MORPHO-PHONOLOGICAL EFFECTS ON THE PHONETIC CHARACTERISTICS OF TENSE CONSONANTS IN KOREAN COMPOUNDS

Chloe Dokyung Kwon, Sam Tilsen, John Whitman

Cornell University
dk837@cornell.edu, tilsen@cornell.edu, jbw2@cornell.edu

ABSTRACT

This study examines the phonetic characteristics of Korean word-medial stops in two morphological contexts - compound and simple nouns - and probes the realization of compound tensification, a process whereby plain medial stops surface as tense consonants. Using utterances extracted from a large-scale Korean conversational corpus to see what speakers actually do, phone duration, F0, and spectral tilt were measured. These were used to investigate the differences in the phonetic patterns of plain and tense stops, described to be categorically distinct regardless of morphological context. However, the results show that morphological context does play a role. Specifically, the plain medial stops in compound nouns exhibit phonetic characteristics in the direction expected for tense stops, while the compound tense medial stops exhibit the characteristics in the direction expected for plain stops. The results call into question the assumption that compound tensification results in categorically distinctive tense consonants.

Keywords: phonetics, morphophonology, Korean, corpus study, stop consonants

1. INTRODUCTION

Korean obstruents show a typologically unusual three-way laryngeal contrast not including a voicing distinction: plain (lenis), tense (fortis), and aspirated [1]. Plain consonants are characterized by slight aspiration word-initially, tense consonants are characterized by tenseness in the vocal folds, and aspirated consonants are characterized by strong aspiration [2]. Acoustic correlates of these characteristics are well documented, including duration [3], [4], [5], [6], F0 [2], [4], [7], and spectral tilt [2], [8]. It is still uncertain how different morphological contexts affect the distribution of the three-way contrast and its implications for the morphological and phonological representations of words.

Korean compound tensification or *sai-sios* exemplifies the ambiguity in the relationship between the morpho-phonological status of words and their phonetic realizations. One of the common types of Korean nominal compounds is a subordinate compound [1]. It consists of two component words, of which the meaning of the first (W_1) modifies the meaning of the second (W_2), the head. Compound tensification refers to the process whereby a plain obstruent, onset of W_2 in a subordinate compound, becomes tense.

	W_1		W_2	
(1)	/pom	+	**p**i/	
	[pom		**p***i]	‘spring rain’
	spring		rain	
(2)	/pom	+	**p**at^h/	‘spring field’
	[pom		**b**at]	
	spring		field	

The historical account is that compound tensification is triggered by insertion of the reflex of a historical genitive marker, attested as a genitive particle =*s* in middle Korean [1]. This triggers post-obstruent tensification in the following plain obstruent [9], [10], causing the plain stop to surface as tense.

However, the occurrence of tensification is not fully predictable. For example, both (1) and (2) are subordinate compounds with the same component word preceding the bilabial stop in the target position of tensification. Korean dictionaries [11], [12] describe the target consonant in (1) as a tense stop (transcribed with an asterisk) and the one in (2) as a plain stop, which likely undergoes intersonorant voicing given the phonological context [7]. The way these forms are produced does not always agree with what is in the dictionary, as the occurrence of tensification tends to vary across speakers. For instance, [13] surveyed 35 native speakers of Yanbian Korean who were asked to indicate whether they would use a tense or plain pronunciation for 1,171 compound nouns. She found that the responses were variable by speaker, such that all 35 speakers agreed on only four of the words with respect to tensification.

To account for the variation, previous studies considered several possible factors. One of them is the frequency of the whole compound, which was found to be positively correlated with the occurrence of tensification [13], [14], [15]. Another factor is the segment type of the coda in the first component word. The likelihood of tensification increased in the order of vowel, liquid, nasal, and obstruent [13], [14]. The place of articulation of the target consonant also showed an effect, with coronal consonants experiencing tensification more frequently [14].

Although the above studies provide insight into the complex nature of compound tensification, there are two main unresolved issues that the present study aims to address. First, previous studies rely on linguistic judgments to infer phonological representations rather than directly investigating the realizations. In this study, this is done by systematic phonetic analyses of the expected acoustic correlates of the three-way contrast. Second, previous studies assume compound tensification is a categorical phenomenon in which the plain and derived tense consonants can be clearly distinguished, and the derived tense consonants are the same as the underlyingly contrastive tense consonants. To investigate whether this is true in terms of phonetic characteristics, this study examines word-medial plain and tense stops in Korean simplex and compound nouns. Acoustic measures are hypothesized to vary not only by the types of consonants (i.e., plain or tense) but also by their morphological contexts (i.e., compound or simplex).

Compound	C	V	{N, L, V}	#	__	V	(C)
Simplex	C	V	{N, L, V}	.	__	V	(C)

Table 1: Schematization of the morphological contexts. The shaded region shows the position of the target consonant.

In Table 1, the shaded region marks the position of the target consonant, which can be a plain or tense bilabial, alveolar, or velar stop. The preceding sound can be a nasal (N), liquid (L), or a non-high vowel (V), whereas the vowel following the target consonant were controlled to be non-high vowels. Based on prior acoustic studies, we expect tense consonants to be longer than plain consonants [3], [4], [5], [6]. F0 and spectral tilt measures on this vowel were predicted to be higher and lower, respectively, for tense consonants [2], [4], [7], [8]. The acoustic measures are hypothesized to vary depending on whether the sounds occur in a compound noun due to compound tensification, in which the plain consonants are expected to behave more similarly to tense consonants.

2. METHOD

2.1. Data

The speech data was extracted from the Korean Conversational Speech Corpus [16], which consists of conversations between two speakers on a common topic. Each conversation lasts for approximately 15 minutes, totaling about 500 hours of data. The corpus records data from 2,739 native speakers who speak different regional varieties and belong to various age groups. The data comes in a total of 870,682 audio files and 2,232 text transcription files. The transcriptions include speaker metadata such as age and the region of principal residence.

2.2. Data processing

The search for simplex and compound nouns was carried out using the Standard Korean Language Dictionary [11]. It provides information about the orthography, pronunciation after the phonological rules have been applied, and the position of the syllable and morpheme boundaries. From this dictionary, the nouns were extracted if (1) they are two to three syllables long, (2) they are native words, (3) they occur with a plain or a tense stop in the onset of the second or third syllable for simplex nouns or the onset of the second component word for compound nouns, (4) these stop sounds are followed by a non-high vowel (/a, ʌ, e/), and (5) the stops are not preceded by high vowels or obstruents. After compiling a list of the nouns, the phrases were parsed into words using the Kiwi parser [17] and kept only the words that appear at least once in the corpus. Regarding compound types, subordinate compounds were separated from coordinate compounds, which is another common type described not to exhibit compound tensification but were set aside for future analysis. The process resulted in 60 subordinate compounds (626 tokens) and 10 simplex nouns (1935 tokens).

The speech data was force-aligned with Montreal Forced Aligner [18]. After segmentation, the following measures were extracted and processed: the duration of the target consonant, fundamental frequency (F0) of the following vowel, and spectral tilt of the following vowel. The duration was measured using the segmentation information, and 219 (8.5%) outliers were removed. For F0 analysis, the values were extracted over the vowel following the target consonant using the autocorrelation method in Praat [19]. These were then processed to discard outliers and smoothed using a smoothing spline method with the spline parameter set to 9.9×10^{-4}. During this process, 247 tokens (9.6%) were discarded. The onset of F0 was used for the analysis.

For spectral tilt, the differences in the amplitudes of the first harmonics and second harmonics (H1-H2) were measured using Voicebox in MATLAB [20]. The processing involved down-sampling the audio signal to 11kHz for LPC analysis [21] and creating 20-ms frames with the step size of 1-ms. After applying the hamming window and estimating LPC filter coefficients, the inverse-filtered frames were used to estimate the narrowband spectrum. Fifty of the spectra were randomly selected and plotted individually to check whether the first and second harmonics fall within the reasonable range. 327 tokens (12.8%) were removed after this process. The analysis used the average of H1-H2 lying within the first 30% of the vowel interval.

2.3. Analysis

Linear mixed-effects modeling (LME) was implemented with *fitlme()* in MATLAB and was utilized to examine the phonetic realizations of plain and tense consonants in different morphological contexts. The dependent measures are the duration of the stop consonants, F0 of the following vowel, and the spectral tilt of the following vowel. The main predictors are the types of stop consonants (Plain or Tense) and the morphological contexts (Simplex or Compound). The interaction effect of these two predictors were included in the model to examine whether the effects of morphological contexts vary by the types of the stop consonants. Gender (Female or Male) was also considered for the F0 analysis. The model included individual subject and word items as random effects.

3. RESULT

3.1. Target consonant duration

Target consonant duration patterned consistently with the hypothesis that it would vary by the consonant types and morphological contexts. This is illustrated in Figure 1 which shows the differences in the distributions of the duration. The boxes are grouped by the types of consonants, with the blue box on the left representing those in simplex nouns and the orange box on the right representing those in compound nouns. The duration is longer for the tense consonants than the plain consonants, but the duration is shorter for tense consonants in compound nouns. This is opposite from the pattern of the duration of the plain consonants that become longer in compound nouns.

The effect of the consonant type is confirmed by the LME result (Plain: -45.8 $\pm$ 19.2, t(2338)=-2.4, $p<0.05$). There was no significant effect of the morphological context alone but an interaction effect was found between the types of consonants and the morphological contexts (Plain and Compound: 42.7$\pm$20.2, t(2338)=2.1, $p<0.05$). The duration of the plain consonants was longer in compound nouns while the duration of the tense consonants was shorter in compound nouns.

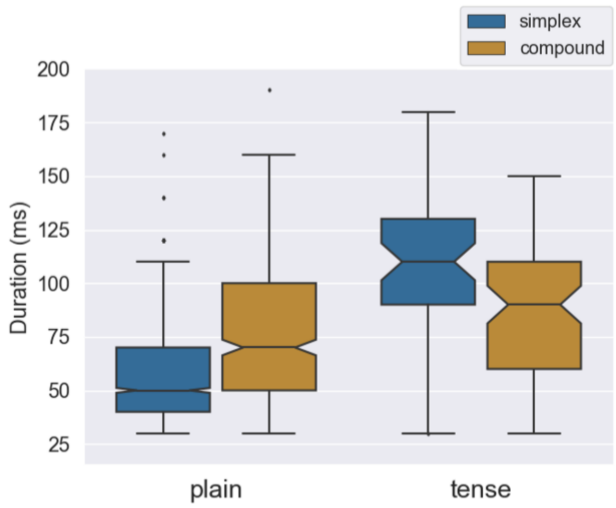

Figure 1: Distribution of phone duration of the target consonants

3.2. F0 of following vowel

Previous studies report that F0 values on the vowel following tense consonants are higher than those of the vowel following plain consonants [2], [4], [7]. Most studies focus on the consonants in the word-initial position, but the results from this study suggest that the pattern also arises in word-medial consonants in simplex nouns.

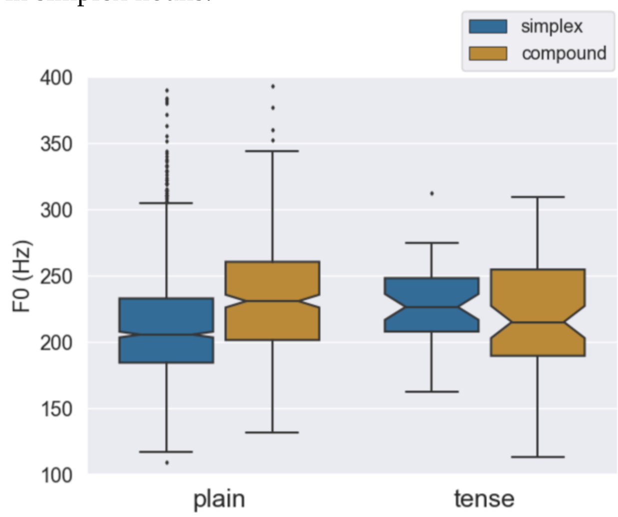

Figure 2: Distribution of F0 from female speakers

However, the plain consonants showed higher F0 than the tense consonants in compound nouns, as shown in Figure 2. Only the effect of gender was confirmed by the LME analysis.

3.3. H1-H2 of following vowel

H1-H2 was measured by comparing the difference between the amplitudes of the first and the second harmonics (H1-H2). It is measured to distinguish between breathy and creaky voicing followed by plain and tense consonants, respectively [2], [8]. The vowels following plain consonants are likely to show higher H1-H2 since the articulation of plain consonants tend to result in breathy vowels. On the other hand, the vowels following tense consonants are likely to be creaky, which would result in lower H1-H2.

Figure 3 shows that the plain and tense consonants exhibit similar H1-H2 values. What is striking is the difference within the tense consonants across the morphological conditions. Tense consonants in compound nouns show higher H1-H2 than the level of H1-H2 for plain consonants. The LME result shows that the differences between the types of consonants and the morphological contexts are significant (Plain: 1.6 ± 0.7, $t(2298)=2.3$, $p<0.05$; Compound: 2.1 ± 0.8, $t(2298)=2.6$, $p<0.01$), as well as their interaction effect (Plain and Compound: -1.9 ± 0.8, $t(2298)=-2.3$, $p<0.05$).

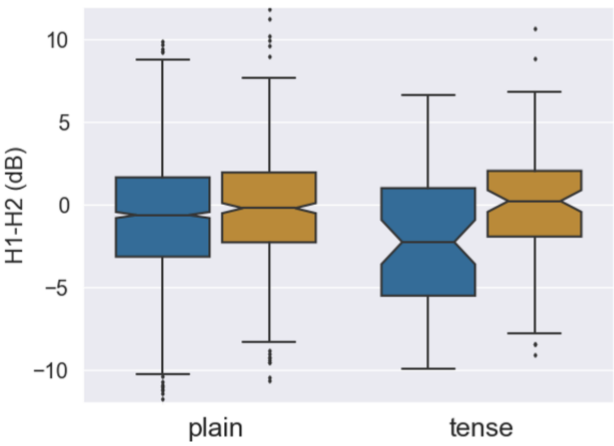

Figure 3: Distribution of averaged H1-H2 over the 30% of the vowel interval following the target consonants

4. DISCUSSION

The results show that phonetic characteristics of the consonants vary not only by consonant type but also by morphological context. Specifically, the duration of the plain consonants in a compound noun showed longer duration than those in a simplex noun. Also, although the differences were not significant, the direction of the effect on F0 was consistent with the hypothesis: it was higher for the plain consonants in a compound noun. It suggests that the plain consonants in compound nouns may be exhibiting characteristics of tense consonants as a result of compound tensification. However, it needs to be further investigated for more conclusive evidence.

All three phonetic measures showed that the tense consonants exhibit different phonetic characteristics depending on the morphological context. Interestingly, the measures patterned in the direction expected for the plain consonants. For example, the duration was longer, F0 was lower, and H1-H2 was higher for tense consonants in compound nouns compared to those in simplex nouns, which are characteristics expected for a plain obstruent in Korean. It seems that the tense consonants are almost neutralized to a plain consonant in a compound noun. It also indicates gradience in the phonetic details of a consonant type. The source of this gradience is a topic for further investigation.

This study offers an important insight into the nature of phonological processes through the examination of naturalistic data. While many studies explore phonetic implementation of phonological features or processes, not many have investigated the exact relation between the phonetic and phonological representations together with the role of the morphological factors. This is especially true in the context of compound tensification in Korean where the orthographic or metalinguistic representation have been taken for granted. The results support the importance of morphological contexts in predicting the phonetic patterns of consonants, which warrants further investigation in a more controlled production study. Another important dimension is to look at how these representations are processed and perceived by the listeners, which will add further insight into how the phonological knowledge is stored in the mind of the speakers/listeners.

While the corpus data set a solid ground for a preliminary analysis, it also came with variation from speakers and other contexts that were difficult to be controlled. For one, factors such as the category of the preceding sounds had to be excluded from the model since it correlated with one of the main predictors, the morphological context. Obtaining a more balanced set of observations from an experiment will allow systematic investigation of inter- and intra-speaker variation. It will also enable examining the difference between the frequency effects and lexicalization. A random-slope analysis will be performed on the more controlled dataset for reliable statistical results.

5. ACKNOWLEDGEMENTS

We thank Abigail C. Cohn for her detailed feedback in developing the ideas and improving this paper. We also thank all the members of Cornell Phonetics Lab for their time and input. Finally, we thank two anonymous reviewers for their thoughtful and helpful comments.

6. REFERENCES

[1] S. Cho and J. Whitman, *Korean: A linguistic introduction*. Cambridge University Press, 2020.

[2] T. Cho, S.-A. Jun, and P. Ladefoged, “Acoustic and aerodynamic correlates of Korean stops and fricatives,” *J. Phon.*, vol. 30, pp. 193-228, 2002.
[3] J.-I. Han, “The phonetics and phonology of “tense” and “plain” consonants in Korean,” Ph.D dissertation, Cornell Univ., Ithaca, NY, 1996.
[4] M.-R. C. Kim, “Acoustic characteristics of Korean stops and perception of English stop consonants,” Ph.D dissertation, Univ. of Wisconsin, Madison, 1994.
[5] D. J. Silva, “The phonetics and phonology of stop lenition in Korean,” Ph.D dissertation, Cornell Univ., Ithaca, NY, 1992.
[6] T. Cho and P. Keating, “Articulatory and acoustic studies of domain-initial strengthening in Korean,” *J. Phon.*, vol. 29, no. 2, pp. 155-190, 2001.
[7] S.-A. Jun, “The phonetics and phonology of Korean prosody,” Ph.D dissertation, Ohio State Univ., Columbus, 1993.
[8] M. R. Kim, P. S. Beddor, and J. Horrocks, “The contribution of consonantal and vocalic information to the perception of Korean initial stops,” *J. Phon.*, vol. 30, pp. 77-100, 2002, doi: 10.006/jpho.2001.0152.
[9] S.-C. Ahn, “The interplay of phonology and morphology in Korean,” Ph.D dissertation, Univ. of Illinois, Urbana, 1985.
[10] K. Chung, “Neutralization in Korean: a functional view,” PhD dissertation, Univ. of Texas at Austin, 1980.
[11] N. Park, “Stochastic learning and well-formedness judgement in Korean phonotactics,” Ph.D dissertation, Seoul National Univ., Korea, 2020.
[12] “Naver Korean Dictionary,” https://ko.dict.naver.com/#/main.
[13] C. Ito, “Compound tensification and laryngeal co-occurrence restictions in Yanbian Korean”, *Phonology*, vol. 31, no. 3, pp. 349-398, 2014, doi: 10.1017/S0952675714000190.
[14] S, Kim, “Phonologtical trends in Seoul Korean compound tensification, M.A. thesis, Seoul National Univ., Korea, 2016.
[15] K. Zurrow, “Predicting Korean sai-siot: phonological and non-phonological factors,” handout, *21st Japanese Korean Linguistics Conference*, Seoul National Univ., Korea, 2011.
[16] The National Institute of Korean Language, “Ilsang daehwa malmungchi [Converstional speech corpus],” https://corpus.korean.go.kr, 2020.
[17] M. Lee, “Kiwi, Korean Intelligent Word Identifier,” Version 0.14.0, 2022, doi: 10.5281/zenodo.7041425.
[18] M. McAuliffe, M. Socolof, E. Stengel-Eskin, S. Mihuc, M. Wagner, and M. Sonderegger, “Montreal Forced Aligner”. Version 1.0. 2017.
[19] P. Boersma, “Praat, a system for doing phonetics by computer,” *Glot Intl.*, vol. 5, pp. 341-345.
[20] “Voicebox: Speech processing toolbox for MATLAB,” http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html.
[21] J. D. Markel and A. H. Gray, *Linear prediction of speech.* New York: Springer-Verlag, 1976.