말소리와 음성과학 제2권 제4호 (2010)
pp. 83~92

A Preliminary Report on Perceptual Resolutions of Korean Consonant Cluster Simplification and Their Possible Change over Time

Cho, Taehong[1)]

ABSTRACT

The present study examined how listeners of Seoul Korean would recover deleted phonemes in consonant cluster simplification. In a phoneme monitoring experiment, listeners had to monitor for C2 (/k/ or /p/) in C1C2C3 when C2 was deleted (C1 was preserved) or preserved (C1 was deleted). The target consonant (C2) was either /k/ or /p/ (e.g., *il**k**-tələto vs. pal**p**-tələto*), and there were two listener groups, one group tested in 2002 and the other in 2009. Some points have emerged from the results. First, listeners were able to detect deleted phonemes as accurately and rapidly as preserved phonemes, showing that the physical presence of the acoustic information did not improve the listeners' performance. This suggests that listeners must have relied on language-specific phonological knowledge about the consonant cluster simplification, rather than relying on the low-level acoustic-phonetic information. Second, listener groups (participants in 2002 vs. 2009), differed in processing /p/ versus /k/: listeners in 2009 failed to detect /p/ more frequently than those in 2002, suggesting that the way the consonant cluster sequence is produced and perceived has changed over time. This result was interpreted as coming from statistical patterns of speech production in contemporary Seoul Korean as reported in a recent study by Cho & Kim (2009): /p/ is deleted far more often than /p/ is preserved, which is likely reflected in the way listeners process simplified variants. Finally, listeners processed /k/ more efficiently than /p/, especially when the target was physically present (in C-preserved condition), indicating that listeners benefited more from the presence of /k/ than of /p/. This was interpreted as supporting the view that velars are perceptually more robust than labials, which constrains shaping phonological patterns of the language. These results were then discussed in terms of their implications for theories of spoken word recognition.

Keywords: Korean, consonant cluster simplification, speech perception, phoneme monitoring, phonological knowledge, language change

1. Introduction

Spoken-word recognition entails a process by which listeners map the surface phonetic patterns in the speech signal onto underlying phonological representations of word forms in the mental lexicon. This mapping is relatively straightforward if the surface patterns correspond faithfully to the underlying representations. There are many cases, however, in which the mapping of the phonetic input onto the lexicon appears to require highly sophisticated linguistic and cognitive competence. One well-known case is when underlying representation surfaces with alternative pronunciations due to the phonological process of place assimilation (see, e.g., Gaskell & Marlsen-Wilson, 1996; Gow, 2001, 2002, 2003; Gow & Im, 2004). For example, when *right berry* versus *ripe berry* are produced in casual speech, they often present a challenge to the listener because /t/ in *right* is assimilated to /b/ in *berry* in terms of place of articulation, resulting in [raɪp bɛrɪ] for both *right berry* and *ripe berry*. Despite the seemingly neutralized pronunciation for both phrases, however, listeners are generally able to distinguish them, presumably relying on fine-grained phonetic cues that differentiate the assimilated (derived) [p] in *righ[p] berry* from the underlying

1) Hanyang University, tcho@hanyang.ac.kr
This work was supported by the research fund of Hanyang University (HY-2007-G).

Received: November 9, 2010
Revision received: December 5, 2010
Accepted: December 15, 2010

www.kci.go.kr

/p/ in *ripe berry* (see Gow 2003 for further discussion on this point under the rubric of the Feature Parsing Theory).

The present study examines an even more radical case, the case of tri-consonantal simplification in Korean, in which an underlying segment is not phonologically modified but completely deleted. In Korean, when a tri-consonantal sequence C1C2C3 (/lkt/ or /lpt/) occurs in a word-form due to morpheme concatenation such as *ilk-ta* ('to read') and *jalp-ta* ('to be thin'), either C1 or C2 can be deleted. An important question is whether, and how, listeners would recover such deleted phonemes in simplified forms of tri-consonantal cluster in Korean.

The present study tests this question in a phoneme monitoring experiment with 104 Seoul Korean listeners. The target was either /k/ or /p/, deleted or preserved in C2 position. The target-bearing words were presented in semantically matched phrasal contexts as in a phrase like *not^{h}i-lil il**k-t**ə*lato* 'although reading the notebook' or *kutu-lo pa**lp-t**ə*lato* 'although stepping (on it) with the shoes'. In the C2-deleted case (e.g., *il**k-t**ə*lato* pronounced as *il-tə*lato)*, target detection required 'restoration' of a consonant that was physically absent. Thus, in contrast to the C2-preserved case, where target detection could be based on acoustic-phonetic evidence alone, target detection in the C2-deleted condition required retrieval of an underlying lexical representation, or some form of inference based on phonological knowledge about deletion. Phoneme monitoring responses might therefore be expected to be faster and more accurate in the C2-preserved condition than in the C2-deleted condition, which will be referred to as *the phonetic superiority hypothesis.* Alternatively, however, given that native listeners have phonological knowledge of C2 deletion as tri-consonantal cluster simplification process, and given that listeners frequently encounter C2-deleted variants and recover the underlying phonemes, retrieval of the underlying segment may well be efficient enough to compensate for the lack of surface phonetic cues to the target, which will be referred to as *the phonological compensation hypothesis.* If that were the case, C2-deleted variants would be processed efficiently, perhaps as efficiently as would be the case with C2-preserved variants.

There is, however, an additional factor that adds a complication to understanding the mapping of the simplified variants onto the lexicon in Korean－i.e., the asymmetry between /k/-bearing versus /p/-bearing consonant clusters (/lkt/ vs. /lpt/) in the way they are simplified. It has been noted that in Seoul Korean the first consonant /l/ is deleted for /lkt/ while for /lpt/, the second consonant /p/ is deleted as shown in (1) (e.g., Kim-Renaud, 1976; Cho, 1990; Iverson & Lee, 1994; Jun, 1998; Martin, 2006).

(1) a. /lkt/ → [{l}kt*]
(e.g., *ilk-ta* 'to read-DEC', *malk-ta* 'to be clear-DEC')
b. /lpt/ → [l{p}t*]
(e.g., *jalp-ta* 'to be thin-DEC', *tʃ*alp-ta* 'to be short-DEC')
(NB: The third member of the cluster (C3) is pronounced as a tense consonant as a result of post-obstruent tensification rule.)

Recently, however, in a production study on statistical patterns of consonant cluster simplification in Seoul Korean, Cho & Kim (2009) reported that although there was an asymmetry between /lkt/ and /lpt/ sequences, it was not categorical as in (1). Instead, as given in Figure 1 (adopted from Figure 2 in Cho & Kim, 2009), when speakers simplified the cluster, they deleted either the first consonant /l/ or the second consonant, regardless of whether the second consonant was /k/ versus /p/. The results nevertheless showed that the asymmetry between /lkt/ and /lpt/ did exist, but it did not lie in whether the first or the second consonant was deleted, but in how frequently /k/ versus /p/ were deleted. As can be seen in Figure 1b, speakers preserved /k/ (30.9%) far more frequently than /p/ (less than 5%), while they deleted /p/ more frequently (57%) than /k/ (44.6%). From a different angle, the data also showed that in production of /l**k**t/ speakers had no clear preference for deleting versus preserving /k/, but in production of /l**p**t/, strong preference was found for /p/-deleted to /p/-preserved variant.

Given that speech production shows different statistical patterns for /lkt/ versus /lpt/, one can also predict that such an asymmetry in speech production is likely to be reflected in speech perception. In phoneme monitoring experiments, it was therefore examined how the statistically-driven production asymmetry between /k/-bearing versus /p/-bearing clusters would be reflected in speech perception. For example, the fact that /p/ is deleted more frequently than /k/ indicates that listeners are exposed more to /p/-deleted variants than /k/-deleted variants, which in turn leads to a prediction that /p/ from a /p/-deleted variant would be recovered more efficiently than /k/ from a /k/-deleted variant.

Another important implication that can be drawn from the results of Cho & Kim (2009) is that the consonantal simplification patterns have indeed changed over time, especially when compared with the traditional observations on the simplification patterns－i.e., for /lkt/, C1 (/l/) is deleted but for /lpt/, C2 (/p/) is deleted. It appears that the change over time explains why the consonantal cluster simplification patterns in the

www.kci.go.kr

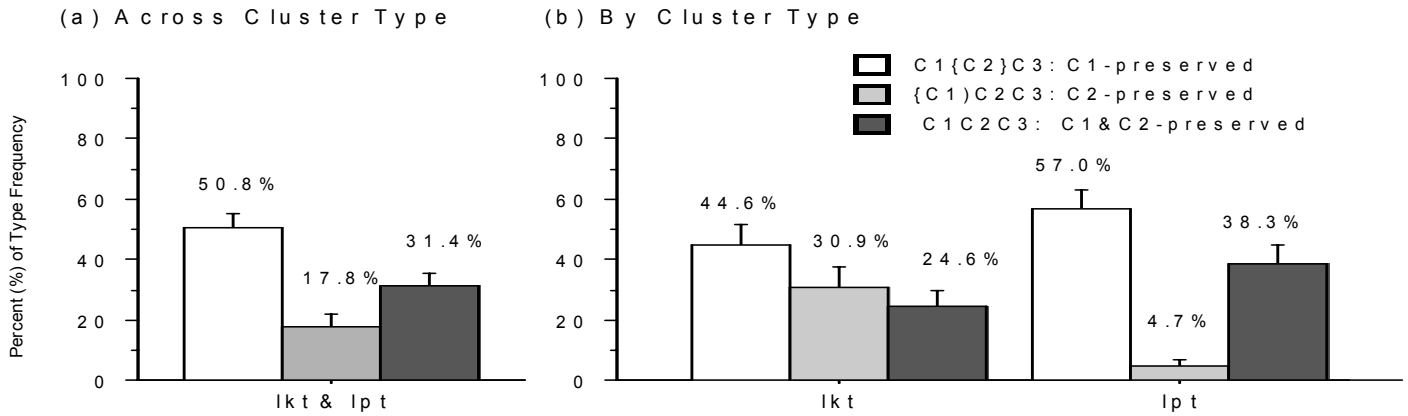

Figure 1. Overall patterns of consonant cluster simplification as a function of C2 Type (/k/ vs. /p/) based on production patterns of twenty speakers of Seoul Korean. (adopted from Cho & Kim, 2009).

contemporary Seoul Korean are so much different from the traditionally assumed simplification patterns suggested by previous phonologists, which dated back as early as 1970s when Kim-Renaud pointed out the cluster simplification asymmetry between /lkt/ and /lpt/ in Seoul Korean as stated in (1) above.

Cho & Kim (2009) indeed discussed this issue by comparing their production data with earlier production data in Cho (1999), and suggested that the way speakers simplified consonant clusters in Seoul Korean appeared to have changed over the past 10 years. However, since the two studies employed different methodologies and experimental protocols, one cannot directly compare the two studies to draw conclusion on exactly how speakers' production has changed over time. This is in fact a more general limitation imposed on studies about sound changes over time. It is not therefore easy to assess how language changes over time both in production and perception, especially when the changes are examined with relatively short time intervals. I have however been lucky to collect phoneme monitoring perception data from more than 100 speakers with 8 years apart, once in early 2002 and once in early 2009, using the exact same speech stimuli and experimental protocols. Comparing the perception data between the two listener groups in different points in time would illuminate how the way listeners process consonant cluster simplification in Seoul Korean has changed over the past 8 years.

In sum, the present study therefore explore two important, interrelated, questions. First, it examines how listeners of Seoul Korean process tri-consonant cluster simplification when the second consonant is physically present (in the C2-present condition) or absent (in the C2-deleted condition), testing the two competing hypotheses: *the phonetic superiority hypothesis* (i.e., listeners would benefit greatly from the presence of the phonetic information for the target in the C2-deleted condition) versus *the phonological compensation hypothesis* (i.e., listeners would use their implicit phonological knowledge to recover phonologically deleted phonemes in an efficient way). Second, it investigates how different listener groups in 2002 and 2009 behave differently in recovering the deleted phoneme out of the C2-deleted variants. In particular, it examines whether the listeners' phoneme detection is asymmetrically influenced by the consonant type (/k/ versus /p/) in connection with statistical patterns of /k/ versus /p/ deletion in contemporary Seoul Korean. Given that /p/ is more frequently deleted than /k/ in contemporary Seoul Korean, it is expected that listeners would process /p/-deleted variants better than /k/-deleted variants. Finally, the results of the present study will allow us to consider whether acoustic-phonetic differences between /k/ and /p/ would influence listeners' performance. Given that the acoustic properties of velars are perceptually more robust than those of labials (Jun, 1995, 2004), it is expected that all else being equal, listeners would benefit more from the presence of /k/ than of /p/.

2. Method

2.1 Subjects

In total one hundred and four Korean students participated in the experiment. The first half of them were students at Korea University, Seoul, and the other half were students at Hanyang University, Seoul, who participated in 2002 and 2009, respectively.

2.2 Materials

The target was either a velar (/k/) or a labial (/p/) stop, which was always the second consonant member of the tri-consonant sequence /lkt/ or /lpt/ (e.g., for /k/, *not^{h}i-lil il**k**-tələto* ‘although

www.kci.go.kr

reading the notebook'; for /p/, *kutu-lo pa**lp-t**ələto* 'although stepping (on it) with the shoes'; note that '-' refers to a morpheme boundary, and that the words are specified with IPA transcriptions). As given in Table 1, six words were used as /p/-bearing test words and eight as /k/-bearing target words. The target bearing words occurred consistently as the second word in a two-word utterance. (Note that the third member of the cluster becomes tensified as a result of post-obsruent tensification rule, which will be discussed more in detail in the discussion section.)

Table 1. List of cluster-bearing test words in a clause

(a) /lpt/-bearing two-word phrases	
(밟다) *kutu-lo pa**lp-t**ələto*	'although stepping (on it) with the shoes'
(떫다) *tʃatu-ka t*ə**lp-t**ələto*	'although the plum is bitter'
(넓다) *sija-ka nə**lp-t**ələto*	'although the view is wide'
(얇다) *tuk*e-ka ja**lp-t**ələto*	'although the thickness is thin'
(짧다) *tʃʰima-ka tʃ*a**lp-t**ələto*	'although the skirt is short'
(엷다) *pitʃʰk*al-i jə**lp-t**ələto*	'although the color is light (pale)'
(b) /lkt/-bearing two-word phrases	
(읽다) *notʰi-lil il**k-t**ələto*	'although reading the notebook'
(맑다) *koŋki-ka ma**lk-t**ələto*	'although the air is clear'
(굵다) *kituŋ-i ku**lk-t**ələto*	'although the pillar is thick'
(밝다) *tʃomjəŋ-i pa**lk-t**ələto*	'although the light is bright'
(긁다) *kaljəwə ki**lk-t**ələto*	'although scratching, because it's itchy'
(늙다) *əntʃen-ka ni**lk-t**ələto*	'although (we) get old some day'
(묽다) *noŋto-ka mu**lk-t**ələto*	'although the concentration is diluted'
(붉다) *noil-i pulk-tələto*	'although the sky is red'

Notes: The basic forms of cluster-bearing words are provided in parenthesis in Korean; '*' refers to the fortis (tensed) consonant.

There were another 14 foils similar to the target bearing two-word utterances. (Note that in this experiment, foils were the auditory stimuli that did not contain the target phoneme specified on the computer screen, such that subjects were not supposed to respond, pressing a 'yes' button. Foils are generally used in phoneme monitoring experiments like the present study in order to prevent subjects from pressing a 'yes' button habitually without paying attention to the auditory stimuli.) Those foils were same as the target words (6 words with /p/ and 8 words with /k/), but they were used in different carrier sentences. For example, the target bearing word '*il**k-t**ələto*' was used as an experimental item in an auditory carrier sentence '*notʰi-lil il**k-t**ələto'* ('although reading the notebook') with a visually specified target /k/ and it was used as a foil in a different auditory carrier sentence '*sinmun-lil il**k-t**ələto'* ('although reading the newspaper') with a visually specified target /p/ which did not exist in the auditory stimulus. Each foil was presented in the running order after its corresponding experimental trial (with the target specified), so that the occurrence of similar target-bearing utterance did not influence listeners' performance. There were also 21 fillers and 21 filler foils (again two-word phrases), with other types of tri-consonantal sequences (e.g., *k* ilh-ta*, (끊다) 'to cut something out') that were not used in the experimental test items. (Note that fillers and filler foils were included as distractors in order to make the experiment complex enough, so that subjects were not biased.)

These materials for testing the consonant cluster simplification were collected as part of a larger study combined with materials for another experiment, which also involved the target phonemes /p/ and /k/ and several other targets. In total, there were 74 trials with /p/ and /k/ as targets (37 target-present and 47 foil trials) and 92 with other targets (46 target-present and 46 foil trials). All targets occurred in similar contexts in terms of prosodic boundaries and target locations. In particular, the prosodic boundary was controlled, so that the word boundary between the first and the second words was embedded inside a phrase. Finally, 28 practice items were constructed. Among those, 15 items were related to consonant cluster simplification, which included both C2-deleted and C2-preserved items with consonant clusters not used in the test words.

2.3 Procedures

Speech materials were recorded in a sound-damped booth by a male native Korean speaker (the author). Two different scripts were read, one containing target-bearing words written with C2 deleted, and one containing target-bearing words written with C1 deleted (and C2 preserved). The C2-deleted and C2-preserved (C1-deleted) forms were deliberately misspelled (i.e., *'pa**lp-t**ələto'* '밟더라도', mis-spelled as C2-deleted *'pa**l-t**ələto'* '발더라도' and as C1-deleted, *'pa**p-t**ələto'* '밥더라도'). The use of such scripts thus ensured that the test words were pronounced without any residual cues for C2 in C2-deleted forms. The two-word phrase was then read several times by the speaker with no phrasal prosodic boundary between the two words. A most naturally produced token was chosen with the intended word (non-phrasal) boundary between the two words, which was later further verified by another Korean prosodic transcriber.

In each year (2002 and 2009), subjects were divided into two groups of 26, so that listeners in each group heard all fillers and one version of each of the test words (C2-deleted or C2-preserved). That is, C2 deletion condition (C2-deleted vs.

www.kci.go.kr

C2-preserved) was counterbalanced across two groups.

The subjects' task was to monitor for a target phoneme from the auditory stimulus. The target was visually pre-specified on the computer screen, printed in *Hangul* for one second; ㅂ(p) and ㄱ (k). Subjects were told that the target occurred in the coda position. They were instructed to press a button as fast and as accurately as possible when they detected a target. Response latencies (reaction time) were recorded relative to the offset of the consonant sequence－i.e., the end of the stop closure for C3, together with errors. Note that since targets were not physically present in the C2-deleted condition, C3 was used as a consistent alignment point.

Each subject heard 166 trials (74 related to the cluster simplification and 92 others related to another experiment, which were combined in the experiment). The inter-stimulus-interval was about 4 seconds, and the entire experiment ran about 20 minutes.

3. Results

3.1 Errors (percent-missing)

One of the basic results on the errors is that listeners were able to detect the target consonant with about 80% detection accuracy (i.e., with only about 20% missing) even when the target consonant was absent (deleted) in the speech input. Repeated measures ANOVAs showed no main effect of C2-Deletion on the error rates, nor was there any interaction between C2-Deletion and other factors such as Consonant Identity (/k/ vs. /p/) and Year (2002 vs. 2009). This revealed that listeners were able to detect the deleted (acoustically absent) targets as accurately as the preserved targets (Figure 2a).

The Consonant Identity and Year factors, however, revealed signifiant main effects (F[1,98]=-79.19, F[1,98]=21.61, respectively, at $p<0.001$). Listeners detected /k/ with less errors (more accuracy) than for /p/ (Figure 2b), and listeners in 2002 performed better with higher accuracy than listeners in 2009 (Figure 2c). As shown in Figure 3a, there was no significant interaction between C2-Deletion and Consonant Identity (F[1,98]<1), but a significant interaction between Consonant Identity and Year was observed (F[1,18=-24.1, $p<0.001$), which stemmed from the fact that listeners in both 2002 and 2009 showed similar error rates in detecting /k/, but in detecting /p/, the listeners in 2002 were far more accurate than those in 2009, as shown in Figure 3b.

In sum, listeners were able to detect deleted phonemes as accurately as preserved phonemes, showing that the physical presence of the acoustic information did not improve the listeners'

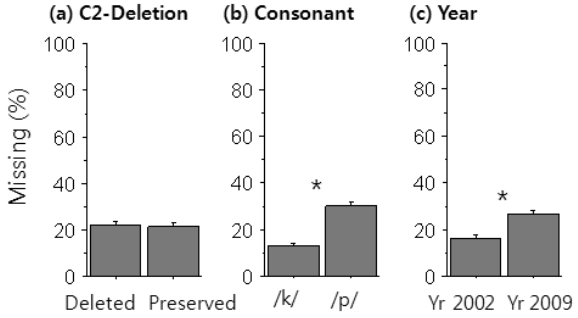

Figure 2. Effects of Consonant, C2-deletion and Year on the error rates (%-missing).

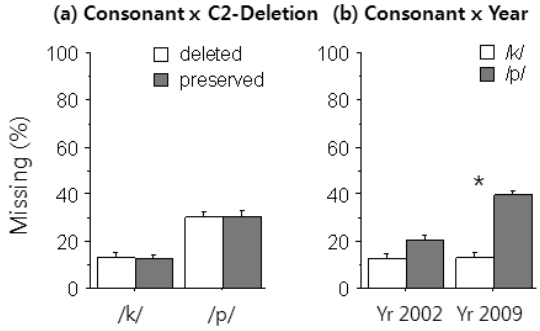

Figure 3. Interaction effects on the errors (%-missing) (a) between Consonant Identity and C2-Deletion and (b) between Consonant, C2-Deletion and Year.

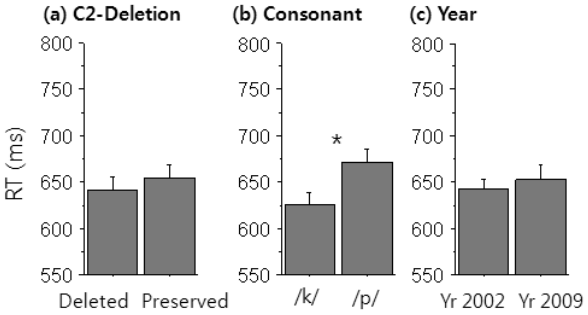

Figure 4. Effects of Consonant, C2-deletion and Year on detection latency (RTs).

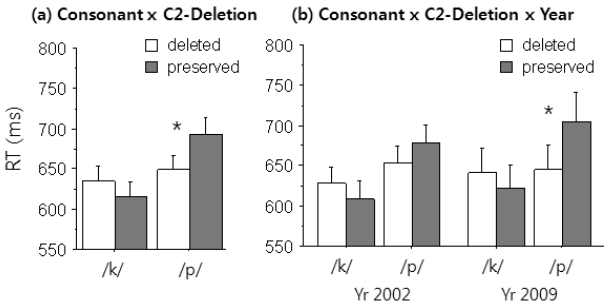

Figure 5. Interaction effects on detection latency (a) between Consonant Identity and C2-Deletion and (b) between Consonant, C2-Deletion and Year.

www.kci.go.kr

performance in terms of detection accuracy. This pattern was consistent for both /k/ and /p/ and for the listeners in 2009 and in 2002. The listener groups, however, differ in processing /p/ versus /k/: listeners in 2009 failed to detect /p/ far more frequently than those in 2002.

3.2 Response latency (RT)

One of the basic results on response latency obtained from repeated measures ANOVAs was that the phoneme detection was not necessarily faster when the target was physically present in the speech signal (in C2-preserved condition) than when it was absent (in C2-deleted condition), as can be seen in Figure 4a. As was the case with the error rates, there was no significant main effect of C2-Deletion on detection latency (RTs).

The Consonant Identity factor (/k/ versus /p/), on the other hand, yielded a significant main effect ($F[1,98]=12.5$, $p<0.001$), showing a consonant-induced asymmetry: listeners were significantly faster in detecting /k/ than /p/, as shown in Figure 4b. There was, however, a significant interaction between Consonant Identity and C2-Deletion ($F[1,98]=6.62$, $p<0.05$). As can be inferred from Figure 5a, the interaction stemmed from two patterns. First, although there was no main effect of C2-Deletion, /p/ revealed a C2-Deletion effect, showing that listeners were faster in detecting *deleted* /p/ than preserved /p/ ($p<0.05$), while no such difference was observed for /k/. Second, the faster reaction time for /k/ than for /p/ (the Consonant Identity effect) was significantly so only in C2-preserved condition ($p<0.005$), but not in C2-deleted condition.

As for the Year effect (2002 vs. 2009), ANOVAs showed no significant main effect ($F[1,98]<1$, Figure 4c). Most importantly, however, there was a three-way interaction between all three factors ($F[1,98]=5.98$, $p<0.05$). As illustrated in Figure 5b, planned pair-wise comparisons showed that the Consonant x C2-deletion interaction was reliable only for the listeners who participated in the experiment in 2009. That is, the significant Consonant Identity x C2-Deletion interaction shown in Figure 5a turned out to be reliably only for the listeners in 2009.

In sum, listeners were not faster in detecting the physically present (preserved) consonant than the physically absent (deleted) consonant, even though the preserved consonant carried more acoustic information for the target consonant. There was also an effect of Consonant Identity, showing that listeners processed /k/ more rapidly than /p/ when the target was physically present (in C-preserved condition). Finally, there was an important three-way interaction among factors, indicating a clear difference between the listener groups in 2002 and in 2009: For /p/, listeners in 2009 were faster in detecting a deleted /p/ than a preserved /p/, but listeners in 2002 did not show such an effect.

4. Discussion

One of the fundamental questions of the present study was how listeners would recover the underlying representations of deleted phonemes in the tri-consonant cluster simplification in Seoul Korean. In connection with this question, two competing hypotheses were considered: *the phonetic superiority hypothesis* versus *the phonological compensation hypothesis*. The phonetic superiority hypothesis was based on the general assumption that the acoustic-phonetic information of the target present in the speech signal (as in C2-preserved condition) would help listeners to recognize that target. On the other hand, the phonological compensation hypothesis was grounded on the assumption that the lack of bottom-up phonetic support would not necessarily hinder the phonological process because listeners would use their phonological knowledge in processing phonological modified segments in their mother tongue.

Results on detection errors and latency in the phoneme monitoring experiment showed that listeners were indeed able to detect deleted phonemes as well as preserved phonemes, suggesting that listeners did not benefit from the presence of the acoustic-phonetic cues for the target phoneme in C2-deletion condition. This runs counter to the phonetic superiority hypothesis, but supports the phonological compensation hypothesis —i.e., listeners could compensate for the lack of bottom-up acoustic-phonetic support by employing their phonological knowledge.

This has implications for theories of spoken word recognition. It has been suggested that phonological processes such as assimilation are not indeed complete, so that, for example, the derived segment [p] due to assimilation (e.g., [p] in *righ[p] berry*) is not phonetically the same as the underlying /p/ (/p/ in *ripe berry*) (Gow, 2002, 2003; Gow & Im, 2004). Gow (2003) argues that phonological modification such as place assimilation does not hinder the process of spoken word, but instead the acoustic-phonetic residuals in the speech signal help listeners in recognizing the word as intended by the speaker. Based on this argument, Gow further proposes that phonological modifications are processed by listeners through universal perceptual mechanisms that utilize the acoustic-phonetic details of phonological modification at the lower-order (auditory-perceptual)

www.kci.go.kr

level, rather than mechanisms driven by language-specific knowledge that operates at a higher-order phonological level. (See Mitterer, et al., 2006, for a similar argument that the recognition of phonologically assimilated words does not depend on language-specific knowledge.) The results of the present study, however, demonstrated that not every phonological modification is processed by a universal perceptual mechanism. The retrieval of the underlying phonemes in processing C2-deleted variants cannot be accounted for by such language-universal auditory-perceptual mechanisms—i.e., Korean listeners did not take advantage of the bottom-up acoustic-phonetic information in recovering deleted phonemes, indicating that the auditory-perceptual mechanism played little role.

The results are instead in line with the basic tenet of the theory of phonological inference proposed by Gaskell & Marslen-Wilson (1996, 1998) that listeners employ their phonological knowledge when processing phonologically modified spoken words (e.g., by undoing a phonological rule: [p] before a labial consonant must have been derived from /t/). Given that there was no acoustic information about the target in C2-deleted condition, Korean listeners must have invoked language-specific phonological knowledge when recovering the deleted C2 which resulted from a language-specific phonological process, the tri-consonantal cluster simplification.

Another important goal of the present study was to explore whether and how the way listeners in 2002 and 2009 would differ in processing simplified forms of consonant clusters. Results showed that the way the simplified variants of the consonant cluster were processed indeed differed between listener groups of 2002 and 2009, especially in terms of how deleted /p/ versus preserved /p/ was processed: Listeners in 2002 were able to recover the physically absent (deleted) /p/ as efficiently (rapidly and accurately) as they detected the physically present (preserved) /p/, but listeners in 2009 showed far better performance in recovering the physically absent /p/ than in detecting the physical present /p/. The results therefore appear to support the hypothesis that was discussed at the outset of the paper: Given the possibility that the way speakers simplify tri-consonant clusters has changed over time, as can be inferred from statistical patterns in production of consonantal cluster simplification reported in Cho & Kim (2009), the way listeners process them may as well have changed.

A question then arises. What would be the driving force that have brought about the difference in processing /p/ between listener groups in 2002 and 2009? More specifically, what could have caused the listeners in 2009 to process deleted /p/ more efficiently than preserved /p/ while the listeners in 2002 did not show such an asymmetry? A possible answer to this question may be found in statistical patterns in speech production in contemporary Seoul Korean, which are likely to be reflected in speech perception. Recall that Cho & Kim (2009) reported that /p/-deleted variants occur far more frequently than /p/-preserved variants (57% versus 4.7%) in contemporary Seoul Korean. Given that the use of phonological knowledge is modulated by the frequency of occurrence of different phonological variants (Connine 2004; Ranbom & Connine 2007), it is possible that different phonological variants of the /p/-bearing words are stored in the mental lexicon, and those separately stored representations are frequency-coded.

The frequency-coded storage of different phonological variants can be further elaborated by an exemplar-based approach (e.g., Goldinger, 1996, 1998; Pierrehumbert, 2001, 2002, 2003). The tenet of this approach is as follows. Speech perception involves the storages of exemplars of specific speech events in a multidimensional phonetic space, and phonetic categories may be developed forming "clouds" in different regions of the space. Each member of a cloud is a remembered instance of a given category (or label), such that each category is associated with a frequency-weighted distribution of phonetic events. In this framework, the more frequently occurring /p/-deleted variants of tri-consonant bearing words are stored more densely in the phonetic space of the lexicon, constituting the basic substrate of the mental lexicon, and therefore retrieval of a /p/-deleted variant is made easier from the dense pool, relative to retrieval of a less frequently occurring /p/-preserved variant.

Why did this happen to listeners in 2009 and why not to listeners in 2002? It is plausible that at the time that listeners participated in the experiments in 2002, the frequency of occurrence of /p/-deleted variants was not high enough to override the acoustic phonetic effects associated with /p/-preserved variants. The frequency of occurrence of /p/-deleted variants may have consistently increased over time, so that in contemporary Seoul Korean, the frequency effect may even exceed the benefit of the presence of the acoustic phonetic cues, as evident in the performance by listeners in 2009.

Finally, another noteworthy of the findings was the fact that /k/ was processed more efficiently (with faster response latencies) than /p/, especially when the target consonant was present in the speech signal (in C2-preserved condition). In other words, listeners appear to have benefited more from the acoustic-phonetic

www.kci.go.kr

cues of /k/ than of /p/. This perceptual advantage of the velars is in line with the general assumption of the so-called "phonetically-driven" phonology: sounds are intrinsically different in terms of their acoustic-phonetic properties with some sounds being perceptually more robust than others, and such differences may constrain certain aspects of the sound patterns of the world languages (Flemming 1995; Hayes et al. 2004; Hume and Johnson 2001; Jun 1995, 2004; Steriade 1999, 2001). For example, Jun (1995, 2004) discusses the listener-oriented production hypothesis (Jun 1995, 2004) which stipulates that speakers make more effort to preserve sounds with more robust acoustic cues for the listener's benefit. This may account for the cross-linguistic tendency that velars undergo phonological alterations less often than labials or alveolars, based on the assumption that velars are perceptually more salient than labials or alveolars, such that speakers preserve /k/ for the listener's benefit, and listeners in turn take advantage of that. The results of the present study (that listeners were faster in detecting /k/ than /p/ when they were both physically present in the speech signal) appear to support the listener-oriented production hypothesis. (See Cho & McQueen, 2008, for further discussion on this point.

Limitations. Thus far I have discussed important findings of the present study along with their implications for theories of speech perception. Some points made based on the results, however, are not entirely immune to controversy. First, although the deleted variants used as stimuli were produced by reading a deliberately misspelled script with the deleted phoneme not being spelled out, one cannot guarantee that the deleted phonemes did not contain the acoustic phonetic residuals of the deleted phoneme, given that the author served as the speaker, and therefore was not entirely naive. Second, there is a possibility that the asymmetry between /k/ and /p/, especially the fact that /k/ was processed more efficiently than /p/, stemmed from the contextual effects. For example, a /k/-bearing phrase like *not^{h}i-lil il**k-t**əlato* 'although reading the notebook' may have a stronger contextual support (i.e., *not^{h}i-lil*, 'notebook') than a /p/-bearing phrase like *kutu-lo pa**lp-t**əlato* 'although stepping (on it) with the shoes', which may induce relatively faster responses in detecting /k/. Some arguments made in the present study therefore must be taken with caution, especially the one that listeners were able to recover phonemes without recourse to any residual phonetic cues and the one that listeners benefited from the presence of /k/ more than that of /p/ because /k/ was perceptually more salient. In order to clarify these issues, additional experiments are underway where the phonetic residual cues have been ensured to be absent and the contextual information was eliminated through low-pass acoustic filtering.

Notes for tensification of the third member of the cluster. A reviewer argued that the results were not viably interpretable as the present study did not take into account the fact that the third consonant becomes tense due to post-obstruent tensification rule. It is true that upon hearing tensified C3, a listener would have made a phonological inference that there must have been an obstruent C2 in the underlying representation as a trigger for tensification of C3. Such a phonological inference, however, does not nullify any of the results and their implications discussed in the present study. First, a fundamental question of the present study was how listeners would recover deleted C2s as compared to preserved C2s, while C3 tensification existed in both the C2-deleted and the C2-preserved conditions. Given that the critical comparison was made between C2-deleted condition (in the absence of the physical information of /k/ or /p/) and C2-preserved condition (in the presence of the physical information of /k/ or /p/), one can still argue that C2-preserved condition carries direct acoustic-phonetic cues of the target, while C2-deleted condition requires phonological inference, which is precisely what the present study has tested. Second, the asymmetry in how listeners perceived /k/ versus /p/ in both C2-deleted and C2-preserved conditions has nothing to do with the fact that C3 was tensified (as opposed to the reviewer's suggestion), not only because C3 tensification did not provide a direct cue for whether the deleted phoneme was /k/ or /p/ in C2-deleted condition, but also because critical comparisons were made in relative terms—i.e., how the deleted /p/, for example, was recovered as compared with the deleted /k/, relative to their preserved counterparts. By the same token, the discussion about the difference between listener groups of 2002 and 2009 remains effective because exactly the same variants were used for both listener groups, and the discussion about the frequency of occurrence in speech production and its potential influence on speech perception is still viable because speech materials used in speech production study (Cho & Kim, 2009) and in the present study contained C3 tensification in all conditions.

5. Conclusion

The most important finding of the present study is that listeners indeed process phonologically modified variants due to consonant cluster simplification in Korean, apparently even

www.kci.go.kr

without acoustic-phonetic support (i.e., when the target was not even physically present). Although this conclusion is subject to further verification when the stimuli are confirmed to contain no acoustic phonetic residuals, such a process is not likely to happen without recourse to language-specific phonological knowledge. The results also indicate that the way listeners perceive the /p/-deleted variants have changed over time, as reflected in the difference in response latencies between listener groups of 2002 and 2009. The fact that listeners in 2009 were even faster in recovering deleted /p/ than detecting preserved /p/ was interpreted as having stemmed from the frequency effects in speech production—i.e., /p/-deleted variants occur far more frequently in contemporary Seoul Korean. Finally, listeners were faster in detecting /k/ than /p/, which could be accounted for by the listener-oriented production hypothesis: velars are perceptually more salient than labials, and therefore velars are processed more efficiently.

Speech variability poses challenges to theories of speech perception, as it is not easy to gauge exactly how listeners cope with it. Dealing with variability coming from phonological processes, researchers have adopted two basic assumptions. Listeners employ language-specific phonological knowledge in processing phonologically-driven speech variation or listeners reply on phonetic details that signal underlying representations using universally-applicable auditory-perceptual mechanisms. The present study has demonstrated a case which imposes strong constraints on models based on universal perceptual processes—i.e., language-specific phonological knowledge is called for when perceiving the results of continuous-speech production processes which leave no acoustic trace of underlying sounds as in the case of complete deletion of a consonant.

Acknowledgement

I thank students of Korea University in 2002 and Hanyang University in 2009 for volunteering to participate in the study. This work was supported by the research fund of Hanyang University (HY-2007-G).

References

Cho, Y. Yu. (1990). “Parameters of consonantal assimilation”, Ph.D. dissertation, Stanford University.

Cho, T. (1999). “Intra-dialectal variation in Korean consonant cluster simplification A stochastic approach”, *Chicago Linguistics Society*, Vol. 35 (CLS 35), pp. 43-57.

Cho, T. & Kim, S. (2009). “Statistical patterns in consonant cluster simplification in Seoul Korean Within-dialect interspeaker and intraspeaker variation”, *Phonetics and Speech Sciences* (Journal of the Korean Society of Speech Sciences), Vol. 1, pp. 33-40.

Cho, T. & McQueen, J. M. (2008). “Not all assimilated sounds are perceived equally Evidence from Korean”, *Journal of Phonetics*, Vol. 36, pp. 239-249.

Connine, C. M. (2004). “It's not what you hear but how often you hear it On the neglected role of phonological variant frequency in auditory word recognition”, *Psychonomic Bulletin & Review*, Vol. 11, pp. 1084-1089.

Flemming, E. (1995). “Auditory features in phonology”, Ph.D. diss. UCLA, Los Angeles, CA.

Gaskell, M. G. & Marslen-Wilson, W. D. (1996). “Phonological variation and inference in lexical access”, *Journal of Experimental Psychology: Human Perception and Performance,* Vol. 22, pp. 144-158.

Gaskell, M. G. & Marslen-Wilson, W. D. (1998). “Mechanisms of phonological inference in speech perception”, *Journal of Experimental Psychology: Human Perception and Performance,* Vol. 24, pp. 380-396.

Goldinger, S. D. (1996). “Words and voices: Episodic traces in spoken word identification and recognition memory”, *Journal of Experimental Psychology: Learning, Memory, and Cognition,* Vol. 22, pp. 1166-1183.

Goldinger, S. D. (1998). “Echoes of echoes?: An episodic theory of lexical access”, *Psychological Review,* Vol. 105, pp. 251-279.

Gow, D. W. (2001). “Assimilation and anticipation in continuous spoken word recognition”, *Journal of Memory and Language,* Vol. 45, pp. 133-159.

Gow, D. W. (2002). “Does English coronal place assimilation create lexical ambiguity?”, *Journal of Experimental Psychology: Human Perception and Performance,* Vol. 28, pp. 163-179.

Gow, D. W. (2003). “Feature parsing: Feature cue mapping in spoken word recognition”, *Perception and Psychophysics,* Vol. 65, pp. 575-590.

Gow, D. W. & Im, A. M. (2004). “A cross-linguistic examination of assimilation context effects”, *Journal of Memory and Language,* Vol. 51, pp. 279-296.

Hayes, B., Kirchner, R. & Steriade, D. (Eds.) (2004). *Phonetically based phonology.* Cambridge: Cambridge University Press.

Hume, E. & Johnson, K. (Eds.) (2001). *The role of speech*

www.kci.go.kr

perception in phonology. San Diego, CA.: Academic Press.

Iverson, G. & Lee, S. (1994). "Variation as optimality in Korean cluster reduction", In ESCOL '94.

Jun, J. (1995). "Perceptual and articulatory factors in place assimilation: An optimality theoretic approach", Ph.D. diss. UCLA.

Jun, J. 1998. "Restrictions on consonant clusters", in B.-S. Park and J. H.-S. Yoon (eds.), *Proceedings of 11th International Conference on Korean Linguistics* (11th ICKL): 382-390.

Jun, J. (2004). "Place assimilation", In B. Hayes, R. Kirchner and D. Steriade (Eds.), *Phonetically based phonology* (pp. 58-86). Cambridge: Cambridge University Press, Cambridge.

Kim-Renaud, Y.-K. (1976). "Korean consonantal phonology", Ph.D. diss. University of Hawaii.

Martin, S. E. (2006). *A reference grammar of Korean: A complete guide to the grammar and history of the Korean language.* Tuttle Publishing.

Mitterer, H., Cspe, V., Honbolygo, F. & Blomert, L. (2006). "The recognition of phonologically assimilated words does not depend on specific language experience", *Cognitive Science,* Vol. 30, pp. 451-479.

Pierrehumbert, J. (2001). "Exemplar dynamics: Word frequency, lenition, and contrast", in J. Bybee, and P. Hopper (Eds.), *Frequency effects and the emergence of linguistic structure* (pp. 137-157). Amsterdam: John Benjamins.

Pierrehumbert, J. (2002). "Word-specific phonetics", In C. Gussenhoven, and N. Warner (Eds.), *Laboratory Phonology VII* (pp. 101-140). Berlin: Mouton de Gruyter.

Pierrehumbert, J. (2003). "Phonetic diversity, statistical learning, and acquisition of phonology", *Language and Speech,* Vol. 46(2,3), pp. 115-154.

Ranbom, J. L. & Connine, C. M. (2007). "Lexical representation of phonological variation in spoken word recognition", *Journal of Memory and Language,* Vol. 57, pp. 273-298.

Steriade, D. (1999). "Phonetics in phonology the case of laryngeal neutralization", in M. Gordon (ed.), *UCLA Working Papers in Linguistics* no. 2 Papers in Phonology 3 (pp. 25-146).

Steriade, D. (2001). "Directional asymmetries in place assimilation a perceptual account", in B. Hume & K. Johnson (Eds.), *The role of speech perception in phonology* (pp. 219-250). San Diego, CA: Academic Press.

• **Cho, Taehong**
Department of English Language and Literature
Hanyang University,
Haengdang-dong 17, Seongdong-gu, Seoul (133-791), Korea
Email: tcho@hanyang.ac.kr

www.kci.go.kr