

Not all sounds in assimilation environments are perceived equally: Evidence from Korean

Taehong Cho^{a,b,*}, James M. McQueen^b

^a*Department of English Language and Literature, Hanyang University, 17 Haengdang-dong, Seongdong-gu, Seoul 133-791, Korea*

^b*Max Planck Institute for Psycholinguistics, Postbus 310, 6500 AH, Nijmegen, The Netherlands*

Received 24 July 2006; received in revised form 15 May 2007; accepted 1 June 2007

Abstract

This study tests whether potential differences in the perceptual robustness of speech sounds influence continuous-speech processes. Two phoneme-monitoring experiments examined place assimilation in Korean. In Experiment 1, Koreans monitored for targets which were either labials (/p,m/) or alveolars (/t,n/), and which were either unassimilated or assimilated to a following /k/ in two-word utterances. Listeners detected unaltered (unassimilated) labials faster and more accurately than assimilated labials; there was no such advantage for unaltered alveolars. In Experiment 2, labial–velar differences were tested using conditions in which /k/ and /p/ were illegally assimilated to a following /t/. Unassimilated sounds were detected faster than illegally assimilated sounds, but this difference tended to be larger for /k/ than for /p/. These place-dependent asymmetries suggest that differences in the perceptual robustness of segments play a role in shaping phonological patterns.

© 2007 Elsevier Ltd. All rights reserved.

1. Introduction

Different speech sounds are produced with distinct articulatory maneuvers, and hence have different acoustic properties. For example, the contact of the tongue dorsum against the point of the soft palate for /k/ forms a relatively larger cavity in front of the constriction. The acoustic consequence is spectral energy predominant in the central frequency region ('compactness'), which makes /k/ sound distinct from /t/ or /p/ (Jakobson, Fant, & Halle, 1952). These distinct acoustic properties of individual sounds are associated with differences in perception. For instance, some sounds in noise are easier to perceive (less confusable) than others (Miller & Nicely, 1955; Peterson & Barney, 1952), and some sounds tend to be recognized earlier than others (e.g., in a gating study, /p/ and /k/ in first position in Dutch diphones tended to be recognized earlier than /t/, Smits, Warner, McQueen, & Cutler, 2003; see also Warner, Smits, McQueen, & Cutler, 2003). We will define these differences in the efficiency and speed of segmental perception as differences in perceptual robustness. We ask here how perceptual robustness might influence continuous-speech processes such as place

*Corresponding author. Department of English Language and Literature, Hanyang University, 17 Haengdang-dong, Seongdong-gu, Seoul 133-791, Korea. Tel.: +82 2 2220 0740; fax: +82 2 2220 0741.

E-mail address: tcho@hanyang.ac.kr (T. Cho).

assimilation. Specifically, do differences in perceptual robustness lead to processing differences in detecting sounds when they are assimilated versus when they are unassimilated?

If some sounds are more perceptually robust than others, due to intrinsically different acoustic–phonetic properties, then it is possible that certain aspects of the sound patterns of the world’s languages are constrained by these perceptual differences. Robust sounds and weak sounds may be treated differently in the phonological systems of the world’s languages, so that, for example, some sounds may undergo certain phonological processes while others may not. This has indeed been one of the primary assumptions of theories of phonetically driven phonology (Flemming, 1995; Jun, 1995, 2004; Hayes, Kirchner, & Steriade, 2004; Hume & Johnson, 2001; Steriade, 1999, 2001). According to such theories, acoustic–phonetic properties constrain how certain sounds pattern together and undergo phonological alterations in ways that are helpful to listeners (although ease of articulation is also assumed to influence these patterns; Flemming, 1995, 2004; Kirchner, 2004). Similar views on the role of phonetics in shaping phonological patterns can also be found in some aspects of phonological acquisition (Kager, Pater, & Zonneveld, 2004) and phonological loanword adaptation (e.g., Kang, 2003; Shinohara, 2004; Silverman, 1992). One version of the view that phonology is grounded in phonetics is the *listener-oriented production hypothesis*: speakers make more effort to preserve sounds with more powerful acoustic cues, especially when sounds occur in potential phonological alteration environments, synchronic or diachronic, and they do so for the listener’s benefit (Jun, 1995, 2004; and see references therein).

Cross-linguistic evidence on place assimilation is consistent with theories of phonetically driven phonology. Across languages, regressive assimilation with codas as targets is more common than progressive assimilation with onset targets. Such directional asymmetries in place assimilation have been taken to be due to differences in perceptual robustness arising from the fact that richer acoustic cues are associated with (pre-vocalic) onsets than with (post-vocalic) codas (Bladon, 1986; Jun, 1995; Steriade, 2001; Wright, 2004). Moreover, place assimilation, whether regressive or progressive, is most likely to occur with underlyingly alveolar targets, less likely to occur with labial targets, and least likely to occur with velar targets (Jun, 1995, 2004). This ranking is consistent with the idea that acoustically weak sounds are more likely to undergo phonological alteration (e.g., Kohler, 1991; Ohala, 1990). Alveolars are thought to have the weakest acoustic cues, due to a shorter formant transition than velars (Ladefoged, 2000), which may be linked with a relatively fast tongue tip movement (Byrd, 1994; Kuehn & Moll, 1976). Velars, in contrast, are taken to have the most robust cues, such as compactness and convergence of F_2 and F_3 (e.g., Jakobson et al., 1952; Stevens, 1989), as well as a comparatively long formant transition due to sluggishness of the tongue body movement.

Few studies, however, have directly investigated the relative perceptual robustness of sounds in assimilation environments. Instead, the proposed ranking between labials and alveolars, for example, is grounded in phonological assumptions about the special status of coronals (including alveolars). Alveolars are claimed to be most vulnerable to phonological alterations because of the phonological underspecification of their coronal features or their unmarked nature (de Lacy, 2002; Hume, 2003; Lahiri & Marslen-Wilson, 1991; Paradis & Prunet, 1991). The proposed ranking of the perceptual robustness of alveolars, labials and velars could, however, account for the ranking of phonological alterability: over time, patterns of place assimilation in a given language could evolve to reflect perceptual robustness, such that, for example, the likelihood of phonological alterations (e.g., place assimilation) across languages might be reduced for perceptually more robust sounds such as velars, but have increased for less robust sounds such as alveolars (cf. Blevins, 2004; Jun, 1995, 2004).

We test here whether different sounds do indeed have different perceptual robustness in assimilation contexts. The case we examine is regressive place assimilation in Korean. Both alveolars and labials can be assimilated in place to following velars in casual Korean speech; velars themselves never undergo place assimilation in Korean (Kim-Renaud, 1976; Jun, 1996). Critically, although it is possible for Korean place assimilation to be phonetically incomplete, assimilation of both labials and alveolars to following velars is often complete and categorical in Korean (Gow & Im, 2004; Jun, 1996; Kochetov, Pouplier, & Son, 2006; Son, Kochetov, & Pouplier, in press). We used fully assimilated materials to control for any effects of residual phonetic cues in sounds which were assimilated. When listeners process assimilated sounds, they are likely to refer to their phonological knowledge (Gaskell & Marslen-Wilson, 1996, 1998; Lahiri & Marslen-Wilson, 1991) or to use acoustic remnants of the speaker’s intentions that might still be present in the speech signal (Gow, 2002; Gow & Im, 2004). But if assimilation is complete, listeners cannot use phonetic information and

must use phonological (including lexical) knowledge to retrieve the underlying segment. In contrast, when sounds are unassimilated, their intrinsic acoustic properties are present, and should of course play a role in perception. The comparison of the perception of underlying sounds that are either completely assimilated or unassimilated thus allowed us to test, in the following way, whether those sounds differ in perceptual robustness.

In Experiment 1 we presented Seoul Korean listeners with two-word utterances, and asked them to monitor for target phonemes in those utterances. The first syllable of the second word always contained an underlying alveolar (/t,n/) or labial (/p,m/) in coda position, followed by a velar /k/, the trigger of regressive place assimilation. The two-word utterances were presented with assimilated and unassimilated targets. Following Gaskell and Marslen-Wilson (1998), for example, listeners were asked to detect the underlying target sounds whether they were physically absent and replaced with another sound (in the assimilated condition) or present (in the unassimilated condition).

The critical predictions concerned differences in phoneme monitoring performance (speed and/or accuracy) between the assimilated and unassimilated conditions for each place of articulation of the underlying target. If labials are perceptually more robust than alveolars, the benefit in monitoring performance for unassimilated over assimilated targets should be larger for labials than alveolars. Listeners should benefit more from the presence of the hypothetically more robust labial consonants than from the presence of the less robust alveolars. Note that a simple comparison of the two unassimilated conditions is inadequate. If responses were faster or more accurate to unassimilated labials than to unassimilated alveolars, this would be consistent with the robustness hypothesis, but could also be due to other factors (e.g., the lexical contexts for the different types of target sound are necessarily different). The comparison of the assimilation effects across underlying place of articulation (i.e., the assimilation by place interaction) controls for contextual (lexical and sentential) factors. That is, because the same target had to be detected in the same context in both conditions but was physically present in only one (the unassimilated condition), differences between the assimilated and unassimilated conditions can be attributed to that sound's perceptual robustness rather than to the context.

In another phoneme-monitoring experiment (Experiment 2), we examined whether Korean listeners are sensitive to labial–velar differences, given that velars are assumed to be more robust than labials. Velars never undergo place assimilation in Korean, and labials are never assimilated when the following sound is an alveolar. Nevertheless, we created phonologically unviable conditions in which velars and labials were illegally assimilated to a following alveolar. Listeners were again asked to monitor for underlying targets both in these illegally assimilated contexts and in unassimilated contexts. If velars are perceptually more robust than labials, the benefit caused by the target being physically present (the difference between the unassimilated and assimilated conditions) should be larger for velars than for labials.

Finally, the size of the prosodic boundary in the target-bearing sequences was manipulated in both experiments. It has been suggested that spoken-word recognition can be influenced by information about prosodic boundaries (e.g., Christophe, Peperkamp, Pallier, Block, & Mehler, 2004; Cho, McQueen, & Cox, 2007). We therefore included a prosodic manipulation here to see whether phoneme recognition in assimilated versus unassimilated environments is also influenced by prosodic information. The two-word utterances used in each experiment were recorded either with an Accentual-Phrase (AP) boundary between the words, or a phrase-internal Word boundary. In Korean, AP boundaries are prosodic boundaries that are intermediate between prosodic-word and Intonational-Phrase boundaries (Jun, 1993, 2000). Articulation is stronger at the onset of larger prosodic domains than at the onset of smaller domains (*domain-initial strengthening*; for Korean, Cho & Keating, 2001; Jun, 1993). It was thus possible that monitoring of unassimilated targets would be better in the AP boundary context than in the Word boundary context.

2. Method

2.1. Participants

Fifty-two Korean student volunteers at Korean University were paid to participate in Experiment 1 (legal assimilation) and another 52 in Experiment 2 (illegal assimilation).

2.2. Materials

In the experimental trials in Experiment 1, target phonemes (/p,m/ or /t,n/) were always the codas of the first syllable in the second word of a two-word sequence. The following consonant was always /k/, and was the onset of the second word's second syllable (e.g., for /p/, [kamʃalil kup-ki-wihɛ] 'in order to bake the potatoes'; for /t/, [kaŋwilil tit-ki-wihɛ] 'in order to hear the lecture'; '-' indicates a morpheme boundary; note that the transcriptions are phonemic in IPA and the vowel /a/ in Korean is central (cf. Yang, 1996)). Both oral and nasal targets were included to see whether manner of articulation interacted with the assimilation manipulation. There were 24 words with labial targets (12 /p/-bearing and 12 /m/-bearing), and 24 with alveolar targets (12 /t/-bearing and 12 /n/-bearing).¹ A full listing of test stimuli for both experiments is given in the Appendix.

In addition to the 48 test sequences, 48 sequences were constructed as *foils*. Foil trials had no instances of the specified target (and thus no response was required on these trials). The presence of these trials made it impossible for participants to predict reliably whether a given trial would contain a target. The second words in the foil sequences were the same as the test words, but the first words were different from those used in the experimental items. Using the same word twice (in experimental and foil trials) could not influence listeners' performance on the experimental items because the foil trials always appeared later in the experiment than the experimental trials with the same word. Furthermore, 14 experimental sequences and 14 matched foils were constructed for another experiment, which was run concurrently. An additional 134 sequences were made, half as fillers and half as foils. There were in total 96 trials with /p,m/ and /t,n/ as targets and 162 with other targets. All targets had similar locations (the end of the first syllable of the second word), and the stimuli were matched across experiments in terms of prosodic boundaries. Twenty-eight practice items were also made. They included assimilated and unassimilated items, which were not used in the experimental trials, and some appropriate for the other experiment.

The materials in Experiment 2 were the same as in Experiment 1, except for experimental and matched foil trials. The target phoneme in the experimental trials (/p/ or /k/) was again the coda of the first syllable of the second word in two-word sequences, and the onset of the following syllable was always /t/. Unlike Experiment 1, there was no manipulation of manner of articulation (i.e., there were no nasal targets in Experiment 2). There were 12 words with labial (/p/) targets, 12 with velar (/k/) targets (see Appendix), and 24 matched foils. All of the fillers and filler foils for Experiment 1, including the materials for the concurrent experiment, were used in Experiment 2.

2.3. Procedure

The stimuli for both experiments were recorded by the first author (a male native Korean speaker and a trained phonetician) in a sound-damped booth. We aimed at obtaining completely assimilated forms, as they often occur in Korean (Gow & Im, 2004; Jun, 1996; Kochetov et al., 2006; Son et al., in press). To ensure complete assimilation for the stimuli, the speaker read two scripts, one with the targets written in their assimilated form and one with them in their unassimilated form. The assimilated forms were misspelled because, according to the *Hangul* system (the Korean orthography system), underlying phonemes should always be written. But because Hangul is phonologically fully transparent, any Korean speaker can consistently pronounce misspellings in the way intended. Therefore, completely assimilated forms were obtained when, for example, /pk/ was written as /kk/ (e.g., *kup-ki* as *kuk-ki*) and pronounced as such.

The same two scripts were read several times, with either an AP boundary or a phrase-internal Word boundary between the words in each sequence. The best token of each sequence (the one that was perceived to be most naturally pronounced as intended) was selected for use in the experiments. The target sounds were never released, and thus were cued primarily by formant transitions in the preceding vowel. The following

¹In almost all cases in Experiment 1, the assimilated forms had single possible underlying forms. Only two of the 24 assimilated forms in each of the place of articulation conditions were lexically ambiguous. The presence of these lexical competitors could have made it harder for listeners to recover the intended underlying targets, but any such effects were balanced across place of articulation.

stops (/k/ in Experiment 1, /t/ in Experiment 2) were always released. These realizations are appropriate for coda and onset stop consonants, respectively, in Korean (Cho & McQueen, 2006).

There were four groups of 13 participants in each experiment. Within each experiment two groups received all the materials recorded with AP boundaries, and two received those with Word boundaries. Two lists were constructed for each experiment, with the assimilation condition (assimilated or unassimilated) counter-balanced over lists. One group of listeners within each pair (i.e., within each experiment and each prosodic boundary condition) heard each list. Each listener therefore heard all the foils and fillers, and one version of each of the test words appropriate for each experiment, but never both the assimilated and unassimilated form of the same test item.

Participants were asked to listen to the stimuli, and to monitor for pre-specified target phonemes in a specific position, namely, within the first syllable of the second word in each utterance. They were instructed to press a button as fast and as accurately as possible when they detected a target. The targets appeared for one second on a computer screen, in Hangul (e.g., ㅂ , /p/ and ㅌ , /t/), and the auditory stimuli began immediately afterwards. A new target was presented on every trial. Response latencies relative to the onset of the target (i.e., the acoustic offset of the preceding vowel) and errors were recorded.

In order to ensure that the speaker had reliably produced utterances with AP and Word boundaries, we measured the duration of the preboundary word and of the initial portion of the postboundary syllable, up to the onset of the target consonant. For items in Experiment 1, ANOVAs showed that preboundary-word duration was longer before AP than Wd boundaries (mean diff. 28 ms; $F[1,92] = 6.35$, $p < 0.02$), with no interaction with Assimilation or Place. The duration of the postboundary CV was longer after AP than Wd boundaries (mean diff. 18.5 ms; $F[1,92] = 8.55$, $p < 0.005$). There were no interactions with other factors. For Experiment 2 items, there were no significant effects on preboundary word duration. But postboundary CV duration was again longer after AP than after Wd boundaries (mean diff. 12.8 ms; $F[1,44] = 4.91$, $p < 0.05$), with again no interactions with other factors. These patterns are consistent with other findings that word-initial segments are produced with significantly increased duration in initial position of APs as compared to phrase-internal words (Cho & Keating, 2001; Jun, 1993), whereas preboundary lengthening before AP boundaries is either small but significant (Cho & Keating, 2001), consistent with the preboundary lengthening effect found for Experiment 1, or unreliable (Jun, 1995), in line with the absence of preboundary lengthening in Experiment 2.

3. Results

3.1. Experiment 1

Mean latencies and error rates are summarized in Fig. 1. In ANOVAs, on latencies with the factors Boundary (AP, Wd), Place (alveolars, labials), Manner (nasal, oral), and Assimilation (assimilated,

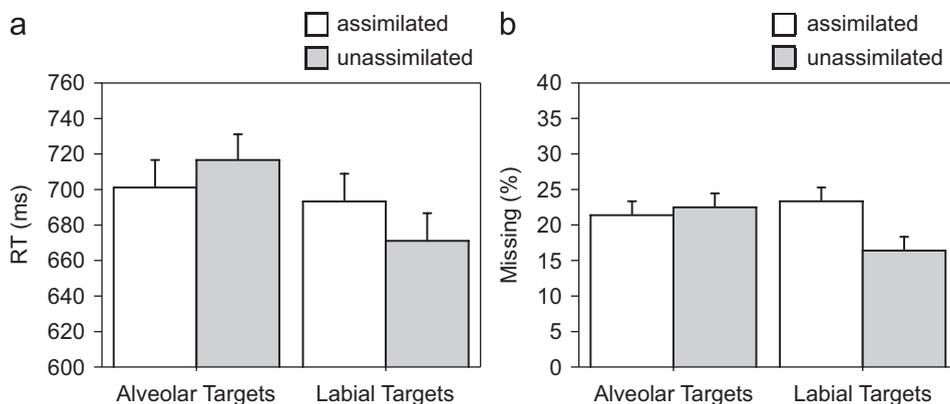


Fig. 1. Legal assimilation (Experiment 1): reaction time (RT, in ms) and percent missing responses for alveolar and labial targets either assimilated or unassimilated to the following /k/ (error bars indicate standard errors).

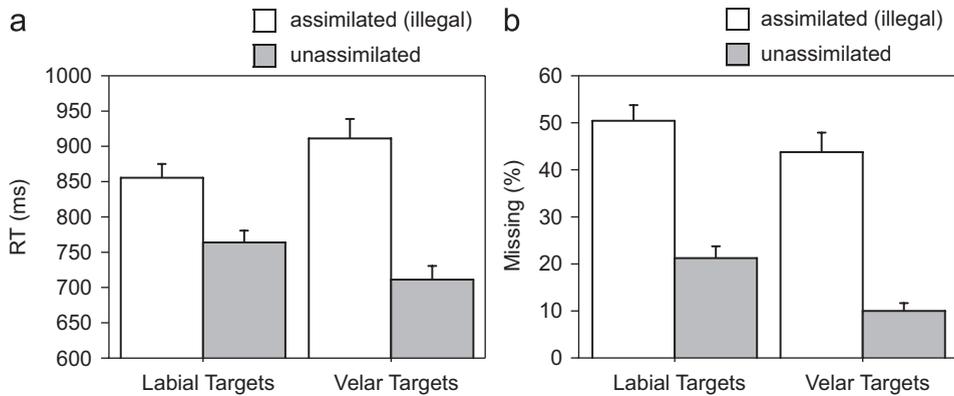


Fig. 2. Illegal assimilation (Experiment 2): reaction time (RT, in ms) and percent missing responses for velar and labial targets either illegally assimilated or unassimilated to the following /t/ (error bars indicate standard errors).

unassimilated), only Place showed a main effect, which was significant by participants ($F_1[1,48] = 10.17$, $p < 0.005$) but not by items ($F_2[1,40] = 2.77$, $p < 0.1$). There was thus a general trend towards more rapid detection of labials than alveolars. The interaction between Place and Assimilation was significant ($F_1[1,48] = 6.92$, $p < 0.02$; $F_2[1,40] = 4.50$, $p < 0.05$). As can be seen in Fig. 1a, for labial targets, unassimilated targets were detected more rapidly than assimilated ones (mean diff. 22 ms; $t_1(51) = 2.15$, $p < 0.05$; $t_2(23) = 2.47$, $p < 0.025$), whereas detection of alveolar targets did not differ as a function of assimilation (mean diff. -14 ms; $t_1(51) = -1.39$, $p > 0.1$; $t_2(23) < 1$). This place-dependent asymmetry was also evident in comparisons between alveolar and labial targets. Unassimilated labial targets were detected faster than unassimilated alveolar targets (mean diff. 46 ms; $t_1(51) = 4.12$, $p < 0.001$; $t_2(46) = 2.41$, $p < 0.02$). No such difference was found when targets were assimilated (mean diff. 8 ms; $t_1(51) < 1$, $t_2(46) < 1$). There were no other significant interactions involving the Assimilation factor, suggesting that the effect of assimilation was not conditioned by prosodic boundary or manner.

The results of ANOVAs on the errors were similar. There were no significant main effects, and the only significant interaction involving Assimilation was with Place, which was significant by participants ($F_1[1,48] = 4.22$, $p < 0.05$), but only marginally so by items ($F_2[1,40] = 3.35$, $p < 0.075$). Labials were detected more accurately when unassimilated than assimilated (mean diff. 6.9%; $t_1(51) = 2.53$, $p < 0.02$; $t_2(23) = 2.47$, $p < 0.025$) while there was no difference in monitoring accuracy for alveolars (mean diff. -1.1%; $t_1(51) < 1$, $t_2(23) < 1$). Comparisons between labials and alveolars confirmed the place-dependent pattern. Detection was more accurate for labials than for alveolars when targets were unassimilated (mean diff., 6%; $t_1(51) = 2.27$, $p < 0.03$; but $t_2(46) = 1.41$, $p < 0.2$), but not when they were assimilated (mean diff., -2%; $t_1(51) < 1$, $t_2(46) < 1$).²

3.2. Experiment 2

Mean latencies and error rates are given in Fig. 2. In ANOVAs on latencies with the factors Boundary (AP, Wd), Place (/k/, /p/) and (Illegal) Assimilation (assimilated, unassimilated), the only significant main effect was

²As a reviewer pointed out, vocalic formant transitions for alveolars are weaker when followed by front vowels than by back vowels. The reviewer also pointed out that the identification data in Mayo and Turk (2005) suggest that the front vowel environment (especially /i/) causes problems for both labial and alveolar identification. In the present study, however, there were 7 front vowels (30%) and 17 non-front vowels (70%) in the alveolar condition and in the labial condition. (In the front vowel environment, there were seven /i/ cases for alveolar targets and six /i/ cases and one /ε/ case for labial targets; in the non-front vowel environment, there were eight /a/ cases for both alveolar and labial targets and the remaining 16 cases involved /ə, o, u, i/; note that the Korean vowels /a, i, ə/ are central in the F_1 - F_2 vowel space, cf. Yang, 1996). Nevertheless, we also re-ran the analyses excluding those items that contained front vowels in alveolar contexts. The results showed virtually the same patterns. There was again a Place \times Assimilation interaction significant both by participants and by items in the latency analysis ($F_1[1,48] = 5.47$, $p < 0.025$; $F_2[1,37] = 4.88$, $p < 0.05$), but not in the error analysis ($F_1[1,48] = 2.33$, $p > 0.1$; $F_2[1,37] = 1.88$, $p > 0.1$). Pairwise comparisons showed there was still no difference between unassimilated and assimilated alveolars (latency, mean diff. -11 ms; $t_1(51) = -1.3$, $p > 0.2$, $t_2(16) < 1$; accuracy, mean diff. -0.25%; $t_1(51) < 1$, $t_2(16) < 1$).

an Assimilation effect ($F_1[1,46] = 77.94, p < 0.001$; $F_2[1,20] = 47.29, p < 0.001$). Illegally assimilated targets were detected far more slowly than unassimilated targets (mean diff. 147 ms). The factor Assimilation again interacted with Place. This interaction was significant by participants ($F_1[1, 46] = 16.92, p < 0.001$), though not by items ($F_2[1,20] = 1.51, p > 0.1$). As shown in Fig. 2a, although detection of both /k/ and /p/ was faster when they were unassimilated than when they were illegally assimilated, the effect was larger for /k/ than for /p/ (mean diff. for /k/, 201 ms; $F_1[1,51] = 8.28, p < 0.001, \eta^2 = 0.573$; $F_2[1,11] = 3.83, p < 0.005, \eta^2 = 0.571$; mean diff. for /p/ 91 ms; $F_1[1,51] = 4.01, p < 0.001, \eta^2 = 0.246$; $F_2[1,11] = 3.08, p < 0.02, \eta^2 = 0.463$). There were no other significant interactions involving the Assimilation factor.

The error analyses also showed a main effect of (Illegal) Assimilation ($F_1[1,48] = 178.43, p < 0.001$; $F_2[1,20] = 23.79, p < 0.001$), but this time there was no Assimilation \times Place interaction ($F_1[1,48] < 1$; $F_2[1,20] < 1$). Illegally assimilated targets were detected less accurately than unassimilated targets regardless of whether they were /k/ or /p/ (see Fig. 2b).³

4. Discussion

Experiment 1 showed that listeners' recognition of assimilated and unassimilated sounds was conditioned by place of articulation, that is, whether the targets were labials or alveolars. Labial targets were detected more rapidly and accurately in the unassimilated than in the assimilated condition; there was no difference between these conditions for the alveolar targets. Although the results of Experiment 2 were less statistically robust, they showed a similar pattern: There was once again an asymmetry across place of articulation in phoneme monitoring performance, such that the difference between the assimilation conditions (unassimilated versus illegally assimilated targets) was larger for velar than for labial targets.

Manner of articulation (oral versus nasal targets) did not modulate the place-asymmetry effect in Experiment 1. The effect thus appears to be a genuine effect of place of articulation, rather than something specific to particular phonemes. The prosodic-prominence manipulation (target-bearing words after AP versus phrase-internal Word boundaries) also had no effect. Differences between the target-present and target-absent (unassimilated versus assimilated) conditions in both experiments thus do not depend upon changes in articulation of the targets (or their assimilated variants) that are prosodically conditioned.

The main pattern in this study, therefore, is a place-dependent asymmetry. There appears to be a benefit to perception when targets are physically present for labials but not for alveolars, and which tended to be stronger for velars than for labials. Our results thus suggest that sounds in assimilation environments differ in perceptual robustness. Velars, by virtue of the place they are articulated and the resulting salient acoustic cues, tend to be perceived more quickly and more accurately than labials, which in turn, and again because of acoustic differences due to place of articulation, tend to be perceptually more robust than alveolars.

These findings are consistent with theories of phonetically driven phonology regarding the role of phonetics in place assimilation. The role of phonetic factors in phonology has been argued to be a diachronic one (e.g., Blevins, 2004), such that, for example, perceptually more robust sounds such as velars may have resisted phonological alterations over time, while the likelihood of phonological alterations may have increased for perceptually less robust sounds such as alveolars. The current findings provide synchronic evidence that listeners benefit more from the acoustic properties of some sounds (e.g., velars or labials) than of other sounds (e.g., alveolars) in line with the diachronically motivated view that the perceptually weakest consonants are the ones that are most vulnerable to phonological alteration (see Jun, 1995, 2004).

³When /k/ and /p/ were illegally assimilated to /t/, lexical competitors were created in most cases (10 and 9 out of 12 items for the underlying /k/ and /p/, respectively). This was hard to avoid because in Korean there is a very robust phonological process which neutralizes all coronal coda consonants (e.g., /t, t^h, t*, s, s*, tʃ, tʃ^h, tʃ*) to [t]. Furthermore, there were two items with front vowels in the /k/ condition but five in the /p/ condition. In order to avoid any confounds coming from the presence versus absence of the competitors and the unbalanced front-back vowel contexts, we conducted ANOVAs excluding items which contained front vowels and whose assimilation did not result in competitors. The results, however, yielded no differences as compared to the original analyses (latency: main effect of Assimilation, $F_1[1, 46] = 24.55, p < 0.001$; $F_2[1,10] = 29.39, p < 0.001$; Assimilation \times Place interaction, $F_1[1,46] = 8.91, p < 0.01$; $F_2[1,10] < 1$; error: main effect of Assimilation, $F_1[1,48] = 28.01, p < 0.0001$; $F_2[1,10] = 43.34, p < 0.0001$; Assimilation \times Place interaction, $F_1[1,46] < 1, F_2[1,10] < 1$).

The current findings may also be considered in terms of the frequency of occurrence of assimilated forms in the language. If alveolars are more frequently assimilated to velars than labials are in present-day Korean and therefore listeners are exposed to assimilated alveolars more often than to assimilated labials, one might expect that they can recover the underlying form of the assimilated alveolars more easily than the assimilated labials. The fact that the assimilated–unassimilated difference in latencies was found for labials but not for alveolars in Experiment 1 could thus be attributable in part to the frequent occurrence of assimilated forms for alveolar targets. While this is an alternative or an additional account for the current findings, it must remain speculative, since little is known about the relative frequency of occurrence of different type of assimilation in present-day Korean. We would predict, however, on the basis of our results, that alveolars do assimilate in place to following velars more often than labials do. Note also that if the present results would prove to be due at least in part to frequency effects, then those frequency effects would still likely reflect effects of perceptual robustness (cf. Cho, 1999). We have already argued that differences in perceptual robustness of consonants across place of articulation may have shaped diachronic phonological change. But they may also influence synchronic phonological alterations, as reflected in the frequency of occurrence of assimilations.

The present study also has implications for theories of speech perception, especially for accounts of how listeners deal with the effects of continuous-speech processes such as place assimilation. A number of accounts of recovery from assimilation have been proposed, including those based on underspecification theory (Lahiri & Marslen-Wilson, 1991), on phonological inference (Gaskell & Marslen-Wilson, 1996, 1998), and on the parsing of acoustic–phonetic features (Gow, 2002). An important issue concerning all three of these theories is whether recovery from assimilation depends on language-specific phonological knowledge or on language-universal perceptual processes (Gow & Im, 2004; Mitterer, Csépe, Honbolygo, & Blomert, 2006; Otake, Yoneyama, Cutler, & van der Lugt, 1996; Weber, 2001). Distinguishing among theories of recovery from assimilation was not the aim of this study. Nevertheless, our results do inform this debate. Because of differences in perceptual robustness, there may, for example, be differences across place of articulation in the listener’s ability to make phonological inferences (Gaskell & Marslen-Wilson, 1996, 1998), or place of articulation differences in the extent to which phonological processes leave residual acoustic–phonetic traces of underlying sounds, upon which feature parsing processes may act (Gow, 2002). Thus, while it remains to be determined exactly what the mechanisms are that listeners use to recover from place assimilation, any account must take into consideration the fact that there may be differences in the inherent perceptual strength of different segments. This conclusion also applies more generally. Theories of spoken word recognition need to take into account the possibility that speech sounds are not all perceived equally—not just those sounds involved in place assimilation but also those involved in other phonological alternations.

Acknowledgments

We thank Kee-Ho Kim for making space available for testing at Korea University. We also thank three reviewers for constructive comments. A preliminary report of this research was presented at the 9th Conference on Laboratory Phonology, Urbana-Champaign, June 2004. This work was supported by the research fund of Hanyang University (HY-2006).

Appendix

The full set of experimental stimuli. Targets and the following phonemes are in bold. The transcriptions are phonemic in IPA

A. Experiment 1: *Legal assimilation (/t/-to-/k/ and /p/-to-/k/ assimilation)*

Alveolar targets (/t, n/)

- | | |
|-------------------------|--|
| 1. ʧantonil patkiwihe | ‘in order to receive the change’ |
| 2. mulil kitkonasə | ‘after drawing water (with a ladle)’ |
| 3. hwepilil kətkonasə | ‘after collecting the membership fees’ |
| 4. simʧika kotketwemjən | ‘if one’s will gets firmed up’ |

5. tɛmunil tatkɪwiɛ	‘in order to close the gate’
6. jækɕjɔŋil titkɔnasə	‘after overcoming the hardship’
7. jɔŋtonil ɔtkɔnasə	‘in order to get (some) allowance’
8. kaŋŋilil titkɪwiɛ	‘in order to hear the lecture’
9. mənʃika mutketwɛmjən	‘if (one) gets dusty’
10. aʃ ^h imhɛka totketwɛmjən	‘when the morning sun rises’
11. tantanhake kutketwɛmjən	‘when (something) gets hardened’
12. somunil mitkɔnasə	‘after believing the rumor’
13. sat ^h ɛka nankamhaja	‘because things get beyond one’s capacity’
14. ʃɛmis*ko sinkihɛsə	‘because things are fun and wondrous’
15. kutulil sinkɔnamjən	‘when one wears (one’s) dress shoes’
16. inhɛŋi ^h ʃankolonin	‘as for the bank account’s balance’
17. kɕjɔŋʃ ^h alɛ sinkohaja	‘by reporting to the police’
18. usimjə pankimjənin	‘if one greets (someone) smiling’
19. ʃikisi nunkamimjən	‘if one closes (one’s) eyes gently’
20. put ^h aki kankokhaja	‘because one’s request is serious’
21. jəuka tunkaphaja	‘as the wolf transforms (itself into something)’
22. kjejak ^h i ^h mankienin	‘when the contract ends’
23. sat ^h ɛka kinkiphake	‘for the situation to become urgent’
24. koʃipil wankohake	‘for one’s stubbornness to be persistent’

Labial targets (/p, m/)

1. ʃamosil ipkiwiɛ	‘in order to wear pajamas’
2. nɛŋmjəni mɛpketwɛmjən	‘if the cold noodles get spicy’
3. munʃɛka swipketwɛmjən	‘if the questions become easy’
4. pəmtʃwɛka mipketwɛmjən	‘when one hates the crime’
5. tɪŋɛta əpkiwiɛ	‘in order to take (someone) on one’s back’
6. kamʃalil kupkiwiɛ	‘in order to bake the potatoes’
7. ʃ ^h imtɛ nupkiwiɛ	‘in order to lie down on a bed’
8. iusil topkɔnamjən	‘when (one) helps out the neighbors’
9. jəlimi tɔpketwɛmjən	‘if the summer is hot’
10. sonilo ʃapɕnasə	‘after grabbing (something) with one’s hand’
11. mulkənil ʃipɕnasə	‘after picking an object up’
12. nals*ika ʃ ^h upketwɛmjən	‘when the weather gets cold’
13. hipjənil samkahaja	‘by refraining from smoking’
14. misinil səmkikosə	‘as one worships the supernatural power’
15. ap*aka ʃamk*otɛlil	‘for the dad to talk in his sleep’
16. salinʃwɛlo kəmkətweə	‘by being arrested for the murder’
17. k*aktukilil tamkɔnasə	‘after making K*aktuki (kind of Kimchi)’
18. tɛmunil ʃamkikosə	‘after closing the gate’
19. tɛt ^h ɔŋljɔŋi ^h imkienin	‘as for the president’s term’
20. məlilil kamkɔnasə	‘after shampooing (one’s) hair’
21. namulil simkɔnasə	‘after planting trees’
22. ʃ ^h uunte kamkitimjən	‘if one catches a cold because of the cold weather’
23. not ^h ilil ʃ ^h amkohaja	‘by referring to the notes’
24. ʃəkʃinil ʃəmkəhaja	‘by occupying the enemy’s post’

B. Experiment 2: Illegal assimilation (/p/-to-/t/ and /k/-to-/t/ assimilation)

/p/ targets

1. ʃamosil iptɛlato	‘even if (one) wears pajamas’
2. nɛŋmjəni mɛptɛlato	‘even if the cold noodles are spicy’

3. muntʃeka swiptəlato	‘even if the questions are easy’
4. pəmtʃweka miptəlato	‘even if the crime is hateful’
5. akilil əptəlato	‘even if the baby cries’
6. kamʃʃalil kuptəlato	‘even if (one) bakes the potatoes’
7. ʃ ^h imtæ nuptəlato	‘even if (one) lies down on a bed’
8. iusil toptəlato	‘even if (one) helps out the neighbors’
9. nals*ika ʃ ^h uptəlato	‘even if the weather is cold’
10. sonilo ʃʌptəlato	‘even if (one) grabs it with one’s hand’
11. mulkənil ʃiptəlato	‘even if (one) picks an object up’
12. jəlɪmi təptəlato	‘even if the summer is hot’
/k/ targets	
1. mosilo paktəlato	‘even if you hammer it with a nail’
2. jəlməka iktəlato	‘even if the fruits become ripened’
3. pʒəntilə ʃuktəlato	‘even if (one) died of illness’
4. əlimi noktəlato	‘even if the ice melts’
5. jənpillo ʃəktəlato	‘even if (one) writes (it) with a pencil’
6. hjuʃjilo taktəlato	‘even if (one) wipes (it) with tissues’
7. kikika ʃəktəlato	‘even if the size is small’
8. ʃ ^h imipil maktəlato	‘even if (one) defends against the invasion’
9. jəkwane muktəlato	‘even if (one) stays at a motel’
10. kimʃ ^h ika saktəlato	‘even if the Kimchi gets fermented’
11. k ^h əp ^h ika siktəlato	‘even the coffee gets cold’
12. nameke soktəlato	‘even if (one) is deceived by someone’

References

- Bladon, A. (1986). Phonetics for hearers. In G. McGregor (Ed.), *Language for hearers* (pp. 1–24). Oxford: Pergamon.
- Blevins, J. (2004). *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.
- Byrd, D. (1994). *Articulatory timing in English consonant sequences*. Ph.D. dissertation, UCLA, Los Angeles, CA.
- Cho, T. (1999). Intra-dialectal variation in Korean consonant cluster simplification: A stochastic approach. *Chicago Linguistics Society*, 35(CLS 35), 43–57 (Also appeared in: M. Gordon (Ed.), *UCLA working papers in linguistics* (Vol. 2), *Papers in Phonology* 3 (pp. 207–226).)
- Cho, T., & Keating, P. (2001). Articulatory and acoustic studies of domain-initial strengthening in Korean. *Journal of Phonetics*, 29, 155–190.
- Cho, T., & McQueen, J. M. (2006). Phonological versus phonetic cues in native and nonnative listening: Korean and Dutch listeners’ perception of Dutch and English consonants. *Journal of the Acoustical Society of America*, 119, 3085–3096.
- Cho, T., McQueen, J. M., & Cox, E. (2007). Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics*, 35, 210–243.
- Christophe, A., Peperkamp, S., Pallier, C., Block, E., & Mehler, J. (2004). Phonological phrase boundaries constrain lexical access. I. Adult data. *Journal of Memory and Language*, 51, 523–547.
- de Lacy, P. (2002). *The formal expression of markedness*. Ph.D. dissertation, University of Massachusetts.
- Flemming, E. (1995). *Auditory features in phonology*. Ph.D. dissertation, UCLA, Los Angeles, CA.
- Flemming, E. (2004). Contrast and perceptual distinctiveness. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetically based phonology* (pp. 232–276). Cambridge: Cambridge University Press.
- Gaskell, M. G., & Marslen-Wilson, W. D. (1996). Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 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*, 24, 380–396.
- Gow, D. W. (2002). Does English coronal place assimilation create lexical ambiguity? *Journal of Experimental Psychology: Human Perception and Performance*, 28, 163–179.
- Gow, D. W., & Im, A. M. (2004). A cross-linguistic examination of assimilation context effects. *Journal of Memory and Language*, 51, 279–296.
- Hayes, B., Kirchner, R., & Steriade, D. (Eds.). (2004). *Phonetically based phonology*. Cambridge: Cambridge University Press.
- Hume, E. (2003). Language specific markedness: The case of place of articulation. *Studies in Phonetics, Phonology and Morphology*, 9, 295–310.
- Hume, E., & Johnson, K. (Eds.). (2001). *The role of speech perception in phonology*. San Diego, CA: Academic Press.

- Jakobson, R. C., Fant, G. M., & Halle, M. (1952). *Preliminaries to speech analysis: The distinctive features and their correlates*. Cambridge, MA: MIT Press.
- Jun, S.-A. (1993). *The phonetics and phonology of Korean prosody*. Unpublished Ph.D. dissertation, The Ohio State University, Columbus, OH.
- Jun, J. (1995). *Perceptual and articulatory factors in place assimilation: An optimality theoretic approach*. Ph.D. dissertation, UCLA.
- Jun, J. (1996). Place assimilation is not the result of gestural overlap: Evidence from Korean and English. *Phonology*, 13.3, 377–407.
- Jun, S.-A. (2000). K-ToBI (Korean ToBI) labeling conventions: Version 3. *Speech Sciences*, 7, 143–169 (also published in *UCLA Working Papers in Phonetics* 99 149–173).
- Jun, J. (2004). Place assimilation. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetically based phonology* (pp. 58–86). Cambridge: Cambridge University Press.
- Kager, R., Pater, P., & Zonneveld, W. (Eds.). (2004). *Constraints in phonological acquisition*. Cambridge: Cambridge University Press.
- Kang, Y. (2003). Perceptual similarity in loanword adaptation: Adaptation of English post-vocalic word-final stops to Korean. *Phonology*, 20, 219–273.
- Kim-Renaud, Y.-K. (1976). *Korean consonantal phonology*. Ph.D. dissertation, University of Hawaii.
- Kirchner, R. (2004). Consonant lenition. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetically based phonology* (pp. 313–345). Cambridge: Cambridge University Press.
- Kochetov, A., Pouplier, M., & Son, M. (2006). *Variability and gradience in Korean place assimilation*. Paper presented at the international conference on East Asian Linguistics, Toronto, November.
- Kohler, K. J. (1991). The phonetics/phonology issue in the study of articulatory reduction. *Phonetica*, 48, 180–192.
- Kuehn, D. P., & Moll, M. L. (1976). A cineradiographic study of VC and CV articulatory velocities. *Journal of Phonetics*, 4, 303–320.
- Ladefoged, P. (2000). *A course in phonetics* (4th ed). Fort Worth, TX: Harcourt Brace Jovanovich.
- Lahiri, A., & Marslen-Wilson, W. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition*, 38, 245–294.
- Mayo, C., & Turk, A. (2005). The influence of spectral distinctiveness on acoustic cue weighting in children's and adults' speech perception. *Journal of the Acoustical Society of America*, 118, 1730–1741.
- Miller, G. A., & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America*, 27, 338–352.
- Mitterer, H., Csépe, V., Honbolygo, F., & Blomert, L. (2006). The recognition of phonologically assimilated words does not depend on specific language experience. *Cognitive Science*, 30, 451–479.
- Ohala, J. J. (1990). The phonetics and phonology of aspects of assimilation. In J. Kingston, & M. E. Beckman (Eds.), *Papers in laboratory phonology: Between the grammar and physics of speech* (pp. 258–275). Cambridge: Cambridge University Press.
- Otake, T., Yoneyama, K., Cutler, A., & van der Lugt, A. (1996). The representation of Japanese moraic nasals. *Journal of the Acoustical Society of America*, 100, 3831–3842.
- Paradis, C., & Prunet, J.-F. (Eds.). (1991). *The special status of coronals: Internal and external evidence*. San Diego: Academic Press.
- Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of vowels. *Journal of the Acoustical Society of America*, 24, 175–184.
- Shinohara, S. (2004). Emergence of universal grammar in foreign word adaptations. In R. Kager, J. Pater, & W. Zonneveld (Eds.), *Constraints in phonological acquisition* (pp. 292–320). Cambridge: Cambridge University Press.
- Silverman, D. (1992). Multiple scansions in loanword phonology: Evidence from Cantonese. *Phonology*, 9, 289–328.
- Smits, R., Warner, N., McQueen, J. M., & Cutler, A. (2003). Unfolding of phonetic information over time: A database of Dutch diphone perception. *Journal of the Acoustical Society of America*, 113, 563–574.
- Son, M., Kochetov, A., & Pouplier, M. (in press). The role of gestural overlap in perceptual place assimilation in Korean. In J. Cole & J. I. Hualde (Eds.), *Papers in laboratory phonology* (Vol. 9), Mouton de Gruyter.
- 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* (vol. 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.
- Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics*, 17, 3–45.
- Warner, N., Smits, R., McQueen, J. M., & Cutler, A. (2003). Phonological and statistical effects on timing of speech perception: Insights from a database of Dutch diphone perception. *Speech Communication*, 46, 53–72.
- Weber, A. (2001). Help or hindrance: How violation of different assimilation rules affects spoken-language processing. *Language and Speech*, 44, 95–118.
- Wright, R. (2004). A review of perceptual cues and cue robustness. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetically based phonology* (pp. 34–57). Cambridge: Cambridge University Press.
- Yang, B. (1996). A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics*, 24, 245–261.