PHONETIC REDUCTION AND ENHANCEMENT IN CONVERSATIONAL KOREAN STOPS

Jae-Hyun Sung

Yonsei University
jsung@yonsei.ac.kr

ABSTRACT

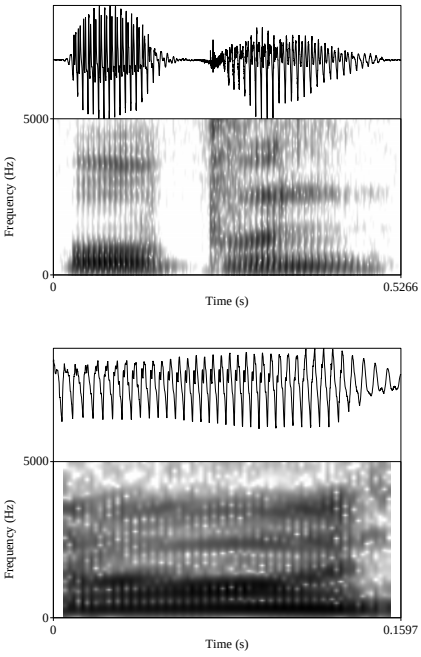
Conversational speech has received considerable attention in recent years, but many instrumental studies have focused on heavily resourced languages and the phonetic reduction in conversation. To diversify and expand on conversational speech research, this study examines the unique system of stop contrast in Korean and their acoustic cues in spontaneous phone conversations. The study investigates some key acoustic properties of intervocalic Korean stops in casual speech, adding to the growing body of research on conversational speech. Korean stops from careful reading and casual conversation show greater acoustic variability in conversation, as previously noted. Despite the phonetic variability in acoustic signals, conversational Korean stops contrast with one another in three acoustic dimensions: intensity difference, VOT, and closure duration, allowing some cues to enhance the stop contrast in conversation. The findings point to manner-specific adjustments in spontaneous Korean stops, and suggest that variability in conversational speech yields both phonetic reduction and enhancement.

Keywords: speech styles, conversation, spontaneous speech, read speech, Korean stops

1. INTRODUCTION

Speech sounds in casual conversation differ dramatically from those found in citation forms, owing to various phonetic reduction processes, differing prosodic constraints, and so forth. For example, a common English word *yesterday* /jɛstə-deɪ/ is phonetically realized as [jɛʃɛ̝ˀ] in casual conversation by a young American adult [23]. This conversational 'reduction' [21] is prevalent across languages and varieties [4, 13], and it has been acoustically represented by shorter duration or deletion of segments or syllables, shorter distance between sound categories, and less available acoustic cues in speech signals. Conversational speech in Korean is not an exception to this. Figure 1 illustrates two instances of a common Korean word /opun/, '5 minutes', produced by one of the speakers in this study. The conversational token in Figure 1 exhibits a clear sign of phonetic reduction in conversational speech, in which the phonetic representation of the word considerably deviates from its canonical/phonological representation. Some of the segments and syllables in the phonological representation are "smeared" or deleted completely. As a result, the underlying two-syllable word is phonetically realized as one, [ə̃n], which is not even remotely close to the canonical form.

Figure 1: Read (upper) and conversational (lower) tokens of /opun/ '5 minutes', Speaker 8.



Considerable bodies of work have explored conversational speech, as reported in [13]. Nevertheless, instrumental acoustic studies are still few for casual/spontaneous speech in non-European languages. Recent phonetic studies, however, have started to examine the production of spontaneous speech in less resourced languages such as Japanese [1, 2] and Chinese [6, 9, 10, 11, 12, 8]. They reported that casual speech in these languages does exhibit substantial phonetic variability just like in English and show similar cue reduction processes. In order

to further extend the scholarship on less spoken languages, this paper examines acoustic cues in use in Korean conversation, focusing on word-medial, intervocalic Korean stops.

Table 1: Word-initial and -medial stop contrast in Korean

manner	word-initial		word-medial	
lenis	/**p**ul/	'fire'	/kua**p**a/	'guava'
fortis	/**p***ul/	'horn'	/a**p***a/	'father'
aspirated	/**p^{h}**ul/	'grass'	/a**p^{h}**a/	'It hurts'

Korean stops are known for its three-way stop contrast, as illustrated in Table 1, in which VOT and f0 of the immediately following vowel serve as important cues to distinguish three stop series [15, 17, 18, 16, 3]. This multi-dimensional stop contrast is not very common cross-linguistically, and thus raises an interesting question how the dynamic nature of stop contrast in Korean is phonetically realized in a casual speech setting, and it is worth investigating how those cues function in casual speech in comparison to careful speech. While domain-initial (e.g., word-initial, phrase-initial, etc.) stops have been extensively studied in the previous literature, a study of word-medial, intervocalic stops can offer new insights into Korean stop contrast, and will provide complementary documentation of the Korean language using an innovative method of data collection in laboratory phonology. The overall goal is to document the acoustic cues at play in conversational intervocalic Korean stops, and add to the growing literature on conversational speech and cue interactions in read vs. spontaneous speech.

2. METHODS

2.1. Participants

Ten native speakers of Korean (three males and seven females, aged 18–24, SD=1.7) were recruited and recorded at the Alberta Phonetics Laboratory in the University of Alberta. Two of the speakers moved to Canada at the age of nine and eleven respectively, but reported that they speak Korean at home and are involved in Korean speaking communities on a regular basis. Further, the two speakers did not behave differently from other eight speakers in terms of their productions of intervocalic stops in two speech registers, and thus were included and analyzed for this study. All other speakers are late learners of English and have not lived in English-speaking countries until their late teens or early twenties.

2.2. Stimuli

This study examined nine Korean stop phonemes /p, p*, p^{h}, t, t*, t^{h}, k, k*, k^{h}/ in various intervocalic (i.e., VCV) environments, some of which are illustrated in Table 2. Roughly 100 to 300 instances of intervocalic stops were extracted from each speaker's phone conversation, and 81 words that contain intervocalic stop phonemes and 19 filler words were chosen for the wordlist. Common loanwords such as /k^{h}ʌp^{h}i/ 'coffee', as well as native Korean words, were also considered for this study. Stuttered words were excluded from the analysis.

Table 2: Conversational and read stops.

manner	segment	example	
lenis	/p/	/o**p**un/ '5 minutes'	/ki**p**un/ 'feeling'
fortis	/p*/	/o**p***a/ 'brother'	/a**p***a/ 'father'
aspirated	/p^{h}/	/ʌʧ^{h}a**p^{h}**i/ 'anyway'	/k^{h}ʌ**p^{h}**i/ 'coffee'

2.3. Procedure

Each speaker was instructed to sit in a sound-attenuated booth, phone their family member or close friend, who is also a native speaker of Korean, using their phone and carry out a brief conversation with them in Korean for fifteen minutes. Only one speaker held a conversation with the investigator as she could not find an interlocutor who was available at the time of her experiment. Only the speaker in the sound booth was consented and recorded (44.1kHz/16-bit WAV) through the high-quality head-mounted microphone (E6 Countryman) which the speaker was wearing. Every conversation was free conversation; no topics were offered. Following the phone conversation, speakers read a prepared list of Korean words in which each word appears three times in a random order. All the instructions were given in Korean throughout the experiment session.

2.4. Measurements

Four acoustic cues of intervocalic stops were measured using Praat [7]: intensity drops from the amplitude peaks of neighboring vowels (henceforth intensity differences), VOT, f0 of the immediately following vowel, and closure duration. The first measure is not a cue to Korean stop contrast, but was chosen for this study as it has been identified as a sign of acoustic reduction of intervocalic stops, in a way that intensity differences are smaller in casual

speech than in reading or storytelling [22, 23]. Intensity difference was defined as the difference between the intensity peaks of neighboring vowels and the trough in the region of the target stop. The rest of the measures–VOT, f0, and closure duration–are known as primary and secondary acoustic cues in Korean stop contrast [17, 18, 20], and thus were considered for this study.

2.5. Analysis

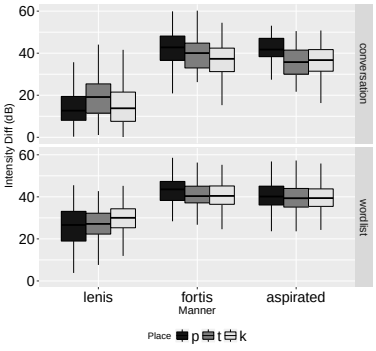
This study employed Linear Mixed Effects modeling [5] using R [19] to determine whether the acoustic signals of Korean stops are statistically significant across two speech registers. Results are reported in the following section for the fixed effects of speech style, manner of articulation, and the interactions of style with manner and with place, as well as the three-way interaction of style, manner and place.

3. RESULTS

3.1. Intensity differences

Figure 2 presents the intensity differences from all stop series produced in two speech styles. The effects of all factors–style, manner, and place–and all the possible interactions among them were statistically significant on the intensity difference measurement: χ^2=13.5, df=1, p<0.001 (style); χ^2=80.8, df=2, p<0.001 (style × manner); χ^2=39.9, df=2, p<0.001 (style × place); χ^2=32.4, df=8, p<0.001 (style × manner × place).

Figure 2: Intensity differences.

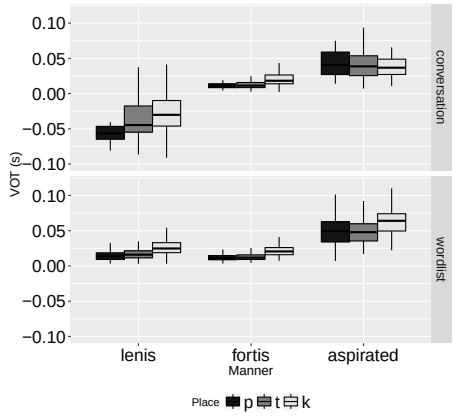


Lenis stops of all three places exhibit significantly smaller intensity differences in conversation (/p/: β=11.2, SE=2.2, t=5.2, p<0.001; /t/: β=8.6, SE=1.9, t=4.6, p<0.001; /k/: β=14.0, SE=1.5, t=9.5, p<0.001). Intervocalic lenis stops in Korean are susceptible to partial voicing and approximation, which decreases the amount of deviation in intensity from neighboring vowels. Style interacts with place, such that the effect of style is significantly larger for /k/ than for the other two stops. In the collected conversation recordings, intervocalic velar lenis stops occurred in frequently used function words such as /nɛ̲ka/ 'I + nominative marker' or /ki̲kʌ/ 'that'. Given that function words are more susceptible to phonetic reduction than content words [14], this might have resulted in a significantly greater effect of the velar place on intensity difference compared to other two places. Unlike lenis stops, fortis and aspirated stops are not affected by speech styles.

3.2. VOT

Figure 3 shows the VOT values from all stops in two speech styles. The effects of all factors and their interactions were statistically significant: χ^2=11.4, df=1, p<0.001 (style); χ^2=35.8, df=2, p<0.001 (style × manner); χ^2=31.2, df=2, p<0.001 (style × place); χ^2=42.4, df=8, p<0.001 (style × manner × place).

Figure 3: VOT.



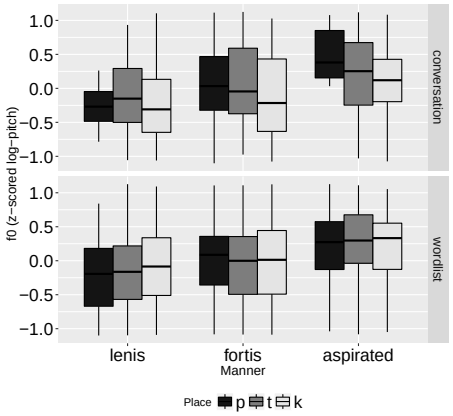
What is noticeable in Figure 3 is that VOT alone creates the three-way stop contrast in conversation, but not in wordlist reading. Intervocalic voicing of lenis stops in conversation, represented by negative VOT values, resulted in a more robust distinction among three stop series in conversation. For lenis stops, VOT for /t/ and /k/ is significantly smaller in conversation than in reading (/t/: β=0.004, SE=0.002, t=2.3, p<0.05; /k/: β=0.008, SE=0.002, t=3.9, p<0.01), demonstrating the conversational reduction in VOT. The effect of style is different across stop places, in which VOT of velar lenis stops is most affected by style. On the contrary, the VOT values for fortis and aspirated stops were not statistically different across two speech styles.

3.3. f0

Figure 4 illustrates the normalized (z-scored) f0 values from all stops in two speech styles, which shows

that f0 does not serve as a reliable cue in either read or conversational speech, at least word-medially. F0 tends to be higher in conversation than in reading for fortis and aspirated stops, but there were no significant effects of place or manner. Of interest here is that there was a significant interaction of style with place for aspirated stops (χ^2=15.7, df=3, p<0.01), in which /p^h/ and /t^h/, but not /k^h/, showed significantly higher f0 values in conversation than in reading.

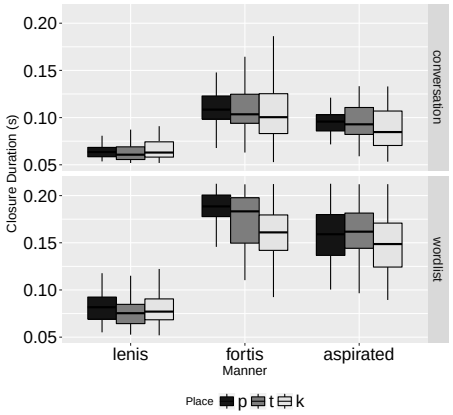
Figure 4: f0.



3.4. Closure duration

Figure 5 presents closure duration values from all stops in two speech styles. Except for the two-way interaction between style and place, all the factors and their interactions were found significant: χ^2=27.6, df=1, p<0.001 (style); χ^2=174.2, df=2, p<0.001 (style × manner); χ^2=0.9, df=2, p>0.1 (style × place); χ^2=68.8, df=8, p<0.001 (style × manner × place).

Figure 5: Closure duration.



Closure duration, known as a secondary cue in Korean stop contrast, successfully distinguishes three stop series in both speech styles, as illustrated in Figure 5. Further, closure duration is the only acoustic measure in this study that demonstrated the significant effect of style for all nine stop phonemes (/p/: β=0.02, SE=0.004, t=4.8, p<0.001; /t/: β=0.02, SE=0.004, t=5.7, p<0.001; /k/: β=0.03, SE=0.004, t=6.7, p<0.001; /p*/: β=0.1, SE=0.01, t=11.1, p<0.001; /t*/: β=0.1, SE=0.01, t=7.6, p<0.001; /k*/: β=0.9, SE=0.01, t=9.6, p<0.001; /p^h/: β=0.1, SE=0.01, t=6.6, p<0.001; /t^h/: β=0.1, SE=0.01, t=9.0, p<0.001; /k^h/: β=0.1, SE=0.01, t=6.3, p<0.001). All three stop series resulted in significantly shorter closure duration in conversation than in reading. Despite the significant decrease in closure duration for all manners of stops, three stop series are still contrastive in conversation.

4. DISCUSSIONS AND CONCLUSION

Taken together, the acoustic contrast of Korean stops is maintained in both careful reading and casual conversation despite acoustic reduction in some cues. Intensity differences come into play for both read and conversational stops, but it serves as a more robust cue in conversation in that the acoustic reduction of intervocalic lenis stop enhances the contrast between lenis and two other stop series. Two previously known cues in Korean stop contrast, VOT and closure duration, also contribute to the contrast in both speech styles, but VOT enhances the three-way contrast in conversation, whereas closure duration makes the contrast less robust. Finally, f0 does not seem to play a role in word-medial stop contrast.

The findings from this study suggest that conversational speech is not always about acoustic reduction, and the term "reduced speech", which is often used interchangeably with conversational speech, needs to be reconsidered. Conversational reduction is not necessarily associated with reduction in contrast as it may sound, and not every aspect of conversational speech exerts reduction. As shown in this study, represented by intensity difference in Figure 2 and VOT in Figure 3, some acoustic cues enhance phonetic contrast across sound categories. Thus, acoustic reduction may not be the only aspect of conversational speech research, and one should look into conversational speech in various perspectives.

On the whole, this paper adds weight to noteworthy differences between careful reading and casual speech, and offers further insight into conversational reduction and the dynamic aspect of Korean stop contrasts. Three stops series in Korean are contrastive in both speech styles, but in different acoustic dimensions. Furthermore, the findings from this paper highlight the importance of diversifying language data beyond the heavily resourced languages such as English or Dutch.

5. REFERENCES

[1] Arai, T. 1999. A case study of spontaneous speech in Japanese. *Proc. 14th ICPhS San Francisco*.

[2] Arai, T., Warner, N., Greenberg, S. 2007. Analysis of spontaneous Japanese in a multi-language telephone-speech corpus. *Acoustical Science and Technology* 28, 46–48.

[3] Bang, H.-Y., Sonderegger, M., Kang, Y., Clayards, M., Yoon, T.-J. 2018. The emergence, progress, and impact of sound change in progress in Seoul Korean: Implications for mechanisms of tonogenesis. *J. Phon.* 66, 120–144.

[4] Barry, W., Andreeva, B. 2001. Cross-language similarities and differences in spontaneous speech patterns. *J. Intl. Phon. Assc.* 31, 51–66.

[5] Bates, D., Mächler, M., Bolker, B., Walker, S. 2015. Fitting linear mixed-effects models using lme4. *J. Statistical Software* 65, 1–48.

[6] Berry, J. 2009. Tone space reduction in Mandarin Chinese. *157th Meeting of the Acoust. Soc. Am.*

[7] Boersma, P., Weenink, D. 2018. Praat: doing phonetics by computer [computer program], version 6.0.37.

[8] Brenner, D. 2015. *The Phonetics of Mandarin Tones in Conversation*. PhD thesis U. Arizona.

[9] Cheng, C. 2012. *Mechanism of Extreme Phonetic Reduction: Evidence from Taiwan Mandarin*. PhD thesis UCL.

[10] Cheng, C., Xu, Y. 2009. Extreme reductions: Contraction of disyllables into monosyllables in Taiwan Mandarin. *Interspeech 2009* 456–459.

[11] Cheng, C., Xu, Y. 2013. Articulatory limit and extreme segmental reduction in Taiwan Mandarin. *J. Acoust. Soc. Am.* 134(6), 4481–4495.

[12] Cheng, C., Xu, Y. 2014. Mechanism of disyllabic tonal reduction in Taiwan Mandarin. *Language and Speech* 58(3), 281–314.

[13] Ernestus, M., Warner, N. 2011. An introduction to reduced pronunciation variants. *J. Phon.* 39, 253–260.

[14] Johnson, K. 2004. Massive reduction in conversational American English. Yoneyama, K., Maekawa, K., (eds), *Spontaneous Speech: Data and Analysis (Proc. of the 1st Session of the 10th International Symposium)* 29–54.

[15] Kang, K.-H., Guion, S. G. 2008. Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. *J. Acoust. Soc. Am.* 124(6), 3909–3917.

[16] Kang, Y. 2014. Voice Onset Time merger and development of tonal contrast in Seoul Korean stops: A corpus study. *J. Phon.* 45, 76–90.

[17] 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. *J. Intl. Phon. Assc.* 42, 145–169.

[18] Lee, H., Politzer-Ahles, S., Jongman, A. 2013. Speakers of tonal and non-tonal Korean dialects use different cue weightings in the perception of the three-way laryngeal stop contrast. *J. Phon.* 41, 117–132.

[19] R Core Team, 2018. *R: A Language and Environment for Statistical Computing [computer program]*. R Foundation for Statistical Computing Vienna, Austria.

[20] Schertz, J., Cho, T., Lotto, A., Warner, N. 2015. Individual differences in phonetic cue use in production and perception of a non-native sound contrast. *J. Phon.* 52, 183–204.

[21] Warner, N. 2011. Reduction. In: van Oostendorp, M., Ewen, C. J., Hume, E., Rice, K., (eds), *The Blackwell Companion to Phonology*. Wiley-Blackwell 1866–1891.

[22] Warner, N., Tucker, B. V. 2010. "Probably, OK, whatever!": Variability in conversational speech stops and flaps. *159th Meeting of the Acoust. Soc. Am.*

[23] Warner, N., Tucker, B. V. 2011. Phonetic variability of stops and flaps in spontaneous and careful speech. *J. Acoust. Soc. Am.* 130, 1606–1617.