Linguistic Research 42(3): 711-737
DOI: 10.17250/khisli.42.3.202512.009

Perception and lexical representation of the innovative variant of /lp/ in Korean*

Tae-Hwan Choi** · **Jeong-Im Han*****
(Catholic Kwandong University · Konkuk University)

Choi, Tae-Hwan and Jeong-Im Han. 2025. Perception and lexical representation of the innovative variant of /lp/ in Korean. *Linguistic Research* 42(3): 711-737. As Korean phonotactic constraints do not allow consonant clusters in the syllable coda position, tautosyllabic clusters are realized as a single (first or second) consonant of a cluster when followed by another consonant. However, recent phonetic studies have reported that clusters with a lateral and obstruent have been undergoing a change in preserving the first consonant, or even both consonants. Across two experiments, this study investigated Korean speakers' perceptual and lexical encoding of an innovative variant (i.e., [lp]) of the cluster /lp/ where the on-going change is most noticeable. Experiment 1 involved a speeded AX discrimination task to examine the perception of the innovative variant. Results revealed that Korean speakers had difficulty discriminating word pairs when one member of the pair contained [lp], suggesting that this variant is not categorically perceived. In Experiment 2, a long-term repetition priming lexical decision task was performed to observe the storage of the innovative variant in the lexicon. Significant priming was found for this variant, which was equivalent to that of the identity pairs, indicating that the innovative variant [lp] is stored as a separate category in long-term memory. **(Catholic Kwandong University · Konkuk University)**

Keywords Korean phonotactics, consonant clusters, variants of /lp/, AX discrimination, long-term repetition priming

* We would like to thank two anonymous reviewers for their valuable comments and suggestions. This research was financially supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2022S1A5A2A01038289).
** First Author
*** Corresponding author

1. Introduction

There is ample evidence that adult listeners can cope with substantial variation that they are confronted with when recognizing spoken words, whether it is phonologically regular, context-dependent, as in assimilation or deletion (Lahiri and Marslen-Wilson 1991; Gaskell and Marslen-Wilson 1996, 1998; Gow 2001, 2002, 2003; Mitterer and Blomert 2003; Sumner and Samuel 2005), or arbitrary, as in mispronunciation (Connine et al. 1993; Frauenfelder et al. 2001). For example, the final consonant of the word *right* in English is pronounced [t] when it appears alone, but [p] when the immediately following word begins with a labial stop, as in *right berries*. This may lead to a mismatch between the speech signal ([p]) and canonical lexical form (/t/). However, this mismatch is accommodated by listeners, as it is not difficult to retrieve the canonical form when recognizing words. The perception of phonological variants is directly related to how listeners represent these variants in their mental lexicons. Although many studies have delineated how speech signals activate lexical representations, two possibilities have been considered. First, only canonical variants are stored in the mental lexicon. Dell (1986) and Levelt et al. (1999) argue that only the canonical form is assumed to be in the lexicon as a lexeme, while other variants are derived from the canonical form through phonological rules. Although there are several phonological variants (e.g., [t] and [p] for the final stop in *right*) on the surface, only one phonological representation is stored in the mental lexicon (e.g., /t/). Thus, the stored form is sometimes abstract and different from the actual pronunciation. Second, multiple representations of the variants are stored in the lexicon. In the episodic model of representation (Goldinger 1998; Ranbom and Connine 2007; Bürki et al. 2010, 2011), for example, representations are fine-grained and subtle details about the particular pronunciation are stored in the lexicon. Every time a variant is produced and heard, a particular form is stored in the lexicon. Thus, there is no apparent benefit for canonical forms as both canonical and non-canonical forms are stored in the mental lexicon. The current study focused on a particular variant in Korean that shows context-independent variation from an on-going sound change. We examined its effects, if any, on speech perception and lexical representations.

1.1 Korean consonant clusters

Korean does not allow consonant clusters to occur in a syllable coda as well as onset positions. This renders tautosyllabic clusters to realize only one of the two consonants in clusters before a word boundary or another consonant ([sam] from /salm/ 'life'; [sam.to] from /salm-to/ 'life also'), while when followed by a vowel, the second consonant of the cluster is resyllabified as an onset of the following syllable ([sal.mi] from /salm/ + /i/ 'nominative marker'). The former has been called 'consonant cluster simplification (CCS)' in the literature, focusing on the choice of consonant to be realized on the surface (Kim-Renaud 1974; Cho 1990; Iverson and Lee 1994; Jun 1998; Sohn 1999; Shin et al. 2012, among others). Shin et al. (2012), for example, proposed that the CCS pattern differs by cluster type; namely, between the two consonants of the clusters, the consonant with the feature [+sonorant] (/ls, nc, lh, nh, lt^h/) or that with the feature [-coronal] (/ps, ks/) is realized on the surface. However, clusters that consist of a lateral and obstruent (/lk, lp, lp^h, lt^h/) show variability; in /lk, lp^h/, the second consonant is preserved, while the first consonant is preserved in /lp/, though both clusters have the feature [+sonorant] in the first consonant and [-coronal] in the second. It has also been reported that the realization of clusters with /l/ plus an obstruent differs by regional dialect (Whitman 1985; Cho 1990) such that [l] is preserved in the Gyeongsang dialect, while the following obstruent is realized in the standard Seoul dialect.

Recent phonetic studies (Cho 1990; Cho and Kim 2009; Kim and Kang 2021) have demonstrated that CCS has been changing over the years to preserve the first consonant of the cluster (i.e., [l] from /lk/) more often. Moreover, the ratio preserving both consonants (e.g., [lk] from /lk/) in the cluster has been increasing.[1] Younger speakers pronounce [l] more frequently than the following consonant ([p, k]) in /lp, lk/ in everyday speech, even though the canonical form is taught to be pronounced in school. Based on a large-scale spontaneous speech corpus, Kwon et al. (2024) found that 1) in /lp/, realization of [l] or [lp] reached a ceiling regardless of the age or regional dialect of the speakers, and 2) in /lk/, younger speakers (born after 1970) showed a significantly higher rate of [l] realization ([l] or [lk]), regardless of dialect

1 The realization of both consonants in /lp, lk/ was recognized in traditional analysis (Kim-Renaud 1974) but this was just based on the author's impressionistic description, not verified by experimental evidence.

(84% for standard Seoul dialect and 93% for Gyeongsang dialect), whereas those born before 1970 showed less preservation of [l] (56% for standard Seoul dialect and 71% for Gyeongsang dialect). Based on these results, Kwon et al. (2024) argued that there may be a sound change in progress toward realizing the first consonant of clusters consisting of /l/ plus /p/ or /k/, namely [l] or even [lk, lp]. They also noted that the innovation initially took place in /lp/, where the process appears almost regular today, while it could have been extended to other clusters where the rate of innovative forms appears to be vigorously increasing among younger speakers. This change is partly reflected in the revised regulations of the Standard Pronunciation of Korean (2018) (The National Institute of the Korean Language, Dictionary of Standard Korean), "/lk/ is simplified to [k], while /lp/, to [l], except the verb /palp-ta/ 'to step on' where [p] is preserved in the coda" (Regulations #10, #11). As such, CCS is dependent on specific cluster types and lexical items.

This variation is distinguished from those discussed in earlier studies, as it is not arbitrary, as in accidental mispronunciations, nor context-dependent, as in assimilation. The realization of the three variants (e.g., [l], [p], or [lp] for /lp/) is not predictable in phonological contexts; they exhibit free variation without a triggering context. Coenen et al. (2001) showed that, unlike regular, context-dependent variations such as place assimilation, context-independent variation can incur a cost in processing. Furthermore, between the two variants of recent changes, preservation of the first consonant ([l]) and preservation of both consonants (e.g., [lp] for /lp/), the first type of change is likely to be perceptually motivated because the first consonant is perceptually more salient with the transitional cue from the previous vowel and the weak formant of [l] compared to the second consonant, [p] or [k] (Steriade 1997; Jun 1998). However, preserving both consonants appears unnatural because it violates the Korean phonotactic constraint of prohibiting consonant clusters in the syllable coda position. Thus, the realization of both consonants in the cluster raises the possibility that when native Korean speakers produce both consonants, they may not produce the two consonants in the cluster as separate segments with independent timing units but may realize both consonants within a single timing unit as in prenasalized stops or complex segments.

1.2 The present study

The present study was concerned with how native Korean speakers perceive and lexically encode the variant with both consonants as well as the other two variants of the cluster with a lateral and obstruent, focusing on /lp/, where the on-going change is most noticeable. We first examined how native Korean listeners perceived spoken words with three variants of the /lp/ cluster using a speeded AX discrimination task, particularly with the purpose of evaluating whether the variants realized with the first ([l]) or second consonant ([p]) and with both consonants ([lp]) are distinguished by Korean listeners. The discrimination task used two distinct Inter-Stimulus-Intervals (ISIs) to induce two putative levels of speech processing: a 100-ms ISI appeared to induce auditory-level processing, while a 1000-ms ISI appeared to induce a phonemic-level analysis of the speech signal. The choice of the two ISIs was based on previous studies (Fox 1984; Werker and Logan 1985; Polka 1991). For example, Werker and Logan (1985) used three ISIs (250 ms, 500 ms, and 1500 ms), each of which was expected to elicit levels of auditory, phonetic, and phonemic processing. As such, a very short 100-ms ISI and a longer 1000-ms ISI were expected to make listeners rely on auditory and more abstract phonological (phonemic) representations, respectively. Second, we examined how Korean listeners encoded the variant realized with both consonants ([lp]) at the lexical level, using a long-term repetition priming task. The purpose of this task was to evaluate how this innovative variant is represented in the long term. In particular, we were interested in whether the innovative variant preserving both consonants of the cluster may have a single abstract representation (e.g., /lp/), or whether it is stored separately (e.g., /l/, /p/, and /lp/).

Regarding perception, we expected native Korean listeners to successfully discriminate variants based on the results of recent phonetic studies reporting a preference for the first consonant and often both consonants of the cluster. However, if the realization of both consonants occurs within a timing unit, Korean listeners may have more difficulty discriminating [l] and [lp] than the other pairs (e.g., [l] and [p]). The realization of both consonants in /lp/ may vary across the ISI condition: two single variants, [l] and [p], which are independent phonemes in Korean, would be well distinguished regardless of the ISI condition. However, the distinction between these phonemes and [lp] would be easy in both auditory (ISI = 100 ms) and more

abstract phonemic levels (ISI = 1000 ms) if Korean speakers have this variant as an independent category, whereas if they do not have it as a distinct category, it would be distinguishable at the auditory level but significantly less distinguishable at the abstract phonemic level. Regarding long-term representation, we expected that if only the abstract representation is maintained in memory, the "canonical" (standard; prescribed in the pronunciation regulation) form is primed only by the canonical form in the long-term repetition-priming task, whereas if exemplar representations of all three variants are stored in memory, each variant is primed by the canonical form.

2. Experiment 1: AX discrimination

2.1 Method

2.1.1 Participants

Thirty native speakers of Korean using the Seoul standard dialect (8 males, 22 females), aged between 18 and 27 years (mean = 21.23 years, SD = 2.27), participated in the experiment. None of the participants reported hearing, reading, or language impairments. They were paid for their participation in the study.

2.1.2 Stimuli

Six Korean disyllabic words (verb or adjectives) with /lp/ in their first syllable (/c'alp-ta/ 'short', /nʌlp-ta/ 'wide', /jalp-ta/ 'thin', /jʌlp-ta/ 'thin' /palp-ta/ 'to step on', and /t'ʌlp-ta/ 'bitter') were selected as the experimental stimuli. The consonant cluster /lp/ is realized in three ways; either the first consonant (e.g., [l]), the second consonant (e.g., [p]), or both (e.g., [lp]).

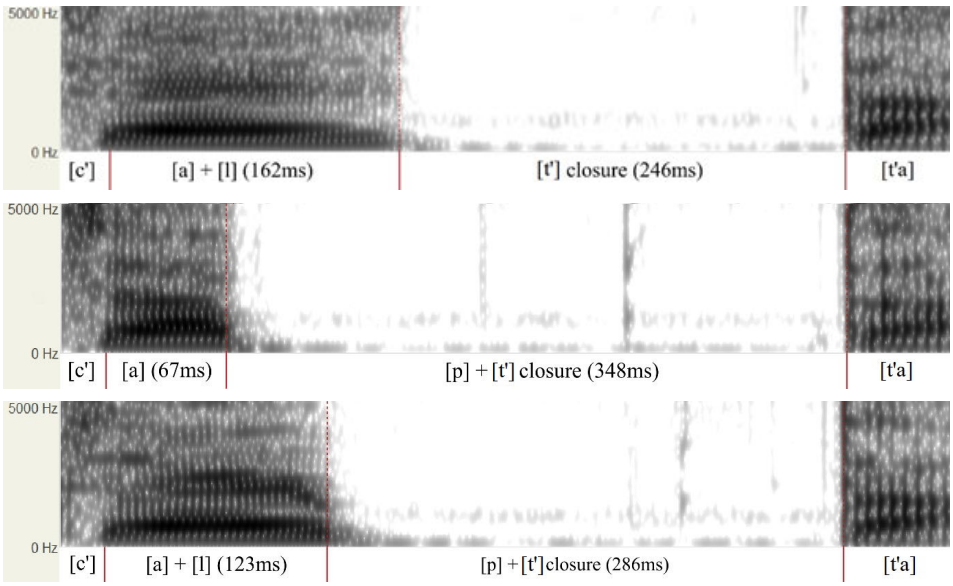

Figure 1. Spectrograms of the three variants of the word /c'alp-ta/ 'short' realized with the first consonant [l] (top), the second consonant [p] (middle), and both consonants [lp] (bottom)

The acoustic characteristics of the three variants of /lp/ for the word /c'alp-ta/ are shown in Figure 1. A comparison of the two variants [l] and [lp] in the two figures (top and bottom) reveals that the closure duration of the stop after the vowel [a] is longer for the [lp] variant (bottom) than for [l] (top), presumably because of the production of [p] in the former. In the spectrograms of [p] and [lp], as shown in the two bottom figures, the F2 of the preceding vowel [a] is shown to drop more rapidly for [p] than for [lp] because the immediately following consonant is labial, not alveolar lateral (Kim and Kang 2021).

Each test word was provided in three different conditions: 1) identical condition (e.g., [jal.ta] – [jal.ta], [jap.ta] – [jap.ta], for /jalp.ta/ 'thin'), 2) target condition (e.g., [jal.ta] – [jalp.ta], [jalp.ta] – [jal.ta], [jap.ta] – [jalp.ta], [jalp.ta] – [jap.ta]) and 3) control condition (e.g., [jal.ta] – [jap.ta], [jap.ta] – [jal.ta]). Under the identical condition, the same variants with a single consonant were provided as the first (A) and second (X) tokens, whereas in the control condition, the two separate phonemes as variants (the first and second consonants of the cluster) were presented as A and X. In the target condition, one test word (A) had a variant of the first or second consonant

of the cluster, whereas the variant of the other word (X) had both consonants.

The stimuli were recorded by two female speakers of Seoul Standard Korean (age: 20 and 25 years) in a sound-attenuated booth using a Tascam HD-P2 solid-state recorder and a Shure KSM44 microphone. Each speaker was directed to produce all variant forms of the /lp/ cluster in the six test words. The test words were provided in the frame sentence (/i.tan.ʌ.nin. __ i.ta/ 'This word is ___'). We carefully examined the productions of the two speakers and found that their overall realization patterns of the target cluster were matched to those in Figure 1.

2.1.3 Procedure

The participants were informed of the details about the experimental procedure, and all provided informed consent before the experimental session. Each participant was tested individually in a sound-attenuated booth using E-prime 2.0 software (Psychology Software Tools) and headphones (Sennheiser, HD 599). The word pairs were provided with two ISI conditions, 100 ms and 1000 ms, to evaluate the participants' listening at purely auditory and linguistic levels, respectively. The two tokens (A and X) in all pairs were spoken by two different speakers and were always physically different even under the identical condition. Trials with a 100-ms ISI were presented first, followed by a five-minute break and then those with a 1000-ms ISI. Each trial began with a fixation mark (+) that remained at the center of the screen for 1000 ms. Next, the participants were presented with an auditory word token, followed by another auditory token over headphones after either 100 ms or 1000 ms. The participants were asked to decide whether the two tokens in each trial were the same or different. They were instructed to press "1" on the keypad as accurately and quickly as possible if they thought they were the same or "3" if they thought they were different. If the participants did not respond within 3 s, a new pair was presented. A total of 96 trials (6 test words x 8 possible conditions (2 identical + 4 target + 2 control) x 2 ISI) were presented in 2 blocks; in one block, the ISI was 100 ms (designed to assess phonetic-level processing), whereas in the other, the ISI was 1000 ms (designed to assess phonological-level processing). Randomization was performed for each block. Before the experiment began, 16 trials were presented for practice which were excluded from the analysis. During the practice session, the volume was adjusted to a comfortable level for each participant. The experiment lasted approximately 15

min. After the AX discrimination task, participants produced and recorded each word in a separate sentence (e.g., /os.i.jalp.ta/ 'The clothes are thin' for /jalp.ta/ 'thin'), to evaluate their own productions of the clusters. Only the target words were extracted from the recorded sentences for acoustic measurements. The two authors, trained phoneticians, coded each token as C1 (the first consonant of the cluster realized), C2 (the second consonant, realized), or CC (both consonants of the cluster realized) based on waveforms and spectrograms as in Figure 1 (See Kim and Kang (2021) for more detailed acoustic characteristics of each realization of the cluster).

The complete list of sentences for the recording is presented in Appendix I. At the end of the experiment, each participant filled out a sheet on their familiarity with the test words, using a 7 Likert scale (1 = least familiar, 7 = most familiar). The mean familiarity score for the test words was 6.17 (range = 1–7, SD = 1.62).

2.2 Results

Trials with reaction times (RT) greater or less than 2.5 SDs from an individual's mean RT were excluded (2.61% of the data). Figure 2 shows the mean discrimination accuracy scores for each word condition in the two ISI contexts.

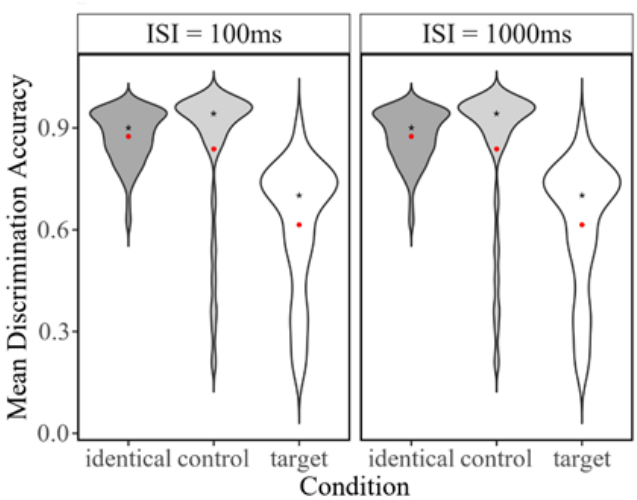

Figure 2. Mean discrimination accuracy scores for identical, control, and target conditions (with 100ms-ISI on the left and 1000ms-ISI on the right) (The asterisk and the small circle in each violin indicate the median and mean, respectively)

The accuracy scores (correct or incorrect) were submitted to a mixed-effects logistic regression model utilizing the *glmer* function from the package *lmerTest* (Kuznetsova et al. 2017) in R (version 4.3.3; The R Foundation for Statistical Computing 2024). The dependent variable was the responses (binary-coded as 0 for incorrect and 1 for correct), and the fixed factors were ISI (100ms, 1000ms) and word condition (identical, control, target). The random factors considered were by-speaker and by-word intercepts, and by-subject random slopes for the word condition. The fixed factors were treatment-coded, with the ISI (100 ms) and the identical condition as the reference levels, to interpret the effects of other conditions relative to these baselines. To build the model, several models were constructed in a stepwise manner from the maximal model containing all factors, such as ISI and word condition, and nested models were compared using the likelihood ratio test of significance. To facilitate model convergence, the BOBYQA optimizer was used with a maximum of 100,000 iterations. A likelihood ratio test revealed a significant main effect of condition ($\chi^2(2) = 47.71$, $p < .001$), but no significant main effect of ISI ($\chi^2(1) = 1.26$, $p = .263$). Additionally, the interaction between ISI and condition was not significant ($\chi^2(2) = 4.2$, $p = .121$). The optimal model was glmer(ACC ~ condition + (1 + condition | Subject) + (1 | word), data = ax, family = binomial, control = glmerControl (optimizer = "bobyqa", optCtrl = list(maxfun = 1e5))). The statistical results of the AX discrimination task are presented in Table 1.

Table 1. Statistical results for perception accuracy in AX discrimination

Fixed effects		β	SE	z ratio	p value
(Intercept)		2.150	0.233	9.231	$p < .001$***
ISI: 1000ms		0.109	0.096	1.129	$p = 0.259$
Condition: control		0.187	0.357	0.522	$p = 0.602$
Condition: target		-1.669	0.261	-6.389	$p < .001$***
Random effects					
Groups	Name	Variance	SD	Corr	
subject	(Intercept)	0.811	0.900		
	Condition: control	2.305	1.518	-0.13	
	Condition: target	1.306	1.143	-0.61	0.84
word	(Intercept)	0.024	0.155		

Pairwise comparison Contrast	β	SE	z ratio	p value
identical - control	-0.187	0.357	-0.522	$p = 0.861$
identical - target	1.669	0.261	6.389	$p < .001$***
control - target	1.855	0.238	7.807	$p < .001$***

Notes. ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$

We observed a main effect of word condition; a follow-up pairwise comparison was conducted using the Tukey test from the emmeans package (Lenth et al. 2021) to further examine the effect of word condition. The mean accuracy scores for the target condition were significantly lower than those for the identical and control conditions. However, the identical and control conditions did not show any significant differences. These results demonstrate that Korean listeners had little difficulty perceiving the variants realized as a canonical form (i.e., [l]) or a single innovative form (i.e., [p]), whereas there was an apparent cost in processing the innovative form realized with both consonants (i.e., [lp]).[2]

Concerning the results related to the ISI, the two ISI conditions (100 ms and 1000 ms) did not influence the perception of the variants of /lp/. The findings for the 100-ms condition were fully replicated for the 1000-ms condition without any significant difference. We expected that these two ISI conditions would elicit auditory and phonemic level of perception, based on the results of Werker and Logan (1985). Our results, however, did not support this finding and rather corroborate previous findings that ISI duration has little effect on discrimination sensitivity in either segments (Boomershine et al. 2008) or suprasegmentals (Wayland and Guion 2004; Soo and Monahan 2023).

In the following, we further assessed whether the critical discrimination scores for the target pairs obtained from Korean participants ([lp] versus [l]/[p]) may be associated with their production ratios for realizing the /lp/ cluster as [lp] (CC production). A close inspection of the cluster production revealed substantial variability across participants. The percentage of consistently producing one of the three variants of the cluster was only 26.67% ([l] for all tokens), whereas in the

2 As suggested by an anonymous reviewer, we further conducted a correlation analysis between word familiarity and accuracy of the AX discrimination task and found that there was no significant relationship between them [coefficient = -0.021, t(2878) = -1.163, $p = .245$].

percentage of producing the cluster with multiple variants, two variants reached approximately 50% (20% for [l] or [p], 30% for [l] or [lp], and 0% for [p, lp]), and three variants reached 23.33% ([l, p, lp]). Looking at the mean production percentage of each variant across the participants, the realizations of /lp/ as [l], [p], and [lp] were 69.4% (range = 16.7 − 100, SD = 30.0), 12.2% (range = 0 − 50, SD = 16.3), and 18.3% (range = 0 − 83.3, SD = 23.7), respectively. Given that the word /palp-ta/ has a different canonical form than the others, we measured the mean percentages of each variant for this word separately; the realizations of /lp/ as [l], [p], and [lp] were 60% (range = 0–100, SD = 49.8), 23.3% (range = 0–100, SD = 43.02), and 16.7% (range = 0–100, SD = 37.9), respectively. Although the overall pattern of production for the word /palp-ta/ across the three variants appeared similar to that of other words, it is notable that the percentage of the [p] variant was much higher for /palp-ta/ than for the remaining words.

Table 2. Acoustic characteristics of the variants of /lp/

variant	duration (ms) of preceding vowel (+ [l])	closure duration (ms) of target stop (+ following stop)	F2 transition (Hz)
[l]	131 (27.9)	206 (33.5)	1619 (182)
[p]	55 (21.9)	245 (35.1)	1141 (139)
[lp]	99 (22.8)	221 (43.9)	1442 (172)

The acoustic measurements in Table 2 demonstrate that 1) the mean closure duration of [l] is shorter than that of [lp] because of the production of [p] for [lp]; and 2) the F2 transition values between the preceding vowel and the target stop are higher for [l] (alveolar) than the remaining stops with a labial and more importantly, the F2 is shown to drop more rapidly for [p] than for [lp] because the immediately following consonant is labial, not alveolar lateral. These results are in good agreement with those of the three spectrograms in Figure 1.

Based on the finding that the mean CC production ratio was 0.183 (range = 0–0.833; SD = 0.24) for /lp/, the relationship between the CC production ratios and discrimination scores for the target pairs was further investigated using Pearson correlation analysis. As shown in Figure 3, the ratio of producing both consonants of the cluster did not correlate with discrimination accuracy [coefficients = 0.068, t = 0.359, df = 28, p = 0.72].

Figure 3 illustrates the non-significant relationship between the participants' own production of /lp/, realized as [lp], and the discrimination accuracy between [lp] and [l]/[p]. Upon inspecting the four quarters of the scatterplot, the top and bottom of the left side appear crowded, suggesting that participants produced a few of the tokens with variants of both consonants but showed a great deal of variation in the target discrimination accuracy.

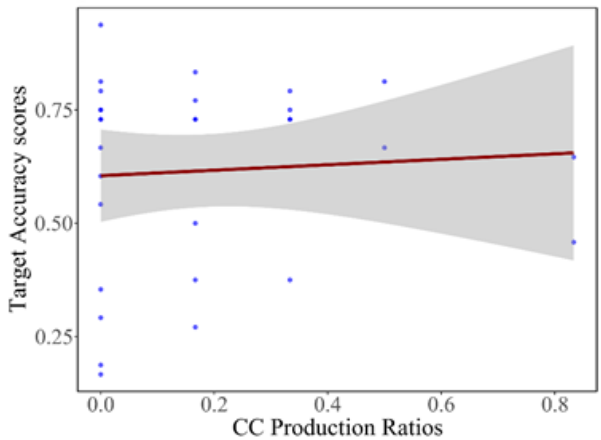

Figure 3. Association between CC production ratios and discrimination accuracy in the target condition.

3. Experiment 2: Long-term repetition priming

Given that the results of Experiment 1 showed poor discrimination between the innovative variant realized with both consonants and that with a single consonant of the /lp/ cluster, Experiment 2 was designed to evaluate how the innovative variant with both consonants and the other two variants are stored and accessed in the long term.

3.1 Method

3.1.1 Participants

Sixty participants who did not participate in Experiment 1 took part in the experiment.

They were all native speakers of Korean using the standard Seoul dialect (26 males and 34 females), aged between 18 and 31 years (mean = 21.07 years, SD = 2.76), and none reported hearing, reading, or language impairments. They were paid for their participation in the study.

3.1.2 Stimuli and procedure

Five Korean disyllabic words (verbs or adjectives) that were selected for the AX discrimination task were utilized for stimulus items (/c'alp-ta/ 'short', /nʌlp-ta/ 'wide', /jalp-ta/ 'thin', /palp-ta/ 'to step on', /t'ʌlp-ta/ 'bitter'). Each participant was presented with a long block where the canonical form of the target cluster was first presented as a prime and then one of the three related forms, namely, the canonical form or the other two innovative variants of the /lp/ cluster were presented approximately 36 to 43 items later as a target. To avoid having the same participant encounter the same test word (e.g., /palp-ta/), three counterbalanced sets were created, each of which had one of the three related forms; the canonical form ([p] for the word /palp-ta/ and [l] for the remaining words) and the other two variants ([l] for the word /palp-ta/ and [p] for the remaining words or [lp] for both words). Participants were randomly assigned to one of the three sets (20 participants per set). To prevent the participants from focusing on the structure of the target cluster, 38 fillers were additionally constructed that showed various phonological alternations in Korean such as consonant cluster simplification of the /lm/ or /lh/ cluster (e.g., [tam.t'a] /talm-ta/ 'resemble'; [il.t^{h}a] /ilh-ta/ 'lose'), merger (e.g., [pwat.t'a] /po-as'-ta/ past form of 'see'), or tensification (e.g., [pok.t'a] /pok-ta/ 'parch'). To balance the number of real and nonword targets in the lexical decision task, the same number of nonwords was constructed. For each word, the coda consonant was replaced by another single coda (e.g., [tan.t'a] or [taŋ.t'a] for [tam.t'a] /talm.t'a/ 'to resemble'). The target word and its nonword shared onset consonants and vowels. However, 10 target words (5 primes + 5 targets) had no corresponding nonwords, and 10 filler nonwords were employed instead. Thus, each block consisted of 96 trials (5 prime words + 5 target words + 38 filler words + 48 filler nonwords). The three types of prime-target pairs for the same word (e.g., prime: [jal.t'a] – target: [jal.t'a], [jap.t'a], [jalp.t'a]) were presented in separate sets. The stimuli were recorded by a female speaker using the Seoul Korean dialect (age = 20 years).

Participants were asked to judge whether each item they heard was a real Korean word or not as accurately and quickly as possible. They were instructed to press "1" on the keypad if they thought it was a real Korean word or "3" if they thought it was not a Korean word. First, a fixation mark was presented for 500 ms, followed by an auditory word token. If the participant did not respond within 3,000 ms, a new token was presented. The reaction time (RT) latencies for correct responses were measured, and the RT differences between the first presented item (prime: canonical form) and its variant form (target: one of the three variant forms), which were presented at long intervals within a block, were analyzed. Before the experiment began, a set of 10 trials was provided as practice, which included only fillers. At the end of the experiment, each participant filled out a sheet for familiarity with the test words (real words only), using a 7 Likert scale (1 = least familiar, 7 = most familiar). The mean familiarity score for the test words was 6.91 (range = 2–7; SD = 0.48).

3.2 Results

The error rate was 7.27% (n = 419 of 5760 (96 trials x 60 participants)). Trials with RTs greater or less than 2.5 SDs of an individual's mean RT were further excluded (2.17%). To evaluate the long-term effects, only the target cluster pairs were included (5 prime-target pairs x 60 participants), removing the fillers. If a member of a pair was excluded based on this criterion, the corresponding member was also removed (22.7%). Thus, the total number of target tokens for the analysis was 232. Priming effects were calculated by subtracting the RT for the second item in a pair (target) from the RT for the first item in a pair (prime). Recall that the regulations of the Standard Pronunciation of Korean for the /lp/ cluster differ between the /palp-ta/ and the remaining words, such that /lp/ is simplified to [p] in the former and [l] in the latter. Thus, the three variants were named differently between the verb /palp-ta/ and the remaining four verbs (/c'alpta/ 'short', /nʌlpta/ 'wide', /jalpta/ 'thin', /palpta/ 'to step on', /t'ʌlpta/ 'bitter'), except for the variant with both consonants (CC) (i.e., [lp]). For the word /palp-ta/, the canonical variant form (CA) was [p], whereas the other single variant (C1) was [l]. In contrast, for the remaining four words, the canonical variant form (CA) was [l], whereas the other single variant (C1) was [p]. The mean priming effects for each condition are shown in Figure 4.

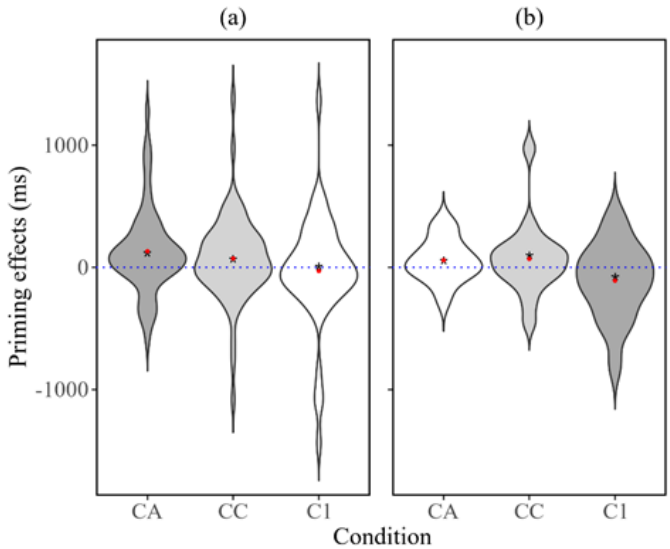

Figure 4. Priming effects (ms) for the four target words with /lp/ (a) and the word /palp-ta/ (b) for three different variant types, canonical variant (CA), variant with both consonants (CC), and the other single variant (C1) (The asterisk and the small circle in each violin indicate the median and mean, respectively)

The mean priming effects for the canonical variant (CA) were 58.5 ms for /palp-ta/ and 132.4 ms for the remaining four verbs, whereas those for the CC variant were 68.9 ms and 71.6 ms and those for the C1 variant were -109 ms and -30.2 ms for /palp-ta/ and the others, respectively. When the canonical form or the form with both consonants was realized in the target, there were considerable priming effects, whereas when the other single-variant form was realized, there was an inhibitory effect. Based on the raw values of the priming effects, the two types of words – /palp-ta/ and the remaining words – revealed differences in the exact amount and pattern of priming. In the word /palp-ta/, the amount of priming for the variant preserving both consonants was slightly larger than but not significantly different from that of the identical (canonical) form, suggesting that these two variants were likely to be equivalent in their ability to promote the activation of a lexical representation in long-term memory. However, in the remaining four words, the repeated canonical form had approximately twice the priming effect of the variant with two consonants. It is worthwhile to note that the priming effect for the variant with both consonants

was comparable between the word /palp-ta/ and the remaining four words (i.e., 68.9 ms versus 71.6 ms), while that for the repeated canonical form was different between these two types of words (i.e., 58.5 ms for /palp-ta/ versus 132.6 ms for the remaining words).

The priming effects (ms) were submitted to a mixed-effects linear regression model utilizing the *lmer* function from the package l*merTest* (Kuznetsova et al. 2017) in R (version 4.3.3; The R Foundation for Statistical Computing 2024). The dependent variable was priming effects and the fixed factor was condition (CA, CC, C1). The random factors considered were by-speaker and by-word to intercept. The independent variable was treatment-coded, with the canonical condition as the reference level, in order to interpret the effects of other conditions relative to the baseline. The optimal model was model_3 = lmer(priming ~ condition + (1 | subject) + (1 | word), data = longterm, REML = FALSE). The statistical results of the long-term repetition priming task are presented in Table 3.

Table 3. Statistical results for long-term repetition priming effects

Fixed effects	β	SE	df	t value	p value
(Intercept)	115.44	44.15	27.51	2.615	$p < 0.05^{*}$
Condition: CC	-43.66	50.91	194.58	-0.858	$p = 0.39218$
Condition: C1	-173.33	54.66	192.50	-3.171	$p < 0.01^{**}$
Random effects					
Groups	Name	Variance	SD		
Subject	(Intercept)	18472	135.91		
Word	(Intercept)	1577	39.71		
Residual		103334	321.46		
Pairwise comparison					
Contrast	β	SE	df	t ratio	p value
CA – CC	43.7	50.9	195	0.858	$p = 0.6676$
CA – C1	173.3	54.7	192	3.171	$p < 0.01^{**}$
CC – C1	129.7	54.0	193	2.403	$p = 0.045^{*}$

Notes. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

There was a significant main effect of the condition ($F(2, 193.41) = 5.31$, $p < .01$). Planned comparisons showed that the repeated canonical form and the innovative form with both consonants yielded priming without significant differences when presented as targets, whereas the variant form with a single innovative form different from the prime word showed a significant difference in priming from the other two

variants. In other words, if the prime (canonical form) was followed by the same form as the target, there was significant priming owing to the complete repetition of the canonical form. This result was a joint contribution from the general identity effect and the canonical form advantage. However, if the prime form (canonical form) was followed by a variant with a single non-canonical consonant as the target, no priming was observed, presumably because the two segments ([l] and [p]) are separate phonemes in Korean. The most interesting was the case where the prime was a canonical form and the target was a form with both consonants, which exhibited a significant priming effect.[3]

5. Discussion and conclusion

The present study examined the perceptual (Experiment 1) and lexical (Experiment 2) encoding of context-independent variants of the consonant cluster /lp/ in Korean. Experiment 1 examined native Korean speakers' perception of all three variants of /lp/ in a speeded AX discrimination task. We found that 1) the degree of difficulty in discriminating between the three variants of /lp/ was not equal; when the word was realized with the variant of both consonants ([lp]), it was least accurately distinguished from words with other types of variants; 2) the ISI did not impact the perception of these variants; and 3) Korean speakers' production pattern of /lp/ was not directly associated with that of its perception. The detailed interpretations and explanations of each of the three results are presented in the following.

First, Korean listeners had no difficulty in perceiving the variants realized as a canonical form (i.e., [p] for the verb stem 'palp-', and [l] for the remaining stems), or a single innovative form (i.e., [l] for the verb stem 'palp-', and [p] for the remaining stems), whereas there was an apparent cost in processing of the innovative form realized with both consonants of the cluster (i.e., [lp]). The former result appears plausible because the canonical and single innovative forms are separate phonemes (/l/ vs. /p/) in Korean; thus, the discrimination of these two variants led to a categorical distinction. In contrast, the latter result indicates that the two consonants of the cluster were not perceptually salient as in the consonant clusters of other languages (e.g.,

3 As in the AX discrimination task, the correlation between word familiarity and priming effects was not shown to be significant [coefficient = -0.02, t(230) = -0.30, p = .767].

milk versus *mill* in English).

Regarding the results on ISI, the two ISI conditions, 100 ms and 1000 ms, did not influence the perception of the variants of /lp/. The speeded AX discrimination task was assumed to be a listening task to tap at an earlier level of processing to examine the possible effects of the presence or absence of surface contrast. However, it has also been claimed that a shorter ISI in a task may reduce language-specific effects, while a longer ISI may elicit phonological effects (Fox 1984; Werker and Logan 1985; Polka 1991). Hence, it was expected that any immediate effect of a canonical or innovative form on the auditory processing level might have faded away over a longer and thus more linguistic course. However, in this study it apparently did not: regardless of the ISI, perception of the innovative form realized with both consonants was likely to undermine the discrimination of spoken words including /lp/. These results are not novel to the present study. Boomershine et al. (2008), for example, examined the impact of allophony versus contrast on speech perception using two tasks; an off-line similarity rating task with a 1000-ms ISI and an on-line AX discrimination task with a shorter 100-ms ISI. These two ISI conditions were designed to tap the phonemic versus phonetic level of processing. However, because similar results were observed between the two tasks, they argued that language-specific influences emerged in both types of task; offline with a long ISI and online with a shorter ISI. This is further supported by the present findings that short versus longer ISIs do not simply induce auditory or phonetic and more abstract phonological processing levels.

A comparison of the results of the AX discrimination task and acoustic analyses of the participants' production of words containing /lp/ revealed a mismatch between speech perception and production. Although the exact frequency and pattern of realizing the /lp/ cluster varied across participants, only 27% of the participants consistently realized the cluster using a single variant, and the remaining participants showed free variation across the three variants. Interestingly, approximately 19% of the participants used the innovative variant preserving both consonants of the cluster ([lp]). This led to the question of whether these participants would be better at perceiving word pairs containing the variant [lp]. The results of the Pearson correlation analysis revealed that the participants who produced the innovative variant preserving both consonants did not necessarily exhibit great performance in discrimination. Previous research examining the link between speech perception and production

showed mixed results, from tight coupling to strict separation of the two modalities (see Baese-Berk 2010). To the best of our knowledge, no study has investigated the perception–production link in the use of phonological variants; however, based on previous findings and the present results, it appears that the frequency of producing a particular variant among many is not directly correlated with the success of its perception.

The poor discrimination of the variant of /lp/ as [lp] raises the possibility that when native Korean speakers produce this variant, they may not produce it with two timing units as in other languages (e.g., *milk* cf. *mill*). Recall that the innovative variant preserving both consonants violates the Korean phonotactic constraints, prohibiting consonant clusters in the syllable coda position. As production of the /lp/ cluster as [lp] is not perceptually (or presumably, acoustically) salient, this variant does not appear to violate the Korean phonotactic constraints in terms of linking both consonants to a single timing unit. Namely, when the /lp/ cluster surfaces, whether as both consonants, or a single consonant, the coda may allow only one timing unit and the three types of variants appear to be distinguished in the melody tier. The exact syllable structure of [lp], however, needs to be systematically verified in further studies.

The results of the AX discrimination task led to the question of whether discrimination differences across variant pairs emerged when a task requires access to a lexicon. To address this question, a long-term repetition priming task was administered in Experiment 2. When the variant containing a canonical single consonant was followed by three variant forms, Korean speakers exhibited similar priming effects between the repeated canonical form and an innovative variant with both consonants ([lp]) as targets, whereas no priming effect emerged for the target of the non-canonical single variant form. As facilitation over longer distances only occurs when there is a complete overlap between the prime and target (Pallier et al. 2001) or when both variants are lexically stored (Sumner and Samuel 2009), both the canonical form which was repeated, and the innovative variant preserving the two consonants appeared to be stored in the mental lexicon. One may point out that as the innovative form with both consonants and the canonical form did not exhibit a significant difference in the amount of priming, the two categories may be stored in the lexicon without distinction. However, this view does not seem plausible as the canonical forms of the /lp/ cluster in the verb /palp-ta/ and the remaining

words are completely different segments, as in [l] for the former and [p] for the latter. Note that when both types of verbs were pronounced with the preservation of two consonants of the cluster, the priming effects were similar between the two types of words. However, when the canonical variant of the cluster was repeated, /palp-ta/ showed less priming than the other words, albeit statistically non-significant. This finding may be associated with younger speakers' /lp/ production. In other words, in the word /palp.ta/, younger speakers prefer to preserve the first consonant ([l]) despite pronunciation regulation; thus, when the canonical form with [p] was completely repeated in the target, its priming effect would not be as strong as that in the repeated canonical form (i.e., [l]) of other words with /lp/. Therefore, the present findings suggest that the two variants, the canonical form and the form with both consonants, were likely to be stored as separate lexical items in long-term memory, whereas the identity priming effect was mitigated for words whose canonical forms differed from the actual pronunciation.

A possible explanation for the poorly discriminated but lexically encoded variant ([lp]) can be found in the literature regarding the effects of orthography on long-term memory. In previous studies, less frequent but canonical forms were observed to show greater priming effects than the corresponding frequent, non-canonical variants. It has been argued that this canonical form advantage results from the impact of orthography, specifically the consistent relationship between unreduced pronunciation and its orthographic form (Ranbom and Connine 2007; Escudero et al. 2008; Racine et al. 2014; Viebahn et al. 2018; Charoy and Samuel 2020). For instance, Escudero et al. (2008) examined the phonetic and lexical mapping of auditorily confusable L2 non-words. They taught them to L2 learners and evaluated their word recognition using an eye-tracking paradigm. When native Dutch speakers learned the English contrast /ɛ-æ/, one group of participants learned by matching the auditory forms to corresponding pictures, while the other group additionally saw the spellings of words. Unlike the former group, speakers with access to spelled forms showed asymmetric word recognition, suggesting that explicit orthographic information may be helpful in building separate lexical representations of similar-sounding words. The present findings are in line with these earlier findings in that the innovative variant with both consonants was directly reflected in its spelled form. The two consonants of the cluster (i.e. [lp]) are directly matched to the corresponding orthographic forms <괘> in which the left orthographic form <ㄹ> is matched to the first consonant

[l] and the right form <ㅂ> to the second consonant [p]. When the participants had access to spelled forms, they constructed distinct representations. While Korean speakers did not proficiently discriminate the innovative variant in the AX discrimination task, which required only phonetic or phonological analysis of the signal, the memory representation for the variant was distinct and linked to separate lexical entries, presumably owing to its orthographic representation.

The results of the long-term repetition priming task do not seem to be accounted for by the two competing models of lexical representation as mentioned earlier. The finding that not only the canonical form but also the new variant preserving both consonants of the cluster exhibited an identity priming effect does not support the abstract view of lexical representation, in which there is a single abstract, canonical representation for perceptual variants. If so, we should have observed identity priming for the canonical variant only, whether it was /l/ or /p/. The substantial and equivalent priming of the innovative variant that preserved both consonants indicates that this innovative variant is stored in long-term memory as a separate category, although it is perceptually confusable from the canonical variant. However, the present findings are inconsistent with the episodic model of lexical representations. Although lexical representations seem to be more fine-grained than those proposed in the abstract model of lexical representation, the present results only partly support or refute the exemplar model of lexical representation (Goldinger 1998; Ranbom and Connine 2007). The canonical form of the /lp/ cluster was shown to be word-specific: in most verbs or adjectives with /lp/, the canonical form is [l], but it is [p] for the verb /palp-ta/ 'step on.' The finding that the lexical representation of a cluster varies from word to word cannot be accounted for by arguing for a single abstract lexical representation. The word-specific phonetic pattern supports a model in which the lexical representations of words include incrementally updated information about the phonetic distribution of each word (Pierrehumbert 2002). However, the present results also demonstrated that among the three variants of the /lp/ cluster, priming effects were observed in two variants, the canonical form and innovative form with both consonants, whereas the non-canonical single consonant variant showed an inhibitory effect. These results show a clear divergence from previous research on the perception and lexical representation of regular variants. For example, Sumner and Samuel (2005) examined the immediate and long-term processing consequences for the three variants of the final /t/ in English (e.g., flute /flut/) which are a fully articulated [t], a

coarticulated, glottalized [ʔtˀ], and a singly articulated glottal stop [ʔ]. Semantic and long-term priming experiments showed that variation did not hinder short-term semantic processing, whereas strong priming was only found for the canonical, unchanged form of /t/ during form processing over time. Based on these results, the authors proposed that surface information is used in immediate processing, whereas exemplar representations of variants are not stored in long-term memory. As such, neither of the competing models of lexical representation mentioned above seems to fit the present findings well.

In conclusion, the results of the AX discrimination and long-term repetition priming tasks in the present study allowed us to draw a distinction between the perceptual and lexical encoding of innovative variants of the /lp/ cluster in Korean. Among the three variants, the variant preserving both consonants ([lp]) was not perceptually salient but lexically encoded as a separate category, probably owing to its orthographic form. Given that no previous study has examined the immediate and long-term effects of context-independent variation emerging from an on-going sound change, the present findings may improve our understanding of how language users accommodate phonological variation by expanding its scope.

References

Baese-Berk, Melissa Michaud. 2010. *An examination of the relationship between speech perception and production.* Phd Dissertation. Northwestern University.

Boomershine, Amanda, Kathleen Currie Hall, Elizabeth Hume, and Keith Johnson. 2008. The impact of allophony versus contrast on speech perception. In Peter Avery, B. Elan Dresher, and Keren Rice (eds.), *Contrast in phonology: Theory, perception, acquisition*, 146-172, The Hague: Netherlands: Mouton de Gruyter.

Bürki, Audrey, F. Xavier Alario, and Ulrich H. Frauenfelder. 2011. Lexical representation of phonological variants: Evidence from pseudohomophone effects in different regiolects. *Journal of Memory and Language* 64: 424-442. https://doi.org/10.1016/j.jml.2011.01.002.

Bürki, Audrey, Mirjam Ernestus, and Ulrich H. Frauenfelder. 2010. Is there only one "fenêtre" in the production lexicon? On-line evidence on the nature of phonological representations of pronunciation variants of French schwa words. *Journal of Memory and Language* 62: 421-437. https://doi.org/10.1016/j.jml.2010.01.002.

Charoy, Jeanne and Arthur G. Samuel. 2020. The effect of orthography on the recognition

of pronunciation variants. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 46(6): 1121-1145. https://doi.org/10.1037/xlm0000781.

Cho, Taehong and Sahyang Kim. 2009. Statistical patterns in consonant cluster simplification in Seoul Korean: Within-dialect interspeaker and intraspeaker variation. *Phonetics and Speech Sciences* 1(1): 33-40.

Cho, Young-mee Yu. 1990. *Parameters of consonantal assimilation.* Phd Dissertation. Stanford University.

Coenen, Else, Pienie Zwitserlood, and Jens Bölte. 2001. Variation and assimilation in German: Consequences for lexical access and representation. *Language and Cognitive Processes* 16(5-6): 535-564. https://doi.org/10.1080/01690960143000155.

Connine, Cynthia M., Debra Titone, and Jian Wang. 1993. Auditory word recognition: Extrinsic and intrinsic effects of word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19(1): 81-94. https://doi.org/10.1037/0278-7393.19.1.81.

Dell, Gary S. 1986. A spreading-activation theory of retrieval in sentence production. *Psychological Review* 93(3): 283-321.

Escudero, Paola, Rachel Hayes-Harb, and Holger Mitterer. 2008. Novel second-language words and asymmetric lexical access. *Journal of Phonetics* 36(2): 345-360. https://doi.org/10.1016/j.wocn.2007.11.002.

Fox, Rober A. 1984. Effect of lexical status on phonetic categorization. *Journal of Experimental Psychology Human Perception and Performance* 10(2): 526-540. https://doi.org/10.1037/0096-1523.10.4.526.

Frauenfelder, Uli H, Mark Scholten, and Alain Content. 2001. Bottom-up inhibition in lexical selection: Phonological mismatch effects in spoken word recognition. *Language and Cognitive Processes* 16(5-6): 583-607. https://doi.org/10.1080/01690960143000146.

Gaskell, M. Gareth and William D. Marslen-Wilson. 1996. Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance* 22: 144-158.

Gaskell, M. Gareth and William D. Marslen-Wilson. 1998. Mechanisms of phonological inference in speech perception. *Journal of Experimental Psychology: Human Perception and Performance* 24: 380-396. https://doi.org/10.1037/0096-1523.24.2.380.

Goldinger, Stephen D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105(2): 251-279.

Gow Jr., David W. 2001. Assimilation and anticipation in continuous spoken word recognition. *Journal of Memory and Language* 45(1): 133-159. https://doi.org/10.1006/jmla.2000.2764.

Gow Jr., David W. 2002. Does English coronal place assimilation create lexical ambiguity? *Journal of Experimental Psychology: Human Perception and Performance* 28(1): 163-179.

Gow Jr., David W. 2003. Feature parsing: Feature cue mapping in spoken word recognition. *Perception and Psychophysics* 65: 575-590. https://doi.org/10.3758/BF03194584.

Iverson, Gregory and Shinsook Lee. 1994. Variation as optimality in Korean cluster reduction.

ESCOL '94: 174-185.

Jun, Jongho. 1998. Restrictions on consonant clusters. *Eoneohag* 23: 189-204.

Kim, Jung-Yun and Eunji Kang. 2021. Phonetic variation of Korean stem-final consonant clusters beginning with a liquid. *Studies in Phonetics, Phonology and Morphology* 27(2): 161-192. DOI: 10.17959/sppm.2021.27.2.161.

Kim-Renaud, Young-Key. 1974. *Korean consonantal phonology.* Phd Dissertation. The University of Hawaii.

Kuznetsova, Alexandra, Per B. Brockhoff, and Rune H. B. Christensen. 2017. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82(13): 1-26. DOI: 10.18637/jss.v082.i13.

Kwon, Soohyun, Tae-Jin Yoon, Sujin Oh, and Jeong-Im Han. 2024. Analogical generalization in progress in stem-final consonant cluster simplification: Evidence from Seoul and Gyeongsang Korean. Ms.

Lahiri, Aditi and William Marslen-Wilson. 1991. The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition* 38: 245-294. https://doi.org/10.1016/0010-0277(91)90008-R.

Lenth, Russell V., Paul Buerkner, Maxime Herve, Jonathon Love, Hannes Riebl, and Henrik Singmann. 2021. *Emmeans: Estimated marginal means, aka least-squares means: R package version 1.6.3.*

Levelt, Willem J. M., Ardi Roelofs, and Antje S. Meyer. 1999. A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22: 1-75.

Mitterer, Holger and Leo Blomert. 2003. Coping with phonological assimilation in speech perception: Evidence for early compensation. *Perception and Psychophysics* 65: 956-969.

National Institute of Korean Language. 2018. hangeul majchumbeob [The standard Korean pronunciation regulations]. Munhwa cheyug gwangwangbu gosi, [Ministry of Culture, Sports and Tourism Announcement].

Pallier, Christophe, Angels Colomé, and Núria Sebastián-Gallés. 2001. The Influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science* 12(6): 445-449. https://doi.org/10.1111/1467-9280.00383.

Pierrehumbert, Janet B. 2002. Word-specific phonetics. In Carlos Gussenhoven and Natasha Warner (eds.), *Laboratory Phonology* 7: 101-139. Berlin: Mouton de Gruyter.

Polka, Linda. 1991. Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions. *Journal of the Acoustical Society of America* 89: 2961-2977. https://doi.org/10.1121/1.400734.

R Core Team. 2024. R: A language and environment for statistical computing (version 4.3.3) [Computer software]. R Foundation for Statistical Computing, Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/.

Racine, Isabelle, Audrey Bürki, and Elsa Spinelli. 2014. The implication of spelling and frequency in the recognition of phonological variants: Evidence from pre-readers and readers.

Language, Cognition and Neuroscience 29(7): 893-898. https://doi.org/10.1080/01690965.2013.832784.

Ranbom, Larissa J. and Cynthia M. Connine. 2007. Lexical representation of phonological variation in spoken word recognition. *Journal of Memory and Language* 57: 273-298. https://doi.org/10.1016/j.jml.2007.04.001.

Shin, Jiyoung, Jieun Kiaer, and Jaeeun Cha. 2012. *The sounds of Korean*. Cambridge: Cambridge University Press.

Sohn, Ho-Min. 1999. *The Korean language*. Cambridge: Cambridge University Press.

Soo, Rachel and Philip J. Monahan. 2023. Phonetic and lexical encoding of tone in Cantonese heritage speakers. *Language and Speech* 66(3): 652-677. https://doi.org/10.1177/00238309221122090.

Steriade, Donca. 1997. *Phonetics in phonology: The case of laryngeal neutralization*. Ms. The University of California, Los Angeles.

Sumner, Meghan and Arthur G. Samuel. 2005. Perception and representation of regular variation: The case of final-/t/. *Journal of Memory and Language* 52: 322-338. https://doi.org/10.1016/j.jml.2004.11.004.

Sumner, Meghan and Arthur G. Samuel. 2009. The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language* 60(4): 487-501. https://doi.org/10.1016/j.jml.2009.01.001.

Viebahn, Malte C., James M. McQueen, Mirjam Ernestus, Ulrich H. Frauenfelder, and Audrey Bürki. 2018. How much does orthography influence the processing of reduced word forms? Evidence from novel-word learning about French schwa deletion. *Quarterly Journal of Experimental Psychology* 71(11): 2378-2394. https://doi.org/10.1177/1747021817741859.

Wayland, Ratree P. and Susan G. Guion. 2004. Training English and Chinese listeners to perceive Thai tones: A preliminary report. *Language Learning* 54(4): 681-712. https://doi.org/10.1111/j.1467-9922.2004.00283.x.

Werker, Janet F. and John S. Logan. 1985. Cross-language evidence for three factors in speech perception. *Perception and Psychophysics* 37(1): 35-44.

Whitman, John B. 1985. Korean clusters. In Susumu Kuno, John Whitman, Ik Hwan Lee, and Young Se Kang (eds.), *Harvard studies in Korean linguistics*, 280–290. Cambridge, MA: Harvard University.

Appendix I. List of the sentences for Experiment 1

	Target	Frame sentence
1	/c'alp-ta/ 'short' <짧다>	/cul.i/ 'line – nom.' + /c'alp.ta/ 'The lines are short'
2	/nʌlp-ta/ 'wide' <넓다>	/ma.taŋ.i/ 'yards – nom.'+ /nʌlp.ta/

		'The yards are wide'
3	/jalp.ta/ 'thin' <얇다>	/os.i/ 'clothes – nom.'+ /jalp.ta/ 'The clothes are thin'
4	/jʌlp.ta/ 'thin' <엷다>	/sek.i/ 'color – nom.' + /jʌlp.ta/ 'The colors are thin'
5	/t'ʌlp.ta/ 'bitter' <떫다>	/kam.i/ 'persimmon – nom.' + /t'ʌlp.ta/ 'The persimmons are bitter'
6	/palp.ta/ 'step on' <밟다>	/pal.il/ 'foot – obj.' + /palp.ta/ 'They step on a foot'

(nom. = nominative marker; obj. = objective marker)

Tae-Hwan Choi
Assistant professor
Verum Liberal College
Catholic Kwandong University
24 Beomil-ro, 579 beon-gil
Gangneung 25601, Korea
E-mail: tchoi@cku.ac.kr

Jeong-Im Han
Professor
Dept. of English, Konkuk University
120 Neungdong-ro, Gwangjin-gu
Seoul 05029, Korea
E-mail: jhan@konkuk.ac.kr

Received: 2025. 05. 13.
Revised: 2025. 08. 27.
Accepted: 2025. 09. 12.