel	rtel	rutél
kloon	kulóon	knuf	kunúf	kpel	kupél	lkant	lukánt
krole	kuróle	kmoov	kumóov	ktel	kutél	rkum	rukúm
blect	buléct	bnuf	bunúf	bdem	budém	lbat	lubát
bract	buréct	bmoov	bumóov	bgeev	bugéev	rbet	rubét
dret	durét	dnese	dunése	dbet	dubét	ldict	ludíct
dreen	duréen	dmong	dumóng	dgeen	dugéen	rduke	rudúke
gloon	gulóon	gnul	gunúl	gbove	gubóve	lgeev	lugéev
gret	gurét	gmute	gumúte	gdeem	gudéem	rgeen	rugéen
fless	fuléss	fnul	funúl	fsent	fusént	lfect	luféct
freen	furéen	fmute	fumúte	fzent	fuzént	rfeet	ruféet

Note: (1) = monosyllabic words, (2) = disyllabic words

## REFERENCES

- ALBRIGHT, ADAM. 2007. *Natural classes are not enough: biased generalization in novel onset clusters*. Manuscript, Cambridge, MA: MIT.
- ALTENBERG, EVELYN. 2005. The judgment, perception, and production of consonant clusters in a second language. *IRAL* 43, 53-80.
- BERENT, IRIS, TRACE LENNERTZ, JONGHO JUN, MIGUEL A. MORENO, and PAUL SMOLENSKY. 2008. Language universals in human brains. *Proceedings of the National Academy of Sciences* 105, 5321-5325.
- BERENT, IRIS, TRACY LENNERTZ, PAUL SMOLENSKY and VERED VAKNIN. 2009. Listeners' knowledge of phonological universals: Evidence from nasal clusters. *Phonology* 26, 75-108.
- BERENT, IRIS, DONCA STERIADE, TRACY LENNERTZ and VERED VAKNIN. 2007. What we know about what we have never heard: Evidence from perceptual illusions. *Cognition* 104, 591-630.
- CARDOSO, WALCIR. 2008. The development of sC onset clusters in interlanguage: Markedness vs. frequency effects. *Proceedings of the 9th Generative Approaches to Second Language Acquisition Conference (GASLA 2007)*, 15-29.
- CLEMENTS, GEORGE N. 1990. The role of the sonority cycle in core syllabification. In John Kingston and Mary E. Beckman (eds.). *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, 283-333. Cambridge: Cambridge University Press.
- DUPOUX, EMMANUEL, KAZUHIKO KAKEHI, YUKI HIROSE, CHRISTOPHE PALLIER and JACQUES MEHLER. 1999. Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perceptual and Performance* 25, 1568-1578.
- DURVASULA, KARTHIK and JIMIN KAHNG. 2015. Illusory vowels in perceptual epenthesis: The role of phonological alterations. *Phonology* 32, 385-416.
- ECKMAN, FRED. 2008. Typological markedness and second language phonology. In Jette G. Hansen Edwards and Mary. L. Zampini (ed.). *Phonology and Second Language Acquisition*, 95-115. Philadelphia: John Benjamins.
- GREENBERG, JOSEPH H. 1978. Some generalizations concerning initial and final consonant clusters. In Edith A. Moravcsik (ed.). *Universals of Human Language: Method and Theory Vol. 2*, 243-279. Stanford, CA: Stanford University Press.
- HALL, A. PIERRE, JUAN SEGUI, ULI FRAUENFELDER, and CHRISTINE MEUNIER. 1998.

- Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human perception and Performance* 24.2, 592-608.
- KABAK, BARIS and WILLIAM J. IDSARDI. 2007. Perceptual distortions in the adaptation of English consonant clusters: Syllabus structure or consonantal contact constraints? *Language and Speech* 50.1, 23-52.
- LEE, SHINSOOK. 2011. The perceptual interaction between schwa deletion and English phonotactic constraints. *Korean Journal of Linguistics* 36.3, 699-733.
- LENNERTZ, TRACY and IRIS BERENT. 2015. On the sonority levels of fricatives and stops. *The Mental Lexicon* 10.1, 88-132.
- MIKHAYLOVA, ANNA. 2009. L2 nonword recognition and phonotactic constraints. *University of Pennsylvania Working Papers in Linguistics* 15.1, 145-152.
- MORETON, ELLIOTT. 2002. Structural constraints in the perception of English stop-sonorant clusters. *Cognition* 84, 55-71.
- PITT, MARK A. 1998. Phonological processes and the perception of phonotactically illegal consonant clusters. *Perception and Psychophysics* 60, 941-951.
- SELKIRK, ELISABETH O. 1984. On the major class features and syllable theory. In Mark Aronoff and Richard T. Oehrle (eds.). *Language Sound Structure*, 107-136. Cambridge, MA: MIT Press.
- SPERBECK, MIEKO and WINIFRED STRANGE. 2010. The perception of complex onsets in English. *University of Pennsylvania Working Papers in Linguistics* 16.1, 194-204.
- STERIADE, DONCA. 1982. *Greek Prosodies and the Nature of Syllabification*. MA Thesis. MIT.

Eunkyug Sung  
 Department of English  
 Cyber Hankuk University of Foreign Studies  
 107 Imun-ro, Dongdaemun-gu  
 Seoul, 130-791, Korea  
 e-mail: eks@cufs.ac.kr

received: November 17, 2016

revised: December 14, 2016

accepted: December 18, 2016