

A perception-based analysis of voice onset time (VOT) dissimilation in Korean

Hijo Kang¹ · Mira Oh^{2,*}

¹Department of English Language Education, Chosun University, Gwangju, Korea

²Department of English Language and Literature, Chonnam National University, Gwangju, Korea

Abstract

This study examines the perceptual motivation behind dissimilation. Consistent with previous arguments suggesting that dissimilation originates from perception rather than production (Coetzee, 2005; Kiparsky, 2003; Scheer, 2013), we hypothesized that an oral stop with short of voice onset time (VOT) would be recognized as non-aspirated more often when it is followed by an aspirated stop with a long VOT. This hypothesis was tested through a perception experiment in which 32 Korean listeners made judgments on the first consonant of C₁VC₂V words manipulated with C₁ VOT and C₂ types. The results revealed that aspirated-based C₁ was recognized as aspirated or tense depending on the duration of VOT, while lenis-based C₁ was consistently recognized as lenis. The dissimilatory effect of aspirated C₂ was confirmed as anticipated, and furthermore, tense C₂ increased the ratio of tense responses more than aspirated C₂. These results provide evidence of a perceptual bias against recurrent aspirated stops, which may play a role in activating a dissimilatory rule or constraint in a language. The assimilatory effect of tense C₂ is in consistent with findings indicating that word-initial tensification is facilitated by the following tense stop in Korean (Kang & Oh, 2016; H. Kim, 2016).

Keywords: dissimilation, perception, aspirated stops, voice onset time (VOT), Korean

1. Introduction

This study investigates the perceptual motivation of dissimilation by means of a perception experiment with Korean oral stops of three distinctive laryngeal features. Dissimilation is a phonological process in which a segment becomes less similar to a nearby segment concerning a specific feature (Bye, 2011). It was initially observed in diachronic studies of language. For instance, the Latin word *arbor* for ‘tree’ evolved into *arbol* in Spanish, and the Classical

Greek word *hepta* for ‘seven’ became *efla* in Modern Greek (Cser, 2013). Subsequently, typological studies demonstrated dissimilation as a synchronic phonological process in various languages (Bye, 2011; Suzuki, 1988; Walter, 2007). It was also found that dissimilation is present in the lexicon as a phonotactic constraint (Davis, 1991; Gallagher, 2010; MacEachern, 1999; McCarthy, 1986; Yip, 1989). Recently, probabilistic dissimilatory patterns in the lexicon were discovered to influence morphophonological processes in Korean (Ito, 2014; Kang & Oh, 2019; Kim S., 2016).

* mroh@chonnam.ac.kr, Corresponding author

Received 30 January 2024; Revised 10 March 2024; Accepted 11 March 2024

© Copyright 2024 Korean Society of Speech Sciences. This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Dissimilation shares similarities with assimilation, although it is conceptually the opposite. Firstly, the trigger and the target can be either adjacent or distant from each other. Secondly, the direction of the effect can be progressive or regressive. Thirdly, the features involved in both processes encompass various phonological features such as place, manner, and laryngeal ones.

Despite having many similarities, dissimilation does not occur as frequently as assimilation does. Diachronically and synchronically, assimilation is by far more productive and extensive than dissimilation. What is the cause of such asymmetry between these two phonological processes? The answer to this question could be explored in the origins of the processes. According to Kiparsky (2003), sound change originates through synchronic variation in the production, perception, and acquisition of language. While the variation in speech production is gradient, arising from natural articulatory processes, the variation stemming from perception and acquisition drives alternating parsing, which can proceed in abrupt discrete steps.

Assimilation is known to be rooted in articulatory variation, even when it occurs between non-adjacent segments. For instance, vowel harmony, a type of assimilation, has been argued to result from vowel-to-vowel coarticulation (Beddor et al., 2002; Cho, 2004; Gay, 1974, 1977; Manuel, 1990; Manuel & Krakow, 1984; Öhman, 1966, among others). Similarly, consonantal harmony is claimed to be motivated by the temporal extension of articulatory gestures (Gafos & Lombardi, 1999; Walker et al., 2008; Whalen et al., 2011).

While assimilation is primarily attributed to articulatory variability, which is gradient, dissimilation is not considered a natural articulatory process. It is thought to arise abruptly through perceptual reanalysis (Kiparsky, 2003). In historical phonology, dissimilation is classified as a psychological change, alongside analogy, haplology, and metathesis (Scheer, 2015). These abrupt changes largely affect individual words, explaining why dissimilation typically manifests as a tendency in the lexicon rather than as rules in phonology (Murray, 2013). Blevins (2013) also asserts that the origin of dissimilation lies in the intrinsic phonological ambiguity of the phonetic signal, leading to divergent percepts.

According to the co-articulation hypercorrection theory (Ohala, 1981, 1993, 1997, 2012), dissimilation occurs when the listener reverses a perceived co-articulation. Therefore, dissimilation is expected to occur only with features having acoustic cues significantly extended in time.¹ Alternatively, dissimilation may arise from challenging processing conditions (Frisch et al., 2004). In this perspective, the repetition of similar sounds is avoided due to the difficulties associated with processing the sequencing of similar segments.

Indeed, Coetzee (2005) provides evidence for the psychological reality of dissimilation.² He observes the absence of /spVp/ and /skVk/ sequences in the English lexicon (while /stVt/ is present). To investigate this phenomenon, three sets of continua were developed

for a perception experiment: [k]–[p], [k]–[t], and [p]–[t]. Each continuum consisted of eight tokens, which were incorporated into six words, as illustrated in Table 1.

Table 1. The stimuli in Coetzee (2005) and their predicted biases

| Words | Predicted bias | Comment |
|---------------|----------------|---------------------------------|
| [skap]–[skak] | [p] | Because *[skak] |
| [spap]–[spak] | [k] | Because *[spap] |
| [spap]–[spat] | [t] | Because *[spap] |
| [stap]–[stat] | [p] | Uncertain, because [stat] legal |
| [skək]–[sket] | [t] | Because *[skək] |
| [stək]–[stet] | [k] | Uncertain, because [stet] legal |

As hypothesized, the continua were generally perceived to differ from the preceding consonant. [p]–[k] stimuli were more likely to be recognized as [p] in [ska_] than in [spa_]. This finding could be attributed to the influence of phonological grammar learned based on the lexicon. But even though [stat] or [stet] is allowed in English, listeners tended to judge the [p]–[t] and [k]–[t] stimuli not as [t] when it is preceded by another [t]. Consequently, the results reinforce the psychological reality of dissimilation with a perceptual bias against repeated consonants.

While voice onset time (VOT) dissimilation could be identified in articulatory variation³, recent findings indicate it is not always the case that there is a phonetic precursor of dissimilation in the articulatory variation of repeated aspirated stops in Korean. Oh et al. (2020) reported that long VOTs of C₁ are shortened when nonadjacent C₂ has also long VOTs in Korean. For instance, C₁ VOT of aspirated stops was significantly shorter before another aspirated stop. However, they found that the VOT shortening effect is bidirectional in that not only C₂ long VOT shortened C₁ VOT, but also C₁ long VOT shortened C₂ VOT. In their successive study, it was replicated that *both* aspirated stops are realized with shortened VOT, not just *one* (Kang & Oh, 2024). These studies lead us to a perception study in search for the motivation of dissimilation.

Building on the previous arguments regarding the origin of dissimilation, we investigate the clues of VOT dissimilation in the perception of repeated aspirated stops in Korean. Similar to Coetzee's (2005) examination of the perceptual motivation of place-OCP in English, this study aims to illustrate how laryngeal dissimilation is driven, using aspirated stops in Korean. Additionally, while Coetzee's study explored the effect of the preceding consonant, we will focus on the impact of the following one. By demonstrating that the perception of an aspirated stop is affected by another aspirated stop in the following syllable, we will claim that VOT dissimilation in Korean is perceptually motivated.

2. Methods

The stimuli were obtained from data collected in Kang & Oh's (submitted) study, where 16 Korean speakers produced 46 different

¹ Aspiration (or long VOT) is regarded as one of temporally extended features along with pharyngealization and laryngealization (Ladefoged & Maddieson, 1996).

² OCP (Obligatory Contour Principle) is adopted instead of dissimilation in Coetzee (2005).

³ For example, it was found that an aspirated stop could be realized with shortened VOT when it is followed or preceded by another stop in Mongolian (Svantesson & Karlsoon, 2012; Svantesson et al., 2005), Georgian (Beguš, 2016), and Aberystwyth English (Jatteau & Hejná, 2016).

types of C₁ C₂ a nonce words.⁴ C₁ and C₂ were chosen from nine Korean oral stops (p, p^h, p'; t, t^h, t'; k, k^h, k') or /n/. The data from one female and one male speaker were selected, and 10 tokens were then extracted from each speaker, as detailed in Table 2. For C₁, one aspirated stop and one lenis stop for each of the three places of articulation were utilized (six in total), and for C₂, alveolar aspirated, lenis, tense, and nasal stops (four in total) were employed.

Table 2. Bases of stimuli (FV: following vowel, PV: preceding vowel, CD: closure duration)

| C ₁ : 6 bases | | C ₂ : 4 bases | |
|---|--------------------|-----------------------------|-------------------------------|
| C ₁ type | Duration (ms) | C ₂ type | Duration (ms) |
| Aspirated (p ^h , t ^h , k ^h) | VOT: 70 +FV: 40 | Aspirated (t ^h) | PV: 40+CD: 100+VOT: 60+FV: 90 |
| | | Lenis (t) | PV: 40+CD: 62+VOT: 20+FV: 90 |
| Lenis (p, t, k) | VOT: 70 +FV: 40 | Tense (t') | PV: 40+CD: 140+VOT: 20+FV: 90 |
| | | Nasal (n) | PV: 40+Nasal: 60+FV: 90 |

C₁ was extracted with 40 ms of the following vowel, and C₂ was extracted with 50 ms of the preceding vowel and so the total length of V₁ was 90 ms. V₂ was also 90 ms, taken out with C₂. The durations of VOT and closure were determined in consultation with the mean values of the entire dataset in Kang & Oh (submitted, see the table in footnote 11). The VOT of C₁ was manipulated into seven steps with 10 ms intervals, ranging from 70 ms to 10 ms.⁵ These steps were then concatenated with four different C₂s, resulting in a total of 168 tokens (6 C₁×7 steps×4 C₂) for each speaker.

A total of 32 Korean speakers (24 females and 8 males) who were college students at a university in Gwangju participated in the perception experiment. They were all born and raised in Gwangju and South Jeolla Province and no one reported a problem with hearing. They were all paid for participation. They were evenly divided into two groups, with 12 female and 4 male speakers in each group. One group listened to stimuli from the male speaker, while the other group listened to stimuli from the female speaker. The experiment consisted of two main blocks in which the tokens were given in random order. Before the main blocks, the participants went through a practice session involving non-manipulated tokens. It was structured as a forced-choice task, where participants were instructed to choose one of the three laryngeal types for C₁ (e.g., p^h, p, p') after listening to each nonce word.⁶ On average, the experiment lasted about 15 minutes.

For each token (168 tokens×32 participants=5,376 tokens), response and response time (RT) were recorded. A total of 41 tokens marked as 'no response' were discarded⁷, resulting in 5,335 valid responses and RTs for statistical analyses.

3. Results

C₁ type clearly determines the response results, as illustrated in Table 3. When the base of C₁ was lenis, it was identified as lenis 99% of the time, irrespective of VOT duration. Additionally, RT was significantly shorter when the base of C₁ was lenis ($p<.001$), indicating that the choice was easier for the participants. Consequently, the subsequent statistical analyses were conducted solely with the data where the base of C₁ was aspirated. From this dataset, the 30 responses indicating lenis were excluded for logistic regression analyses since the responses were not affected by VOT steps nor by C₂. Aspirated and tense responses were coded as 0 and 1, respectively.

Table 3. Basic results

| C ₁ type | Response | | | Total |
|---------------------|---------------|--------------|------------|-------|
| | Aspirated (%) | Lenis (%) | Tense (%) | |
| Aspirated | 2,273 (85.6) | 30 (1.1) | 352 (13.3) | 2,655 |
| Lenis | 13 (0.5) | 2,653 (99.0) | 14 (0.5) | 2,680 |

The data underwent logistic regression analyses using the mixed-effects model in R (R Core Team, 2021). The independent variables considered were C₁ type (aspirated and lenis), C₁ place (coronal, dorsal, and labial), C₁ VOT (ranging from 10 to 70 in seven steps), Voice (female and male), and C₂ type (aspirated, lenis, nasal, and tense).⁸ Participant was treated as a random effect. The gender of participants did not have a significant impact on the results and was therefore excluded from the analysis. The final results are presented in Table 4.⁹

⁴ They also produced 46 different kinds of C₁anC₂a words but none of these were selected for this study. The mean and standard deviation (in parenthesis) values of VOT and CD (closure duration) are as follows:

| C ₁ | C ₁ VOT | C ₂ CD | C ₂ VOT |
|----------------|--------------------|-------------------|--------------------|
| Aspirated | 70.6 (20.8) | 120.9 (35.3) | 59.7 (19.3) |
| Lenis | 58.4 (23.3) | 58.9 (25.8) | 18.4 (13.1) |
| Tense | 16.0 (7.1) | 143.0 (45.7) | 16.3 (7.1) |

⁵ 10 ms of VOT was taken out from the middle of VOT one step at a time.

⁶ The choices were given in the Korean writing system (Hangeul). Although all plausible options (i.e., all the three distinctive laryngeal stops in Korean: aspirated, lenis, and tense) were given, it should be noted that the forced-choice task might have introduced noise to the results.

⁷ These are cases where participants did not respond in three seconds.

⁸ For each of the variable, the first item was given as reference.

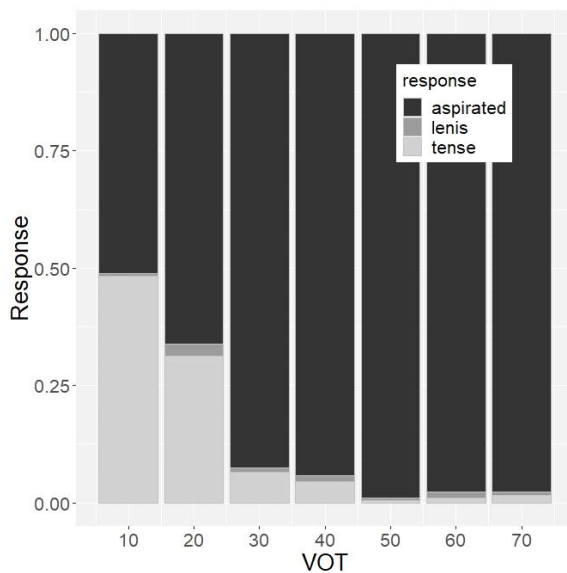
⁹ `glmer(response~VOT+C2+C1_place+voice+VOT*C1_place+(1|participant), family=binomial, data=d)`. The interaction of VOT and C₂ was excluded because it turned out not to be significant at all.

Table 4. The results of logistic regression (response)

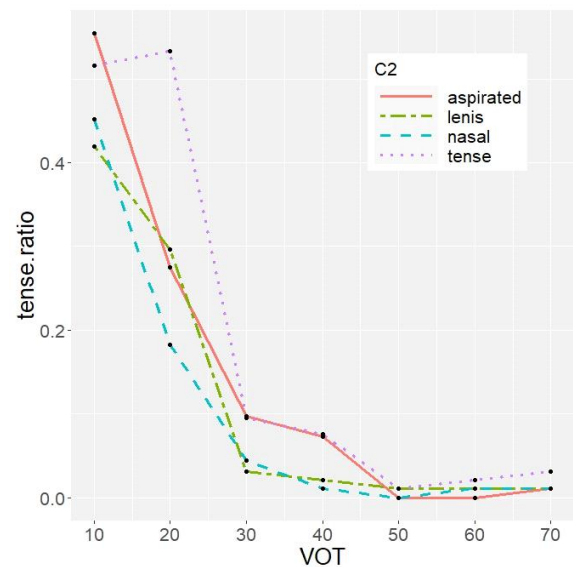
| | Estimate | Std. error | z-value | Pr(> z) |
|--------------------------------|----------|------------|---------|-------------|
| (Intercept) | 1.398 | 0.381 | 3.67 | 0.000243*** |
| VOT | -0.147 | 0.013 | -10.86 | <2e-16*** |
| C ₂ : lenis | -0.396 | 0.205 | -1.93 | 0.053457 |
| C ₂ : nasal | -0.586 | 0.209 | -2.79 | 0.005209** |
| C ₂ : tense | 0.456 | 0.192 | 2.37 | 0.017442* |
| C ₁ _place: dor | -0.568 | 0.369 | -1.53 | 0.124131 |
| C ₁ _place: lab | -1.390 | 0.379 | -3.66 | 0.000246*** |
| Voice: male | 0.964 | 0.336 | 2.86 | 0.004186** |
| VOT*C ₁ _place: dor | 0.043 | 0.016 | 2.70 | 0.006904** |
| VOT*C ₁ _place: lab | 0.050 | 0.016 | 3.03 | 0.002370** |

* $p < .05$, ** $p < .01$, *** $p < .001$.

As anticipated, C₁ was more likely perceived as tense as the VOT of C₁ became shorter. However, even when VOT was at its shortest (10 ms), tense responses remained below 50%. In contrast, when VOT was 50 ms or longer, the majority of tokens were perceived as aspirated (refer to Figure 1).

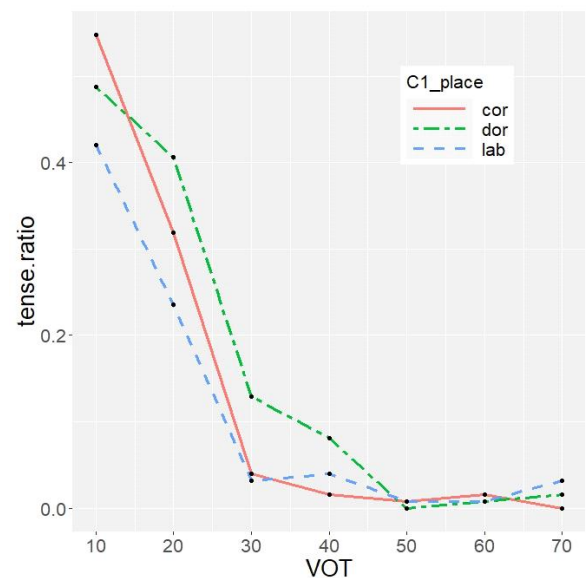
**Figure 1.** Response ratio as a function of VOT.

The judgment was also influenced by C₂. The effect of C₂ on tense responses is shown in Figure 2. Generally, a tense response was more probable when C₂ was aspirated compared to when it was nasal. However, there was no significant difference in tense responses between aspirated C₂ and lenis C₂ conditions.

**Figure 2.** Ratio of tense response as a function of C₂ type.

The ratio of tense responses was higher at 10 ms, 30 ms, and 40 ms VOT when C₂ was aspirated or tense compared to when it was lenis or nasal. At 20 ms, the ratio was predominantly higher when C₂ was tense. Interestingly enough, a tense C₂ significantly elevated the ratio of tense response compared to an aspirated C₂. To be specific, when C₁ VOT is 20 ms, the ratio of tense response is very high when nonlocally followed by a tense C₂ but it is drastically dropped when nonlocally followed by an aspirated C₂.

The effect of C₁ place on tense responses was also found as shown in Figure 3. The ratio was significantly lower when C₁ was labial compared to other places of articulation. The difference between coronal and dorsal C₁ with respect to the ratio of tense responses was not significantly different.

**Figure 3.** Ratio of tense response as a function of C₁ place.

Next, we will examine the analysis of RT. RT was subjected to regression analyses with the same factors. Table 5 demonstrates the

results of regression analysis regarding RT.

Table 5. The results of regression (response time)

| | Estimate | Std. error | z-value | Pr(> z) |
|-------------------------|----------|------------|---------|-------------|
| VOT | -4.126 | 0.389 | -10.58 | <2e-16*** |
| C ₂ lenis | 42.807 | 22.052 | 1.94 | 0.052348 |
| C ₂ nasal | 2.529 | 22.095 | 0.11 | 0.908849 |
| C ₂ tense | 83.839 | 22.027 | 3.80 | 0.000144*** |
| C ₁ placedor | 48.173 | 19.116 | 2.52 | 0.011798* |
| C ₁ placelab | 58.354 | 19.084 | 3.05 | 0.002254** |
| Voicemale | 179.556 | 53.716 | 3.34 | 0.002238** |

* $p < .05$, ** $p < .01$, *** $p < .001$.

Figure 4 shows the results of RT as a function of C₂ type. The shorter the VOT, the longer it took to make a judgment. Tense C₂ required a significantly longer time for the choice. Coronal C₁ significantly shortened RT. The male voice required more time to make a judgment than the female voice.

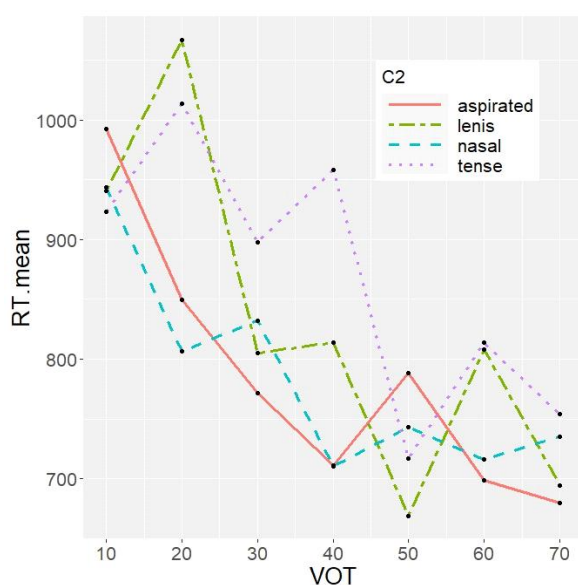


Figure 4. Response time as a function of C₂ type.

To summarize the results from the perception experiment, lenis-based C₁ was mostly perceived as lenis but aspirated-based C₁ was judged as aspirated or tense. Short VOT raised the ratio of tense response, which was also elevated by tense C₂, followed by aspirated/lenis C₂ and nasal. For these effects, RT was long proportionally to the ratio of tense response. This suggests that when a participant made a decision toward tense, the token was perceptually ambiguous, in general.

4. Discussion and Conclusion

This study aimed to investigate the phonetic motivation behind VOT dissimilation in Korean, specifically examining whether and how an aspirated stop influences the perception of another aspirated stop in the preceding syllable. The findings revealed that Korean listeners exhibit a perceptual bias against consecutive aspirated-onsets. An aspirated stop in the second syllable decreased the likelihood of perceiving another stop in the first syllable as

aspirated. When combined with the results of the previous study (Coetzee, 2005), it becomes evident that dissimilation is more attributable to perception than production. While Kang & Oh (submitted) do not clearly show the origin of dissimilation in the variation patterns of production data, this study decisively demonstrates that dissimilatory patterns can emerge in the process of perception. However, it is noteworthy that C₁ tokens constructed from an aspirated stop were perceived as aspirated more than 50 percent of the time, even when VOT was at its shortest (10 ms). This is ascribed to the 40 ms of vowel taken from aspirated stops with VOT, which contained other acoustic cues for aspirated stops. From this result, it is suggested that such a misperception would be least probable, if not possible, giving a hint for the unproductivity of dissimilation.

While aspirated-aspirated pairs were perceived toward dissimilatory patterns, aspirated-tense pairs leaned toward assimilatory ones in perception. Namely, aspirated C₂ retarded the recognition of aspirated C₁ whereas tense C₂ accelerated that of tense C₁. This result provides insight into why tense-tense pairs are preferred in Korean (Kang & Oh, 2016, 2019; Kim H., 2016). The preference of tense-tense pairs is supported by two facts. Firstly, tense-tense pairs are overrepresented in the Korean lexicon. Secondly, word-initial lenis stop more likely undergoes tensification when followed by a tense stop in the following syllable. This preference seems to be, partially, based on the perceptual bias found in the current study. The asymmetry of aspirated and tense C₂ effects is also found in the transitional patterns of responses. The ratio of tense response gradually decreased in the aspirated C₂ condition, but it sharply dropped between 20 ms and 30 ms in the tense C₂ condition (refer to Figure 2). In addition, a tense C₂ significantly delayed the judgment (refer to Figure 4). Taking these differences into account, we speculate that tense C₂ affects the perception of C₁ in a different way from aspirated C₂. When the listeners are exposed to tense C₂, the judgment can be influenced not only during the online processing of speech sounds, but when retrieving the first consonant from their short-term memory. This hypothesis will be explored in future study.

This study affirmed the critical role of F₀ in distinguishing lenis stops from others, consistent with previous research (Ahn, 2000; Kang 2014; Kang & Guion, 2008; Kim & Duanmu, 2004; Lee & Jongman, 2012; Lee et al., 2020; Silva, 2006a, 2006b, among others). When listening to tokens constructed from lenis stops, 99% of responses were lenis, regardless of the VOT duration. Conversely, with tokens from aspirated stops, 98.9% of responses were either aspirated or tense. It is inferred that the 40 ms vowel portion taken from the base provided sufficient acoustic information, particularly in terms of F₀, for the distinction. In addition, lenis stops in Korean are permitted to have a broad range of VOT depending on their positions. They are manifested by long VOT in word-initial position but by short VOT in intervocalic position due to intersonorant voicing. Thus, Korean speakers tend to have a more lenient attitude toward the VOT values of lenis stops. The results clearly demonstrate that the low F₀ of the vowel onset is firmly associated with perceiving lenis stops in Korean.

Finally, the effect of place on stop category perception can be attributed to phonetic differences among stops. VOT in aspirated stops is in general long in the order of dorsal > coronal > labial (Kim, 2019, 2021; Lee & Yoon, 2016; Silva, 2006a, etc.).¹⁰ Thus, when the VOT values are 30 ms to 40 ms in Figure 3, dorsal C₁ was

more perceived as tense compared to labial or coronal C₁s since VOT values of dorsal stops need to be long enough to be perceived as aspirated.¹¹ In contrast, labial stops were least likely to be perceived as tense, presumably because labial aspirated stops are characterized by relatively short VOT. Considering that the duration of VOT is not significantly different between labial and coronal stops in many studies, we need to investigate whether identical place of articulation between C₁ and C₂ can affect the C₁ stop category perception. Since the C₂s in this study were all coronal, the ratio of tense responses might have increased in the coronal C₁ condition because of *identity* effect. As aspirated C₂ hindered the recognition of aspirated C₁, which is *identical* to C₂, the identity of place could have been another reason for the difference between coronal and labial. This aspect should be tested in the future, along with the effect of short-term memory (Goldinger et al., 1999).

References

- Ahn, H. (2000). On the lenis stop consonants in Korean. *Language Research*, 36(2), 361-379.
- Beddor, P. S., Harnsberger, J. D., & Lindemann, S. (2002). Language-specific patterns of vowel-to-vowel coarticulation: Acoustic structures and their perceptual correlates. *Journal of Phonetics*, 30(4), 591-627.
- Beguš, G. (2016, September). The phonetics of aspirate dissimilation: Grassmann's law in Georgian. *Poster presented at the South Caucasian Chalk Circle*. Paris, France.
- Blevins, J. (2013). Evolutionary phonology: A holistic approach to sound change typology. In P. Honeybone, & J. Salmons (Eds.), *The Oxford handbook of historical phonology* (pp. 485-500). Oxford, UK: Oxford University Press.
- Bye, P. (2011). Dissimilation: Volume III. Phonological processes. In M. van Oostendorp, C. J. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell companion to phonology* (pp. 1408-1433). Hoboken, NJ: John Wiley & Sons.
- Cho, T. (2004). Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics*, 32(2), 141-176.
- Coetzee, A. W. (2005). The obligatory contour principle in the perception of English. In S. Frota, M. Vigário, & M. J. Freitas (Eds.), *Prosodies: With special reference to Iberian languages* (pp. 223-246). Berlin: De Gruyter Mouton.
- Cser, A. (2013). Basic types of phonological change. In P. Honeybone, & J. Salmons (Eds.), *The Oxford handbook of historical phonology* (pp. 193-204). Oxford, UK: Oxford University Press.
- Davis, S. (1991). Coronals and the phonotactics of nonadjacent consonants in English. In C. Paradis, & J. F. Prunet (Eds.), *The special status of coronals: Internal and external evidence: Phonetics and phonology* (Vol. 2, pp. 49-60). San Diego, CA: Academic Press.
- Frisch, S. A., Pierrehumbert, J. B., & Broe, M. B. (2004). Similarity avoidance and the OCP. *Natural Language and Linguistic Theory*, 22(1), 179-228.
- Gafos, A., & Lombardi, L. (1999). Consonant transparency and vowel echo. *North East Linguistics Society*, 29(2), 8.
- Gallagher, G. (2010). Perceptual distinctness and long-distance laryngeal restrictions. *Phonology*, 27(3), 435-480.
- Gay, T. (1974). A cinefluorographic study of vowel production. *Journal of Phonetics*, 2(4), 255-266.
- Gay, T. (1977). Articulatory movements in VCV sequences. *The Journal of the Acoustical Society of America*, 62(1), 183-193.
- Goldinger, S. D., Kleider, H. M., & Shelley, E. (1999). The marriage of perception and memory: Creating two-way illusions with words and voices. *Memory & Cognition*, 27(2), 328-338.
- Ito, C. (2014). Compound tensification and laryngeal co-occurrence restrictions in Yanbian Korean. *Phonology*, 31(3), 349-398.
- Jatteau, A., & Hejné, M. (2016). Dissimilation can be gradient: Evidence from Aberystwyth English. *Papers in Historical Phonology*, 1, 359-386.
- Kang, H., & Oh, M. (2016). Dynamic and static aspects of laryngeal co-occurrence restrictions in Korean. *Studies in Phonetics, Phonology and Morphology*, 22(1), 3-34.
- Kang, H., & Oh, M. (2019). The asymmetric tense consonant effects in compound and word-initial tensions in Korean. *Studies in Phonetics, Phonology and Morphology*, 25(1), 3-30.
- Kang, H., & Oh, M. (2024). A phonetic study of bidirectional duration modulation in Korean oral stops. *Korean Journal of Linguistics*, 49(1), 31-50.
- Kang, K. H., & Guion, S. G. (2008). Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. *The Journal of the Acoustical Society of America*, 124(6), 3909-3917.
- Kang, Y. (2014). Voice onset time merger and development of tonal contrast in Seoul Korean stops: A corpus study. *Journal of Phonetics*, 45, 76-90.
- Kim, H. (2016). Contextual distribution of English loanword word-initial tensification in Korean. *Studies in Phonetics, Phonology and Morphology*, 22(2), 245-288.
- Kim, M. R. (2019). A study of L1 phonetic drift in the voice onset times of Korean learners of English with long L2 exposure. *Phonetics and Speech Sciences*, 11(4), 35-43.
- Kim, M. R. (2021). Voice onset time in English and Korean stops with respect to a sound change. *Phonetics and Speech Sciences*, 13(2), 9-17.
- Kim, M. R., & Duanmu, S. (2004). "Tense" and "lax" stops in Korean. *Journal of East Asian Linguistics*, 13(1), 59-104.
- Kim, S. (2016). *Phonological trends in Seoul Korean compound tensification* (Master's thesis). Seoul National University, Seoul, Korea.

¹⁰ In some studies, labial and coronal stops are not different in terms of VOT duration.

¹¹ One of the reviewers pointed out that some tokens are possibly heard as real words, which could have affected the results. Considering that some tokens ended with *ta*, *t'a*, and *na* (same as verbal ending forms) it is plausible, even though the participants were told they would listen to nonce words. While both aspirated and tense forms of coronal (*t'a* and *t'a*) and labial (*p'a* and *p'a*) are frequently used as the initial part of verbal stems in Korean, aspirated form of dorsal (*k'a*) would not be as frequent as tense form (*k'a*) in Korean verbal stems. This might be an additional reason for the high ratio of tense response in dorsal. However, RT was not affected by this.

- Kiparsky, P. (2003). The phonological basis of sound change. In B. D. Joseph, & R. D. Janda (Eds.), *The handbook of historical linguistics* (pp. 311-342). Hoboken, NJ: Blackwell.
- Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford, UK: Blackwell.
- Lee, H., Holliday, J. J., & Kong, E. J. (2020). Diachronic change and synchronic variation in the Korean stop laryngeal contrast. *Language and Linguistics Compass*, 14(7), e12374.
- Lee, H., & Jongman, A. (2012). Effects of tone on the three-way laryngeal distinction in Korean: An acoustic and aerodynamic comparison of the Seoul and South Kyungsang dialects. *Journal of the International Phonetic Association*, 42(2), 145-169.
- Lee, Y., & Yoon, K. (2016). A study on the voice onset times of the Seoul Corpus males in their twenties. *Phonetics and Speech Sciences*, 8(4), 1-8.
- MacEachern, M. R. (1999). *Laryngeal cooccurrence restrictions*. New York, NY: Garland.
- Manuel, S. Y. (1990). The role of contrast in limiting vowel to vowel coarticulation in different languages. *The Journal of the Acoustical Society of America*, 88(3), 1286-1298.
- Manuel, S. Y., & Krakow, R. A. (1984). Universal and language particular aspects of vowel-to-vowel coarticulation. *Haskins laboratories status report on speech research* (No. 77/78). Retrieved from <https://files.eric.ed.gov/fulltext/ED247626.pdf#page=73>
- McCarthy, J. J. (1986). OCP effects: Gemination and antigemination. *Linguistic Inquiry*, 17(2), 207-263.
- Murray, R. W. (2013). The early history of historical phonology. In P. Honeybone, & J. Salmons (Eds.), *The Oxford handbook of historical phonology* (pp. 11-31). Oxford, UK: Oxford University Press.
- Oh, M., Kim, D., & Kang, H. (2020). Duration modulation in Korean stops: Nonlocal similarity avoidance vs. timing regulation. *Studies in Phonetics, Phonology and Morphology*, 26(1), 103-125.
- Ohala, J. J. (1981). The listener as source of sound change. In C. S. Masek, R. A. Hendrick, & M. F. Miller (Eds.), *Papers from the parasession on language and behavior* (pp. 178-203). Chicago, IL: Chicago Linguistic Society.
- Ohala, J. J. (1993). The phonetics of sound change. In C. Jones (Ed.), *Historical linguistics: Problems and perspectives* (pp. 237-278). London, UK: Longman.
- Ohala, J. J. (1997). The relation between phonetics and phonology. In W. J. Hardcastle, & J. Laver (Eds.), *The handbook of phonetic sciences* (pp. 674-694). Oxford, UK: Blackwell.
- Ohala, J. J. (2012). The listener as a source of sound change: An update. In M. J. Solé, & D. Recasens (Eds.), *The initiation of sound change: Perception, production, and social factors* (pp. 21-26). Amsterdam, The Netherlands: John Benjamins.
- Öhman, S. E. G. (1966). Coarticulation in VCV utterances: Spectrographic measurements. *The Journal of the Acoustical Society of America*, 39(1), 151-168.
- R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria. Retrieved from <https://www.R-project.org/>
- Scheer, T. (2013). How diachronic is synchronic grammar? Crazy rules, regularity, and naturalness. In P. Honeybone, & J. Salmons (Eds.), *The Oxford handbook of historical phonology* (pp. 313-336). Oxford, UK: Oxford University Press.
- Silva, D. J. (2006a). Variation in voice onset time for Korean stops: A case for recent sound change. *Korean Linguistics*, 13(1), 1-16.
- Silva, D. J. (2006b). Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology*, 23(2), 287-308.
- Suzuki, K. (1988). *A typological investigation of dissimilation* (Doctoral dissertation). University of Arizona, Tucson, AZ.
- Svantesson, J. O., & Karlsson, A. (2012). Preaspiration in modern and old Mongolian. *Suomalais-Ugrilainen Seuran Toimituksia*, 264, 453-464.
- Svantesson, J. O., Tsendina, A., Karlsson, A., & Franzen, V. (2005). *The phonology of Mongolian*. Oxford, UK: Oxford University Press.
- Walker, R., Byrd, D., & Mpiranya, F. (2008). An articulatory view of Kinyarwanda coronal harmony. *Phonology*, 25(3), 499-535.
- Walter, M. A. (2007). *Repetition avoidance in human language* (Doctoral dissertation). Massachusetts Institute of Technology, Cambridge, MA.
- Yip, M. (1989). Feature geometry and cooccurrence restrictions. *Phonology*, 6(2), 349-374.
- Whalen, D. H., Shaw, P., Noiray, A., & Antony, R. (2011, Aug). Analogs of Tahlitan consonant harmony in English CVC syllables. *Proceedings of the 17th International Congress of the Phonetic Sciences* (pp. 2129-2132). Hong Kong, China.

• Hijo Kang

Associate Professor, Department of English Language Education
Chosun University, 309 Pilmun-daero, Dong-gu,
Gwangju 61452, Korea
Tel: +82-62-230-7323
Email: hijokang@chosun.ac.kr
Areas of interest: Phonetics, Phonology

• Mira Oh, Corresponding author

Professor, Department of English Language and Literature
Chonnam National University, 77 Yongbong-ro, Buk-gu
Gwangju 61186, Korea
Tel: +82-62-530-3165
Email: mroh@chonnam.ac.kr
Areas of interest: Phonetics, Phonology