Published in final edited form as:

J Phon. 2013 March ; 41(2): 117–132. doi:10.1016/j.wocn.2012.12.002.

Speakers of tonal and non-tonal Korean dialects use different cue weightings in the perception of the three-way laryngeal stop contrast

Hyunjung Lee,
University of Kansas

Stephen Politzer-Ahles, and
University of Kansas

Allard Jongman
University of Kansas

Abstract

The current study investigated the perception of the three-way distinction among Korean voiceless stops in non-tonal Seoul and tonal Kyungsang Korean. The question addressed is whether listeners from these two dialects differ in the way they perceive the three stops. Forty-two Korean listeners (21 each from Seoul and South Kyungsang) were tested in a perception experiment with stimuli in which VOT and F0 were systematically manipulated. Analyses of the perceptual identification functions show that VOT and F0 cues trade off each other for the perception of the three stops. However, the trading relationship differs between the two dialects. Logistic regression analyses confirmed the two dialects use the perceptual cues differently for the lenis and aspirated stops. While Seoul listeners rely primarily on F0 for making lenis responses and on VOT and F0 for aspirated responses, F0 plays a less important role in modulating both lenis and aspirated responses for Kyungsang than for Seoul listeners. It is proposed that different tonal systems between the two dialects and the ongoing diachronic sound change in the stops of Seoul Korean contribute to the inter-dialect difference in cue weighting for the three-way stop distinction. The results suggest that although the difference in phonology between the two dialects influences the phonetic realization, the phonetic trade-off among multiple cues allows each dialect to maintain the phonemic distinction in a unique way.

Keywords

Korean; stops; perception; Seoul; Kyungsang; lexical pitch

1. Introduction

The purpose of this study was to investigate how a difference in tonal systems in the same language affects the perception of laryngeal contrast. We tested a well-known three-way

Address for correspondence: Hyunjung Lee, Linguistics Dept., Univ. of Kansas, 427 Blake Hall, 1541 Lilac Lane, Lawrence, KS 66044, Phone: 785-218-0720, hyunjung@ku.edu.

laryngeal distinction among Korean voiceless stops, comparing non-tonal Seoul and tonal Kyungsang dialects of Korean. A majority of previous studies have reported that both the VOT of stops and the fundamental frequency (F0) of following vowels are important, perceptually relevant acoustic properties to distinguish the three-way distinction among Korean voiceless stops, indicating that Korean speakers use not only consonant properties but also vowel properties to classify the three contrastive stops (e.g., Lisker & Abramson, 1964; C.-W. Kim, 1965; Cho, 1996; Cho, Jun & Ladefoged, 2002; Francis & Nusbaum, 2002; Kim, Beddor & Horrocks, 2002; Kim, 2004). Most of these studies have focused only on Seoul standard Korean, and consequently our understanding of the perception of the three-way laryngeal distinction of stops is limited to this variety.

Unlike standard Seoul Korean, which is non-tonal, North and South Kyungsang dialects of Korean (spoken in the southeastern part of the Korean peninsula) have lexical pitch accent contrasts (e.g., Ramsey, 1975; Kenstowicz, Cho & Kim, 2008). In a pitch accent language, pitch prominence within a phonological word and a different placement of high pitch may cue meaning differences. That is, the use of tone in the Kyungsang dialects is different from that in canonical tone languages (e.g., Chinese) in that tone patterns play a phonological role, but not absolute F0 differences for each syllable. An example in (1) shows three words of South Kyungsang Korean, which are segmentally homonymous, but contrast in the placement of high pitch within the word.

1.

Kyungsang	káci HL 'type'	kací LH 'eggplant'	kácí HH 'branch'
Seoul	kaci 'type', 'eggplant' or 'branch'		

Figure 1 presents an example of pitch tracks showing high pitch on the first and second syllable for HL and LH accent patterns, respectively, and comparable peaks on the two consecutive syllables for HH in South Kyungsang Korean.

Figure 1 indicates that the LH pattern is different from HL and HH patterns regarding F0 values of the initial syllables for the purpose of lexical accent contrasts. Kyungsang Korean also has the three-way laryngeal contrast in word- or phrase-initial position. Given that F0 in a following vowel is a cue for stop type in non-tonal Seoul Korean, the presence of lexical tone in Kyungsang raises an issue regarding the effects of tone on the three-way stop distinction. Since in Kyungsang Korean high and low F0 of the initial syllable is a cue to lexical accents (LH vs. HL, HH) and may compete with the use of F0 as a cue to word-initial voiceless stops, dialectal differences in F0 as a function of laryngeal distinction may well exist between non-tonal Seoul and tonal Kyungsang Korean.

This kind of linguistic inquiry has been pursued in other languages. Specifically, it is well known that F0 at vowel onset is correlated with the voicing of the preceding consonants in many languages such as English, French, Dutch and Japanese (e.g., Hombert, 1978; Abramson & Lisker, 1985; Löfqvist, Baer, McGarr & Seider Story, 1989; Oglesbee, 2008), in which F0 values are lower after voiced stops than voiceless stops. Physiological factors (i.e., aerodynamic factors, vocal fold tension) have been proposed as possible explanations, indicating that the effect of consonant voicing on F0 is intrinsic (Hombert, 1978). However, evidence from tonal languages suggests that speakers can control this effect to some degree. While lower F0 after voiced stops has also been found in tonal languages such as Thai, Yoruba, Taiwanese, and Cantonese, the F0 perturbation due to consonant voicing does not spread as far into vowels in tonal languages as it does in non-tonal languages (Gandour, 1974; Hombert, 1978; Lai, 2004; Francis, Ciocca, Wong & Chan, 2006). For this difference between non-tonal and tonal languages, Hombert (1978) stated that a tonal language speaker may tend to minimize the voicing effect on F0 to maximize the tonal contrast. This indicates that F0 difference is predicted not only by phonetic properties, but also by phonological properties. Moreover, Kingston and Diehl (1994) noted that F0 did not vary according to the voicing of a preceding consonant in Tamil, providing further evidence that the F0 perturbation is not determined solely by the phonetic difference in consonant voicing. Therefore, the evidence suggests that phonology does influence phonetic realization, supporting the notion that the phonetic realization is intentionally controlled or a combination of controlled and automatic processes (Kingston & Diehl, 1994; Francis et al., 2006).

As speakers of another tonal language, Kyungsang speakers also use pitch differences as a cue for two different purposes, laryngeal contrast and pitch accent contrast. Under the notion of controlled phonetics influenced by phonological structure, the different tonal systems between Seoul and Kyungsang Korean motivate us to test whether tonal Kyungsang speakers use F0 to differentiate laryngeal distinctions in the same way as non-tonal Seoul speakers do, and thus to test whether speakers' use of F0 to cue laryngeal contrasts is controlled in part by their phonology. If, as suggested by previous research, the effect of preceding consonants on F0 is not entirely an automatic process determined solely by physiology, the use of F0 for the three-way laryngeal distinction may be different between non-tonal Seoul and tonal Kyungsang Korean, and the competition in using F0 as a cue to accent and consonant in Kyungsang may diminish the use of F0 as a cue to the laryngeal distinction compared to Seoul.

Importantly, if the function of F0 is different between the two dialects of Korean, the function of other cues (i.e., VOT) that contribute to the three-way laryngeal distinction is also expected to show inter-dialect differences. According to the notion of trading relations in speech perception, when multiple cues work together to signal phonemic contrast, these cues trade with each other depending on their importance (Repp, 1982). For example, VOT and F1 trade with each other to cue voicing of initial stops in English; if onset F1 frequency is low, increased VOT might compensate the weakened F1 to signal a voiceless stop (Summerfield & Haggard, 1977). To cue voicing of intervocalic stops, closure duration and glottal pulsing during the closure also trade with each other: while long closure durations make the stop more likely to be perceived as voiceless, glottal pulsing can make stops with long closure durations still seem voiced (Lisker, 1986; Parker, Diehl & Kluender, 1986). That is, when one cue is not sufficiently salient to signal a phonetic distinction, the importance of the other cue is increased to obtain the phonetic equivalence. In this sense, if the role of F0 in stop categorization patterns differently between Seoul and Kyungsang Korean due to the different tonal systems, it is predicted that other perceptually relevant acoustic properties will also pattern differently.

Although there are a few acoustic studies on the three-way contrast of stops in Kyungsang Korean (e.g., Kenstowicz & Park, 2006; Lee & Jongman, 2012), the perception of stops in

Kyungsang Korean has not yet been explored. Therefore, the current study aimed to investigate the perceptual cues for the three-way distinction among Korean voiceless stops in both Seoul and Kyungsang Korean. By comparing the perceptual pattern for the Korean stops between the two dialects, we hope to better understand the relation among multiple perceptual cues and the way in which the presence vs. absence of tone in a speaker's phonology may affect the role of F0 as a cue to the laryngeal distinction. We start by reviewing previous research on the three-way distinction among voiceless stops in Seoul and Kyungsang Korean, and then specify the research hypothesis of the current perception study.

1.1. Previous investigations of the three-way laryngeal distinction in Seoul Korean

In word- or phrase-initial position, Korean has a three-way laryngeal distinction among voiceless stops and affricates (called fortis, lenis and aspirated). Each of the three stops occurs at three places of articulation (bilabial, alveolar, velar), and the affricates occur at palato-alveolar. Previous acoustic studies on stops have shown that both consonantal VOT and vocalic F0 cues play an important role in distinguishing the three-way contrast in Seoul Korean (e.g., Lisker & Abramson, 1964; C.-W. Kim, 1965; Han & Weitzman, 1970; Cho, 1996; Cho, Jun & Ladefoged, 2002). Specifically, although VOT increases from fortis to lenis to aspirated, the overlap of VOT values between fortis and lenis stops, and between lenis and aspirated stops indicates that VOT alone cannot be a reliable acoustic cue. Accordingly, previous research suggested that the laryngeal specification of the consonant is also cued by F0 of the following vowel, reporting that F0 of the following vowel is lower for the lenis than for the fortis and aspirated stops, with slightly higher F0 for the aspirated stop. Notably, the role of F0 has become more important as a result of a sound change in the stops of Seoul Korean where the VOT difference between the lenis and aspirated stops has decreased over decades (e.g., Silva, 2006; Wright, 2007; Kang & Guion, 2008; Perkins & Lee, 2010; Lee, 2013).

Several perception studies on Korean stops also indicate the importance of information conveyed by both the consonant and vowel. Cho (1996) investigated to what extent the vowel following a stop could help listeners distinguish the fortis, lenis and aspirated stops. Cho's (1996) subjects were asked to identify preceding stop consonants based only on the vowel portion. Cho (1996) reported 67% accuracy with only the vowel portion, and argued that native Korean speakers indeed use the acoustic information in the vowels in perceiving the three-way distinction among stops. Kim, Beddor and Horrocks (2002) tested the contribution of both the consonant and vowel to Seoul Korean listeners' perception of the stop with cross-spliced and vowel-only stimuli. Kim et al. (2002) noted a majority of lenis responses for the cross-spliced tokens containing a lenis vowel (i.e., 92% for fortis consonant + lenis vowel pair, 81% for aspirated consonant + lenis vowel pair), but only few lenis responses for the cross-spliced tokens with a lenis consonant (i.e., 4% for lenis consonant + fortis vowel pair, 6% for lenis consonant + aspirated vowel pair). Kim et al. (2002) attributed these perception results to the large F0 difference in vowels between lenis and fortis/aspirated stops; that is, the low F0 value associated with the lenis stop dominates the VOT consonantal cue in the perception of the lenis stop. For the perception of the fortis and aspirated stop, Kim et al. (2002) reported that while listeners were not successful in distinguishing the fortis from the aspirated stop for the vowel-only tokens, they could distinguish the two stops for the cross-spliced stimuli that contain relevant consonant portions combined with either the fortis or the aspirated vowel. Kim et al. (2002) concluded that the vowel portion alone was a sufficient cue for perceiving the lenis stop, while the perception of the fortis and aspirated stops relied on the combination of both consonant and vowel properties.

While the studies by Cho (1996) and Kim et al. (2002) explored the importance of each of the consonant' and vowel' portions for identifying the stop, Kim (2004) investigated trading relations between VOT and F0 in the perception of Korean stops. Kim (2004) used monosyllabic stimuli /C + a/ in which /C/ varied across the three types of laryngeal stops. Kim (2004) manipulated F0 in 10 Hz steps. But for VOT, Kim (2004) did not use manipulated tokens; rather, she used a great number of original tokens with a large VOT range. Kim (2004) reported that stimuli manipulated from the fortis stop, which has a relatively short VOT, were not influenced by changes in F0, showing 94% perception accuracy. Kim (2004) suggested that this is because of the extremely short VOT of the fortis stop. For the analysis of the lenis and aspirated stops, Kim (2004) noted that the perception of the lenis and aspirated stop indicates a phonetic trading relation between VOT and F0; F0 values at the category boundary become smaller as VOT values become larger. The results suggest that although a long VOT triggers more aspirated than lenis responses, a stimulus is likely to be perceived as lenis when the F0 of the following vowel is low. Conversely, although a short VOT triggers more lenis than aspirated responses, a stimulus is likely to be perceived as aspirated when the F0 of the following vowel is high. Based on these results, Kim (2004) argued that there is a phonetic trading relation between VOT and F0 regarding the lenis-aspirated distinction of Korean stops, and concluded that both VOT and F0 are equally important to the perception of Korean stops.

Overall, these previous perception studies successfully argued for the importance of the stop (VOT) and following vowel (F0), but the examination is limited to non-tonal Seoul Korean. As mentioned earlier, however, Kyungsang Korean is a pitch-accent language. That is, Kyungsang speakers already use pitch differences conveyed in vowels to cue lexical pitch accent contrasts as well as laryngeal contrasts. This means that, between Seoul and Kyungsang Korean, the roles of VOT and F0 in signaling distinctions among voiceless stops may differ.

1.2. Effects of lexical tone on the three-way laryngeal distinction in Kyungsang Korean

In a recent acoustic study, Lee and Jongman (2012) explored the three-way laryngeal distinction among stops in non-tonal Seoul and tonal Kyungsang Korean by examining stop properties (VOT) as well as following vowel properties (F0 and H1-H2). Lee and Jongman (2012) questioned to what extent the presence of tones in Kyungsang Korean affects Kyungsang speakers' stop production. Lee and Jongman (2012) demonstrated dialectal variation in the classification of the Korean stop, showing a different use of F0 and VOT cues for the stop in the two dialects. Adapted from Lee and Jongman (2012), Figures 2 and 3 present the F0 and VOT ranges, respectively, for the three-way stop distinction in Seoul and Kyungsang Korean.

Figure 2 shows that while the interquartile ranges of F0 (measured at the onset of the following vowel) are well separated across the three stops for Seoul, the F0 distribution for Kyungsang overlaps, in part because of variation between low and high tones of initial syllables. Lee and Jongman (2012) demonstrated that F0 is not a sufficiently reliable acoustic cue for Kyungsang speakers to distinguish the Korean stops, and suggested that its unreliability as a cue may be due to the presence of lexical pitch in Kyungsang speakers' phonology. Kenstowicz and Park (2006) compared their Kyungsang data with Seoul data from Cho et al. (2002) and also proposed that F0 was not a reliable cue for distinguishing stops for Kyungsang speakers, suggesting that H1-H2 may provide additional cues to compensate. However, in a direct comparison, Lee and Jongman (2012) did not find differences in H1-H2 between the two dialects.

VOT, however, clearly distinguishes the three stops for Kyungsang speakers, as shown in Figure 3. On the other hand, the VOT ranges overlap between lenis and aspirated stops in Seoul Korean. Statistics in Lee and Jongman (2012) indicated that VOT plays a bigger role for Kyungsang Korean than for Seoul in categorizing the three contrastive stops. Accordingly, they argued that VOT is a more reliable cue to the three-way contrast in Kyungsang Korean than in Seoul, and VOT makes up for the weakened role of F0 in Kyungsang. That is, with the results of F0 overlap but better classification based on VOT in Kyungsang Korean, Lee and Jongman (2012) concluded that Kyungsang speakers primarily use VOT to distinguish the three-way contrast, while Seoul speakers use a combination of F0 and VOT, and the use of the F0 cue for the laryngeal distinction is diminished for Kyungsang speakers due to their use of F0 for the purpose of pitch accent.

These acoustic findings by Lee and Jongman (2012) motivated the current study to test this dialectal variation for the Korean stops in perception in addition to production, exploring if the dialectal variation of Korean stops observed acoustically would also be observed in perception. If the different tonal systems between Seoul and Kyungsang Korean, namely the different phonologies of the two dialects, affect perception as well as production, we expect that listeners of tonal and non-tonal Korean dialects use different cue weightings in the perception of the three-way laryngeal stop contrast. If there is an inter-dialect difference in perception corresponding to the acoustic patterns, the importance of F0 in identifying stops would be greater for Seoul listeners than Kyungsang listeners; on the other hand, VOT would play a more important role for Kyungsang than Seoul listeners. Based on the acoustic findings in Lee and Jongman (2012), we specify our prediction for the perception of the fortis, lenis and aspirated stops.

First, according to its acoustic properties, the fortis stop would be well perceived with shorter VOTs and higher F0s by Korean listeners. However, based on the reported dialectal difference in VOT ranges (Fig. 3), we expect that although VOT and F0 cues work complimentarily in both dialects of Korean, the complimentary relationship between the two perceptual cues would differ between the two dialects. In Kyungsang Korean, the proximity of the VOT range for the fortis and lenis stops may form a moderately[1] ambiguous region of VOT between the two stops. Thus, for Kyungsang listeners, high F0s would be needed in the moderately ambiguous VOT region (e.g., 30 ms) to ensure fortis responses. In Seoul Korean, on the other hand, the fortis VOT range is well separated from the lenis stop, leaving almost no ambiguous VOT region between fortis and lenis. Thus, as long as a short VOT duration is provided Seoul listeners would perceive the fortis stop more accurately than Kyungsang listeners, relying less on the high F0 than Kyungsang. listeners.

Second, whereas identification of the lenis stop would be affected more by VOT than F0 for Kyungsang listeners, it would be affected more by F0 than VOT for Seoul listeners. Figure 3 indicates that the VOT of the lenis stop overlaps substantially with that of the aspirated stop in Seoul Korean, suggesting that VOT alone is not a reliable acoustic cue in distinguishing the lenis from the aspirated stop; on the other hand, the VOT range between the lenis and aspirated stops is well separated in Kyungsang Korean. Figure 2 shows that the F0 following lenis stops in Seoul Korean is significantly lower than following fortis or aspirated stops, whereas in Kyungsang Korean it overlaps with that following fortis stops. Thus, if this acoustic property is reflected in perception, we can expect that while VOT would play a robust role for Kyungsang listeners in the lenis perception, VOT would play a much smaller role for Seoul listeners, presumably due to severe confusion with the aspirated stop. For Seoul listeners, previous literature has also verified that VOT alone is insufficient for perception of the lenis stop, and the lenis percept is instead driven by low F0 (e.g., Cho 1996; Kim et al., 2002).

An inter-dialect difference is also expected for the perception of the aspirated stop. Across Seoul and Kyungsang Korean, longer VOTs and higher F0s would make both Seoul and Kyungsang listeners perceive an aspirated stop. However, as mentioned above, the VOT of the lenis and aspirated stops overlaps substantially in Seoul Korean, but not in Kyungsang, which led us to predict that VOT would be a better cue for Kyungsang listeners and F0 a better cue for Seoul listeners. In other words, a stimulus with a long VOT appropriate for the aspirated stop would be perceived as aspirated in the absence of a high F0 by Kyungsang listeners; on the other hand, Seoul listeners would not identify the same stimulus as aspirated unless a high F0 is provided.

2. Methods

To test the research hypotheses, we investigated the relative contributions of VOT and F0 to the perception of Korean stops in both Seoul and Kyungsang dialects. To observe the relation between the two acoustic parameters in the two dialects of Korean, we partially followed Kim's (2004) experimental design where perception was tested with manipulated F0. Furthermore, unlike Kim's study (2004) in which VOT was not manipulated, we manipulated VOT along a continuum, which allowed us to interpret the perception pattern in a systematic manner.

[1]This prediction might be more complicated if we consider the degree of overlap. The overlap between the fortis and lenis stops in Kyungsang Korean seems *moderate* compared to the overlap between lenis and aspirated stops in Seoul Korean that is *substantial*. Thus, consideration of the degree of overlap may lead to the prediction that although Kyungsang listeners would rely on F0 for the moderately overlapping VOT for the fortis stop perception, the reliance on F0 would not be greater than that in Seoul's lenis and aspirated percept.

2.1. Original base token

From the Korean triplet *p^{h}ul* grass', *pul* fire' and *p'ul* horn' the lenis stop *pul* produced by a male Kyungsang speaker (27 years old, South Kyungsang (Pusan city)) was chosen as the base token. Although the base token was selected from Kyungsang, not from Seoul, we would expect similar results with a Seoul token. This is because the acoustic properties (i.e., VOT and F0) of this base *pul* token were not particularly biased toward those of Kyungsang Korean according to Lee and Jongman (2012) which examined 8 male speakers of each dialect. Figures 2 and 3 indicate the minimum and maximum values of VOT and F0 between Seoul and Kyungsang Korean reported by Lee and Jongman (2012). The base token *pul* with the lenis stop had a VOT of 69 ms and an F0 of 110 Hz at vowel onset. Figure 3 indicates that the VOT duration of 69 ms for the base token is in the range of the lenis stop in Seoul Korean; Figure 2 shows that the F0 value of 110 Hz for the base token is in the range of the lenis stop in both Seoul and Kyungsang Korean. Thus, the selection of a base token with neutral acoustic properties minimized potential perceptual bias toward a particular dialect.

2.2. Manipulated stimuli

We manipulated all the stimuli from the single *pul* token with the lenis stop, which enabled us to control unintended variables such as vowel duration, intensity, or phonation type. For the stimulus manipulation, this study selected the maximum and minimum values of the VOT and F0 continua in order to fully encompass the VOT and F0 ranges observed in both dialects in Lee and Jongman (2012) as shown in Figures 2 and 3.

For the present perception test, the VOT duration of the base token was manipulated to span the range from 10 ms to 142 ms in 12 ms steps. To create these shorter and longer VOTs, we compressed and expanded the duration using the Praat manipulation function instead of reducing or adding to the aspiration portion. The manipulation procedure was as follows. First, we defined VOT as the interval between the release of the stop and the onset of voicing in the *Duration manipulation tier* of Praat and extracted the *Duration Tier* for this VOT portion. Next, we changed the original VOT value into the desired outcome ratio value. For example, to shorten VOT to 34 ms from the original duration of 69 ms, we used a ratio value of 0.492 (34 ms divided by 69 ms). Finally, we synthesized a new token with modified VOT by replacing the original duration tier with the new tier.

As for the F0 manipulation, since Lee and Jongman (2012) had only used the vowel /a/, F0 values for the vowel /u/ were estimated from Yang (1996), taking into account the intrinsic F0 difference between the /a/ and /u/ vowels. The F0 range was fixed from 99 Hz to 209 Hz in 10 Hz steps, slightly extending the F0 values of 99 Hz to186 Hz reported by Lee and Jongman (2012) for /a/. The large F0 range encompasses potential minimum and maximum values for the three Korean stops with the vowel /u/. In addition, since a monotone F0 from the onset to the offset of the vowel made the manipulated tokens sound unnatural, we used F0 contours. We estimated the F0 contour by calculating the percentage of the mean F0 change between the onset and offset of the vowel across the three stops; F0 was reduced from onset to offset by 11%.

These lower and higher F0s were created in the *Pitch manipulation tier* of Praat. In the *Pitch manipulation tier*, we lowered or raised the F0 contours by moving each pitch point that represents the F0 value at that point. For example, to raise the onset F0 to 209 Hz from the original value of 110 Hz, we moved the initial pitch point to a value of 209 Hz, and moved the final pitch point to 186 Hz, which is 11% lower than 209 Hz. The remaining F0 points between the onset and offset of the vowel were interpolated in Praat. This F0 manipulation was done using tokens for which VOT had already been manipulated, replacing the original F0 tier of the VOT-manipulated token with the new F0 tier. Overall, a total of 144 stimuli were created (12 levels of VOT × 12 levels of F0).

2.3. Subjects

Forty-two Korean listeners participated: twenty-one each from Seoul (4 females, 17 males) and South Kyungsang (5 females, 16 males). The listeners in the Seoul group were either from Seoul city or Kyunggi region where non-tonal standard Korean is spoken; the listeners in the Kyungsang group were either from Pusan or Ulsan cities, both of which are located in the South Kyungsang region. The age in the Seoul group ranged from 22 to 29 years old (mean = 26.2, SD = 2.2); the age in the Kyungsang group ranged from 20 to 65 years old (mean = 36.7, SD = 12.7)[2]. All of the speakers in each dialectal group had lived and been educated in the target dialect region with parents who spoke the same target dialect for at least 20 years. All the subjects were recruited in Seoul, Korea, and most of the subjects were graduate or undergraduate students at Seoul National University and Hanyang University. None of the speakers in either dialect group reported any speech or hearing disorders, and all of the speakers were literate in Korean.

2.4. Procedure

A total of 432 trials (144 stimuli × 3 repetitions) were randomly presented to each listener using the Paradigm software package (Perception Research Systems, Inc). Before the test, each subject was instructed to identify what they heard by clicking one of the three response options on a computer monitor, which were the words 풀(*p*^{h}*ul*), 불(*pul*), and 뿔(*p'ul*) in Korean orthography. The order of the response options was always the same. Prior to the actual test, 10 trials, a subset of actual stimuli in the main test, were given to the subjects with no feedback as a practice for familiarization with the test. After clicking one option among the three, the next stimulus was automatically presented to the listeners, and accordingly there was no time pressure. The experiment was conducted using headphones (SONY, MDR-V6) connected to a computer (Compaq CQ62-114TX). The test was conducted in quiet locations on campus, and the entire experimental session took approximately 30 minutes. Subjects were paid for their participation.

3. Results and discussion

3.1. Overall perception pattern

We start by discussing several aspects of the data plots: 1) the identification functions for VOT and F0 across Seoul and Kyungsang Korean, 2) overall dialectal difference in using

---

[2]Two Kyungsang participants were in their 60s; two were in their early 50s, and three were in their 40s.

VOT and F0, and 3) inter-dialect difference in the trading relation between the two perceptual cues.

Figures 4 and 5 present the identification patterns as a function of VOT and F0, respectively, for Seoul and Kyungsang dialect groups; Figure 4 only represents the function of VOT, and responses are averaged across the 12 different F0 values. Likewise, Figure 5 only represents the function of F0, and responses are averaged across VOT values.

Averaging across F0, Figure 4 indicates that VOT plays an important role in distinguishing the fortis from the aspirated stop in both Seoul and Kyungsang Korean. An extremely short VOT (e.g., 10 or 22 ms) provides Korean listeners with a salient perceptual cue for the fortis stop. In contrast, as long as Korean listeners are given a long VOT (e.g., 106 ms or longer), they tend to identify stops as aspirated. Slightly more lenis responses were obtained in the VOT range from 34 to 70 ms than at other VOT values for both Seoul and Kyungsang groups. However, unlike the fortis and aspirated stops that show a clear increase or decrease in responses depending on the long or short VOT, a change in VOT does not seem to provide a strong perceptual cue for the perception of the lenis stop. Instead, low F0 values increase the lenis stop responses, as indicated in Figure 5. In Figure 5, the responses for the lenis stop and the aspirated or fortis stops show opposite patterns; the proportion of lenis responses decreases as F0 increases, while the aspirated and fortis stop responses increase as F0 increases, though the positive effect of F0 for fortis responses seems less apparent than that for aspirated responses. These results suggest that VOT is more likely to be used as a perceptual cue for differentiating the fortis from the aspirated stop, while F0 is used for distinguishing the lenis stop from the other two.

We noted some dialectal differences in using VOT for the stop classification. First, while a short VOT of 10 or 22 ms was perceived as fortis more than 90% of the time by Seoul speakers, the same VOT range was perceived as fortis less often by Kyungsang speakers. Second, although both Seoul and Kyungsang listeners show more aspirated responses as VOT increases, long VOTs (e.g., 118–142ms) elicit fewer aspirated responses for Seoul listeners compared to Kyungsang. Third, Seoul listeners heard the stimuli with long VOTs as lenis stops more than Kyungsang listeners did, whereas Kyungsang listeners were more likely to perceive the stimuli with short VOTs as lenis stops than Seoul listeners. It is also noted that the two dialect groups use F0 in a different way in their stop perception. While Seoul listeners perceived 20% of the stimuli with low F0 values (e.g., 99 or 119Hz) as the aspirated stop, the proportion of aspirated responses was almost 40% for Kyungsang listeners. In other words, while low F0 values are effective for distinguishing the lenis from both the fortis and aspirated stops in Seoul Korean, the same F0 range is effective mainly for the distinction between the lenis and fortis stop in Kyungsang Korean. In sum, these plots indicated dialectal differences as well as the overall perception pattern for the Korean stop. Seoul and Kyungsang groups differed in the VOT ranges at which they showed perceptual confusion, and low F0 functions differently between the two dialects, particularly for distinguishing the aspirated stop from the other two.

3. 2. Heat plots and logistic regression analysis

The current study closely examined dialectal similarities and differences in the relationship between VOT and F0 as cues to the stop choice through a three-dimensional plot referred to as a *heat plot.* A heat plot uses gradations of color to display the identification responses to stimuli in which the two perceptual cues of VOT and F0 are varied simultaneously. The horizontal and vertical axes represent the12 levels of VOT and 12 levels of F0, respectively. The darkness of each cell represents the proportion of identification responses: the higher the proportion of responses of the target category is, the darker the cell is. While two-dimensional plots like Figures 4 and 5 are familiar and useful for understanding the general pattern due to their ability to show all three responses together, the three-dimensional heat plot provides a clearer illustration of the presence or absence of a trading relation between the two perceptual cues without having to compare multiple plots.

For statistical assessment, the current study analyzed participants' choice proportions using binary logistic regression, which allowed us to create models predicting participants' classification of items as a function of VOT and F0, and to predict whether there were differences in the means or slopes of these regression models between Dialect groups. We adopted a procedure for repeated-measures regression analysis based on Lorch and Myers (1990; see also Benki, 2001). This procedure involves first pooling the trials from all subjects together into a single analysis to determine which predictors to include in the regression model (referred to as the pooled model), and then using those predictors to fit a regression model to each subject. The final regression model (the population-averaged model) is derived by, for each predictor, averaging the coefficients derived for each subject. The coefficients in the population-averaged model can be tested statistically using *t*-tests. In the current study, the overall significance of each predictor across dialects was tested using one-sample *t*-tests (comparing the coefficient against zero), and inter-dialect differences were tested using independent samples *t*-tests (comparing the coefficient for one dialect against that for another). This procedure was repeated three times, using Fortis, Lenis, or Aspirated response as the outcome variable, respectively. Each model maximally included an intercept, two continuous predictors (VOT and F0), one categorical predictor (Dialect), and all possible interactions. In order to meaningfully interpret mean differences between categorical predictors in the presence of interactions with continuous predictors, the continuous predictors (VOT and F0) were centered by subtracting the mean of that predictor from each observation. The Seoul dialect group was used as the baseline, and thus effects of Dialect indicate how much more or less likely Kyungsang listeners were to choose a given outcome (Fortis, Lenis, or Aspirated). For each analysis, we first included each of VOT, F0 and Dialect predictors (blocks 1 and 2), and then sequentially added the interaction terms of Dialect by VOT (block 3), Dialect by F0 (block 4) and Dialect by VOT by F0 (block 5). Based on Chi-squared tests, we determined whether the addition of a given term or set of terms significantly improved the model, and chose an optimal model to report the results.

3. 2. 1. Fortis stop identification—Figure 6 presents the identification responses by each dialect group for the fortis stop.

First, we examined perceptual similarities in the fortis response between Seoul and Kyungsang Korean. Figure 6 shows that items with VOTs between 10 and 46 ms were sometimes perceived as fortis in both dialects. At extremely short VOTs (10 or 22 ms), listeners mainly perceived the fortis stop, with the darkest cells indicating a fortis identification percentage of over 80%; at moderately short VOTs (34 or 46 ms), listeners' fortis choices decreased as indicated by the grey shading. Figure 6 also suggests that F0 plays a role in the perception of the fortis stops at moderately short VOTs (34 or 46 ms): at a low F0, listeners are less likely to perceive the tokens as fortis, but at a high F0 they are more likely to perceive the tokens as fortis. This suggests that when a token has ambiguous VOT (34 or 46 ms), the other cue, F0, contributes to perception, whereas when the VOT is unambiguous (10 ms) it alone is sufficient to identify the stop without input from the F0 cue. In other words, when one cue is ambiguous (e.g., VOT), the role of the other cue is strengthened (e.g., F0), suggesting these two perceptual cues trade with each other in identifying the fortis stop.

The logistic regression analysis confirmed the pattern suggested in Figure 6. As a preliminary estimation, a series of nested logistic regression models were built from the pooled data from all subjects, adding the following terms at each step: 1) VOT, F0, and VOT by F0 (χ^2 (3) = 12010.822, p < 0.001), 2) Dialect (χ^2 (1) = 125.189, p < 0.001), 3) Dialect by VOT (χ^2 (1) = 3.956, p = 0.047), 4) Dialect by F0 (χ^2 (1) = 1.184, p = 0.276), 5) Dialect by VOT by F0 (χ^2 (1) = 41.747, p < 0.001). The model was significantly improved by adding the three-way interaction term, Dialect by VOT by F0, and thus we modeled each subject's responses using the full model with all predictors included. To account for covariance within subjects, we then constructed a saturated logistic regression model for each subject and tested the significance of the coefficients across subjects using one-sample *t*-tests. Table 1 shows the results of one-sample *t*-tests for Fortis, which tested the coefficients averaged across all subjects for each predictor.

Across subjects, the coefficient for VOT is significantly less than 0 (p < 0.001), indicating that as VOT increases, both Seoul and Kyungsang listeners become less likely to identify a token as fortis. However, the coefficient for F0 does not significantly differ from 0 (p = 0.643). The averaged coefficient for the VOT by F0 interaction was also significantly less than 0 (p < 0.001), indicating that as F0 increases, Seoul and Kyungsang listeners are even less likely to choose an item as fortis at longer VOTs (i.e., as F0 increases, the effect of VOT is boosted (see also Fig. 7)).

Figure 6 also suggests that there may be a dialectal difference in the fortis responses. Seoul listeners show relatively stable fortis responses across F0 levels at the VOT range of 10–22 ms, whereas Kyungsang listeners' likelihood of choosing fortis appears to be affected by F0. However, the apparent dialectal difference in Figure 6 was not supported by the statistical model in fortis identification. To statistically test for dialectal differences, we compared the regression coefficients for Seoul and for Kyungsang speakers using independent samples *t*-tests.

Table 2 shows that the averaged coefficients for VOT, F0 and VOT by F0 were not significantly different between Seoul and Kyungsang groups. These results suggest that

Seoul and Kyungsang listeners use VOT and F0 in a similar way to perceive the fortis stop. Figure 7 shows the choice proportions predicted by the model along with the actual choice proportions for Fortis at three levels of F0 (low, mid, and high; top) and the relationship between F0 and fortis probability at three levels of VOT (short, mid, and long; bottom). Because the differences between speaker groups were not significant, a single curve is used to model responses by both groups.

Figure 7 illustrates that, across levels of F0 and across dialects, as VOT durations increase, fortis responses decrease; while the probability of fortis responses was over 80% at the short VOT of 10 ms, at the mid and long VOTs of 70 ms and 142 ms fortis stops are not predicted at all. Notably, the effect of VOT is modulated by F0: while the shortest VOT of 10 ms is associated with only about an 80% chance of choosing "fortis" at an F0 of 99 Hz, the same VOT duration is associated with a near-100% chance of fortis choice at 209 Hz.

In sum, consistent with our prediction, the logistic regression assessment for Fortis showed that short VOT duration is crucial to trigger the fortis judgment, and the effect of short VOT is boosted at high F0 for Seoul and Kyungsang listeners. Contrary to predictions, however, the functions of VOT and F0 do not differ between dialects.

3. 2. 2. Lenis stop identification—Figure 8 presents the identification responses by each dialect group for the lenis stop.

Lenis responses were obtained mainly in the VOT range from 34 to 70 ms for both Seoul and Kyungsang, as indicated by the darkest cells representing VOT-F0 combinations which elicited over 80% of lenis identification. Unlike the fortis responses, which show a steady decrease as VOT increases, the relationship between VOT and lenis responses is non-monotonic: the VOT change from short (e.g., 10 ms) to intermediate (e.g., 70 ms) increases the lenis responses, but the change from intermediate (e.g., 70 ms) to long (e.g., 142 ms) decreases the lenis responses. Figure 8 also suggests that the lenis choice for both Seoul and Kyungsang listeners depends on F0 as well as VOT; the lower the F0, the more lenis responses are obtained. Notably, the dependency on F0 varies according to how salient or ambiguous VOT is. Specifically, in Figure 8, VOTs shorter than 34 ms and longer than 70 ms are more ambiguous relative to the VOT range from 34 to 70 ms for lenis responses as indicated by the reduced lenis responses. In the more ambiguous VOT territory, listeners tended to hear the stimuli as lenis stops, relying more on a low F0 (e.g., 99–119 Hz). This suggests that VOT and F0 cues trade with each other in judging the lenis stop.

The logistic regression analysis confirmed the pattern suggested in Figure 8. A pooled logistic regression model for Lenis was conducted as a preliminary estimate of the model, adding the following terms at each step: 1) VOT, F0 and VOT by F0 (χ^2 (3) = 5659.696, $p <$ 0.001), 2) Dialect (χ^2 (1) = 6.173, p = 0.013), 3) Dialect by VOT (χ^2 (1) = 139.762, $p <$ 0.001), 4) Dialect by F0 (χ^2 (1) = 4.019, p = 0.045), 5) Dialect by VOT by F0 (χ^2 (1) = 14.06, $p <$ 0.001). Since the model was significantly improved by adding the three-way interaction term, Dialect by VOT by F0, we modeled each subject's responses using the full model with all predictors included. In addition, because the lenis response proportion clearly had a non-monotonic relationship with VOT (i.e., increasing VOT increased the lenis response proportion at low levels of VOT but decreased it at high levels), we also included VOT^2 as a term in the model. As a second phase of statistical assessment, the population-averaged model tested how consistent the individual coefficients are for Lenis. Table 3 shows the results of one-sample *t*-tests for Lenis, which tested the overall significance of each coefficient regardless of dialect.

Across subjects, the coefficient for F0 is significantly less than 0 ($p < 0.001$), indicating that as F0 increases, the likelihood of lenis responses decreases. Moreover, the significant coefficient of VOT^2 further indicates that the effect of VOT is not monotonic. That is, at low levels of VOT, the likelihood of lenis responses increases as VOT increases; as VOT reaches higher levels, though, the likelihood of lenis responses decreases as VOT continues to increase (see Fig. 9). The averaged coefficient for VOT by F0 was also significantly less than 0 across the two dialects ($p < 0.001$), suggesting that the effect of F0 is boosted as VOT increases; as VOT increases, Seoul and Kyungsang listeners are even less likely to perceive a high-F0 stimuli as lenis.

Figure 8 also suggests a dialectal difference in the perception of the lenis stop. While Kyungsang listeners showed a gradual decrease in lenis responses as VOT increases, Seoul listeners did not; even at long VOTs (i.e., 106–142 ms) Seoul listeners perceived tokens with low F0 values (i.e., 99–119 Hz) as lenis about half of the time. This observation may suggest that while the long VOTs alone provide a more salient perceptual cue for the non-lenis judgment (i.e., aspirated) for Kyungsang listeners than for Seoul, Seoul listeners' lenis judgment depends more on F0 than Kyungsang listeners' even at long VOTs. To statistically test for dialectal differences, the regression coefficients for Seoul and for Kyungsang speakers were compared using independent samples *t*-tests.

Table 4 shows that the coefficient for VOT is significantly more negative for Kyungsang listeners than Seoul listeners, indicating that the effect of VOT is stronger for Kyungsang listeners ($p = 0.041$). On the other hand, there was no significant inter-dialect difference for F0 ($p = 0.809$). Finally, the inter-dialect difference in the coefficient for VOT by F0 was near-significant at $p = 0.064$, indicating that the dialectal difference in the effect of VOT tends to be modulated by F0 and vice versa. Figure 9 shows the choice proportions predicted by the model along with the actual choice proportions for Lenis at three levels of F0 (low, mid, and high; top) and the relationship between F0 and fortis probability at three levels of VOT (short, mid, and long; bottom).

Regarding the inter-dialect difference, Figure 9 (top) illustrates that the effect of VOT is greater (steeper) for Kyungsang than for Seoul listeners—an increase in VOT causes a greater reduction in lenis responses for Kyungsang speakers than for Seoul speakers. This suggests that VOT is a better predictor for lenis responses in Kyungsang than Seoul listeners. The extent to which the VOT and F0 effects modulate one another tends to be greater in Seoul speakers than Kyungsang speakers. As indicated in the top half of Figure 9, the effect of VOT (i.e., the shape of the regression curve) changes more drastically between different levels of F0 in Seoul speakers than it does in Kyungsang speakers. The same is true of the bottom half of Figure 9: the effect of F0 changes more across VOTs in Seoul speakers than in Kyungsang speakers.

The dialectal difference for the perception of the lenis stop is consistent with the acoustic characteristics of the lenis stop between Seoul and Kyungsang Korean, which confirms our prediction. Based on the acoustic characteristics of the lenis stop (Figs. 3 and 4), we predicted that Seoul listeners' lenis judgments may depend more on low F0 values than those of Kyungsang listeners to compensate for the severe overlap of VOT between the lenis and aspirated stops; on the other hand, the lack of ambiguity in VOT range in Kyungsang may help Kyungsang listeners distinguish the lenis and aspirated stops with less dependence on F0 than Seoul.

To sum up, the logistic regression models revealed that listeners in both Seoul and Kyungsang Korean are more likely to perceive the lenis stop with intermediate VOT and lower F0. Importantly, there were several noteworthy inter-dialect differences. Consistent with our prediction, VOT plays a larger role in lenis stop identification for Kyungsang listeners than it does for Seoul listeners, and VOT tends to interact with F0 more strongly in Seoul listeners than Kyungsang listeners.

3. 2. 3. Aspirated stop identification—Figure 10 presents the identification responses by each dialect group for the aspirated stop.

Across Seoul and Kyungsang Korean, stimuli with long VOT and high F0 were more likely to be perceived as aspirated. Figure 10 shows that, at long VOT values from 106–142 ms, listeners are more likely to perceive the aspirated stop, as indicated by darker cells compared to moderately long VOTs from 58–94 ms. Moreover, F0 values above 149 Hz trigger more aspirated responses than F0 values below 149 Hz. Figure 10 also revealed the complementary use of VOT and F0: although moderately long VOTs (e.g., 58–94 ms) alone are not sufficiently robust to trigger aspirated responses, listeners do perceive the aspirated stop in this VOT range when provided with higher F0 values.

The logistic regression analysis confirmed the pattern suggested in Figure 10. As a preliminary estimation, the pooled logistic regression model was first conducted for Aspirated in the following sequence: 1) VOT, F0 and VOT by F0 (χ^2 (3) = 13654.261, $p <$ 0.001), 2) Dialect (χ^2 (1) = 131.041, $p < 0.001$), 3) Dialect by VOT (χ^2 (1) = 38.403, $p <$ 0.001), 4) Dialect by F0 (χ^2 (1) = 9.46, $p = 0.002$), 5) Dialect by VOT by F0 (χ^2 (1) = 3.588, $p = 0.058$). Since the model was not significantly improved by adding the three-way interaction term, Dialect by VOT by F0, we modeled each subject's responses using a model with all predictors and two-way interactions, but did not compare the VOT by F0 coefficient between groups in the follow-up analysis. Table 5 shows the results of the one-sample *t*-tests for the coefficients in the Aspirated model.

Across dialects, the averaged coefficients for VOT and F0 are each significantly greater than 0 ($p < 0.001$ for each), indicating that as VOT and F0 increase, Seoul and Kyungsang listeners were more likely to choose aspirated over non-aspirated (fortis, lenis). In addition, the interaction of VOT by F0 was significantly greater than 0 across dialects ($p < 0.001$), suggesting that the effect of VOT is boosted as F0 increases, and vice versa.

Figure 10 also suggests a dialectal difference in judging the aspirated stop. Comparison between Figure 8 and Figure 10 indicates that as VOT increases from 82 to 142 ms, the gradual decrease in lenis responses in Kyungsang Korean coincides with an increase in aspirated responses; on the other hand, in Seoul Korean, the presence of more lenis responses at long VOTs (82–142 ms) is linked to the presence of fewer aspirated responses in this VOT range. The different proportion of the aspirated responses between Seoul and Kyungsang Korean is related to F0 change. In the long VOTs (106–142 ms), Seoul speakers' likelihood of making an aspirated response decreases more at low F0s than Kyungsang speakers' likelihood of making an aspirated response does. This observation seems to suggest that a long VOT alone is a more reliable perceptual cue for Kyungsang than for Seoul listeners for non-lenis (aspirated) responses, and F0 affects Kyungsang listeners' perception of the aspirated stop less than Seoul listeners. Independent samples *t*-tests confirmed the dialectal difference by showing the group difference in the regression coefficient in aspirated responses.

In Table 6, the averaged coefficients for VOT and F0 were not significantly different between Seoul and Kyungsang groups (p = 0.16 for VOT, p = 0.82 for F0), indicating that Seoul and Kyungsang listeners use VOT and F0 in a similar way to perceive the aspirated stop. However, the intercept is significantly greater for Kyungsang listeners than for Seoul (p = 0.038), indicating that Kyungsang listeners were overall more likely than Seoul listeners to choose the aspirated stop. Figure 11 shows the choice proportions predicted by the model along with the actual choice proportions for Aspirated at three levels of F0 (low, mid, and high; top) and the relationship between F0 and fortis probability at three levels of VOT (short, mid, and long; bottom).

Figure 11 (top) shows positive-sloping regression curves for VOT across F0 levels and Dialect; this means that the longer the VOT is, the more aspirated choices are predicted for both Seoul and Kyungsang groups. Figure 11 (bottom) shows a positive-sloping curve for F0 as well, and shows that the relationship between F0 and perception of aspirated stops is stronger at mid and long VOTs than at short VOTs.

Finally, Kyungsang listeners perceived more aspirated stops than Seoul listeners as indicated by a higher regression line in Figure 11, indicating that the perceptual boundaries for VOT and F0 occur earlier in Kyungsang than Seoul. In other words, Kyungsang listeners hear items as aspirated at shorter VOT and lower F0 than Seoul listeners.

In sum, the statistical assessment for Aspirated showed that Seoul and Kyungsang listeners are more likely to perceive the aspirated stop at long VOTs and high F0s. The perception of the aspirated stop at a long VOT is facilitated with a high F0. Regarding the dialectal difference, Kyungsang speakers are more likely to hear a stop as aspirated, although the dialect effect does not interact with VOT and F0.

3. 3. Discussion

The plot observation and logistic regression analysis indicated that Seoul and Kyungsang Korean listeners use the perceptual cues of VOT and F0 differently to distinguish Korean stops, as reflected by differences in the phonetic trading relation between VOT and F0.

First, we noted similarities in the relationship between VOT and F0 in the perception of the three-way stop distinction between Seoul and Kyungsang Korean (Figs. 4 and 5). Consistent with previous research (e.g., Lisker & Abramson, 1964; C.-W. Kim, 1965; Han & Weitzman, 1970; Cho, 1996; Cho et al.,, 2002; Kim et al., 2002), short and long VOTs play an important role in distinguishing the fortis from the aspirated stop, and low F0 values provided a strong cue to the lenis stop in both Seoul and Kyungsang Korean. Heat plots (Figs. 6, 8 and 10) further suggested that Korean listeners use VOT and F0 complementarily to identify the stops, and the trading relation between the two cues is present not only in Seoul but also in Kyungsang Korean. The trading between VOT and F0 in the perception of Korean stops is not surprising. Repp (1983, pp. 341) states that “When several different acoustic cues contribute to the perception of a phonetic distinction, a trading relation among the cues can be demonstrated in an identification task as long as the speech stimuli are phonetically ambiguous”. Whalen, Abramson, Lisker and Mody (1993) further noted that F0, a secondary cue to voicing in English stops, also provides voicing information even for unambiguous VOTs, suggesting use of all acoustic information in speech perception. The perception of Korean stops also involves multiple acoustic cues, namely at least VOT and F0, and we established that these two cues trade with each other when one cue is phonetically ambiguous or when an inappropriate cue is accompanied by an unambiguous cue.

The present perception pattern is in line with Kim’s findings (2004) of a phonetic trading relation between VOT and F0 in Seoul Korean. Specifically, the current study showed that the perception of the fortis stop at extremely short VOTs (10–22 ms) was robust for Seoul listeners, and the non-overlapping VOT range of the fortis stop with the other two stops explained the non-dependence on F0. Kim (2004) noted that perception of the fortis stop with its clearly distinctive VOT mostly relies on VOT and is less influenced by F0. For the lenis and aspirated stop perception, Kim (2004) reported a phonetic trading relation between VOT and F0. Therefore, Kim (2004) pointed out that investigating the relations among various perceptually relevant acoustic cues is as important as investigating the cues separately. The results in the present perception study support the argument on cue relations. In most of the previous acoustic and perception studies, the roles of VOT and F0 were investigated separately; these previous studies mostly described the fortis stop as having a short VOT and high F0, the lenis stop as having an intermediate VOT and low F0, and the aspirated stop as having a long VOT and high F0 (e.g., Lisker & Abramson, 1964; Han & Weitzman, 1970; Kim, 1994; Cho, 1996; Kim et al., 2002). However, as seen in the present study, the perception patterns in Seoul Korean did not always match these reported characteristics. Specifically, Seoul listeners were not always likely to perceive the low F0 tokens as the lenis stop, particularly if the token had an extremely short VOT. Likewise, a long VOT does not always guarantee the aspirated stop response; if Seoul listeners were given a low F0, they tended to perceive the lenis stop. Therefore, although the three-way distinction of Korean stops may be characterized with those separate cues in terms of acoustic properties, we should be aware that those cues work together and affect each other in perception.

Second, regarding the dialectal difference, the heat plot observation suggested that although Seoul and Kyungsang listeners use the VOT and F0 cues in somewhat similar ways to identify the three-way stop distinction, each of these dialect groups showed notable differences in their use of VOT and F0. While the two dialects use both VOT and F0 perceptual cues complementarily, the degree of the dependency on one perceptual cue is not always consistent between the two dialects. The logistic regression analyses confirmed the inter-dialect difference in the VOT and F0 cue weightings for the perception of the lenis and aspirated stops. Figure 12 presents a collective heat plot representing identification percentages for all three stops. Dark and light gray indicate cells in which a certain stop was chosen over 75% or 50–74% of the time, respectively. White indicates cells in which the three stops were confused (e.g., 49% of fortis, 49% of lenis, 2% of aspirated identification). In each portion of the figure, the solid line indicates the region where lenis responses were made. Figure 12 allowed us to examine the perception pattern for all three stops simultaneously in a VOT and F0 domain and to readily see which stop is likely to be perceived in a particular VOT/F0 condition by Seoul and Kyungsang listeners.

For the perception of the fortis stop, we predicted that the proximity in VOT range between the fortis and lenis stops in Kyungsang Korean compared to Seoul would require Kyungsang listeners to more depend on high F0s than Seoul listeners to ensure fortis responses (see Fig. 3). Contrary to the prediction, however, there was no such reported dialectal difference in the logistic regression. That is, for the perception of the fortis stop, an extremely short VOT alone was sufficient to hear stops as fortis for both Seoul and Kyungsang listeners, and the proximity in VOT between fortis and lenis stops for Kyungsang listeners does not affect their use of F0 unique to Seoul listeners. For fortis responses, although F0 change alone does not predict fortis responses in the two dialects, the effect of VOT is modulated by F0 in that a short VOT with high F0 is more likely to be perceived as a fortis stop than an item of the same VOT with low F0.

The results for the lenis and aspirated stops support our hypothesis, indicating that the dialectal difference in the VOT range may account for the different use of the two perceptual cues of VOT and F0. We predicted that the perception of the two stops by Seoul listeners depends less on VOT than for Kyungsang listeners due to VOT overlap, and thus F0 may compensate for the ambiguity in the VOT cue for Seoul listeners. As observed in Section 3.2. and Figure 12, extremely long VOTs alone were more reliable for Kyungsang listeners for the aspirated percept across F0 levels than for Seoul. On the other hand, these long VOTs were less salient for Seoul listeners, showing perceptual confusion between the aspirated and lenis stops.

For lenis judgments, the logistic regression results revealed that VOT change predicted the lenis percept better in Kyungsang Korean than in Seoul. Figure 12 also shows that VOT becomes an unambiguous cue for the aspirated stop much earlier for Kyungsang than for Seoul listeners: while a VOT longer than 82 ms provides Kyungsang listeners with a reliable cue for the non-lenis judgment (i.e., aspirated), the same VOT is ambiguous for Seoul listeners. In addition, the dialectal difference in VOT seems to be reflected in the contribution of F0 between the two dialects. Specifically, the role of low F0 values (i.e., 99–139 Hz) for Kyungsang listeners is restricted to ambiguous VOTs between 46 and 82 ms in the lenis and aspirated stop distinction. But Seoul listeners use the low F0 values throughout VOTs longer than 34 ms to distinguish the lenis from the aspirated stop.

Finally, for the aspirated stop perception, the logistic regression showed that the two dialects have different perceptual boundaries for VOT and F0 as shown by different intercepts; the perceptual boundaries for VOT and F0 occur earlier in Kyungsang than Seoul, suggesting that Kyungsang listeners hear a stop as aspirated at shorter VOTs and lower F0s as compared to Seoul listeners. Figure 12 also shows that while Kyungsang listeners consistently perceive the aspirated stop starting at a VOT of 46 ms with F0s of 159 Hz and above, Seoul listeners' aspirated percept at 46 ms is limited to the extremely high F0s of 199 and 209 Hz, verifying the earlier perceptual boundary for VOT. Regarding the boundary for F0, Kyungsang listeners prevailingly heard stops as aspirated even at the lowest F0 of 99 Hz as long as stimuli have VOTs longer than 94 ms, while Seoul listeners' aspirated percept at 99 Hz is limited to the longest VOT of 142 ms. In other words, Kyungsang listeners do not require high F0s as much as Seoul listeners do in order to ensure their aspirated stop perception.

Importantly, these results verify that the dialectal difference observed acoustically shows up perceptually in accordance with the dialectal difference in VOT range and cue weightings. A recent acoustic study, Lee and Jongman (2012), reported that Seoul speakers are more likely to use a combination of VOT and F0, while Kyungsang speakers primarily use VOT as a single strong cue. The acoustic finding in Lee and Jongman (2012) is compatible with the present perceptual finding that the role of VOT is greater in Kyungsang than in Seoul with respect to the lenis and aspirated distinction, and Kyungsang listeners are less sensitive to F0 change in the perception of the aspirated stop.

We suggest two possible factors that may cause the dialectal differences in stop perception as well as production. One is the different tonal system between the two dialects and the other is the diachronic change in VOT of the aspirated stop in Seoul Korean.

In their acoustic study, Lee and Jongman (2012) suggested that the presence of pitch accent in Kyungsang Korean makes F0 a less reliable acoustic cue for the purpose of the three-way laryngeal distinction. Lee and Jongman (2012) argued that "since Kyungsang speakers are already using F0 to distinguish tonal contrasts, the use of F0 for the purpose of classifying the three laryngeal distinctions seems to be diminished (p. 167)". The argument by Lee and Jongman (2012) is illustrated with Figure 2 of the present paper (p. 10). Specifically, Figure 2 indicates that Seoul and Kyungsang speakers with different tonal systems have different F0 spaces. In non-tonal Seoul Korean, the F0 space can be distinctly divided according to the three-way laryngeal contrast, patterning from high to low in the order of aspirated, fortis and lenis stops. In Kyungsang Korean, the F0 space is more gradually divided both by the three stops and the initial high and low lexical tones. Kyungsang Korean also showed the F0 pattern of Lenis < Fortis < Aspirated across the initial high and low lexical tones; Kyungsang speakers showed high and low F0 for the initial high and low lexical tones, respectively, across the three stops. This acoustic pattern was observed not only by Lee and Jongman (2012), but also by Kenstowicz and Park (2006); both studies suggested that tonal Kyungsang Korean needs a different phonological representation for the F0 space than non-tonal Seoul Korean. Kenstowicz and Park (2006), for example, noted that the three stops can be specified with the feature of [±stiff vocal folds] (Halle & Stevens, 1971) in non-tonal Seoul Korean, assigning [+stiff vocal folds] and [−stiff vocal folds] for the aspirated/fortis and for the lenis stops, respectively[3]. However, for tonal Kyungsang Korean, Kenstowicz and Park (2006) noted that the feature of [±upper] is also required for specifying the initial high and low lexical tones, and suggested four feature combinations of F0 to capture both the lexical pitch and the three-way stop contrast as in 1) [+upper, +stiff vocal folds] for initial high toned aspirated/fortis, 2) [−upper, +stiff vocal folds] for initial low toned aspirated/fortis, 3) [+upper, −stiff vocal folds] for initial high toned lenis, 4) [−upper, −stiff vocal folds] for initial low toned lenis. Kenstowicz and Park (2006) explained that while F0 space should be shared for the realization of lexical contrast and laryngeal contrast in tonal Kyungsang Korean, the entire F0 space is available for laryngeal contrast in non-tonal Seoul Korean (p. 6). Comparable explanations have been proposed by previous studies examining consonant voicing effects on F0 in tonal languages such as Yoruba and Cantonese, implying that less of the F0 space is available for cueing contrasts in the prevocalic consonants for speakers of tonal languages compared to non-tonal languages (e.g., Hombert, 1978; Francis et al. 2006). In this sense, since Seoul Korean specifies the stop only with [±stiff vocal folds] in F0, Kyungsang speakers who need to specify the stop both with [±stiff vocal folds] and [±upper] in the phonological F0 dimension were theoretically assumed to be less able than Seoul speakers to readily use F0 differences to signal high pitch for aspirated/fortis [+stiff vocal folds] and low pitch for lenis [−stiff vocal folds] because of the restricted availability of F0 to signal the laryngeal contrast. This theoretical assumption was empirically supported in Lee and Jongman (2012) where the initial low tone fortis and high tone lenis stops in Kyungsang were not distinct in terms of F0.

Therefore, we might conclude that the F0 cue to the three-way laryngeal stop is less effectively used by Kyungsang than Seoul speakers because tonal Kyungsang speakers are likely to use F0 differences to signal the high and low lexical tones. The current results support this argument by showing 1) decreased sensitivity to F0 in perceiving the aspirated stop, and 2) more stable lenis responses across F0 values. This leads us to conclude that different phonologies between the two dialects influence not only the production but also the perception of speech, which might be in line with the influence of L1 phonology on L2 speech production and perception in cross-linguistics studies (e.g., Miyawaki, Strange, Verbrugge, Liberman, Jenkins & Fujimura, 1975; Williams, 1979; Beddor & Strange, 1982; Gottfried, 1984; Yamada & Tohkura, 1992; Best & Strange, 1992; Strange, 1995; Chen, Robb, Gilbert & Lerman, 2001; Xu, Gandour & Francis, 2006).

Notably, although the role of F0 is reduced in Kyungsang Korean compared to Seoul, VOT played a stronger role in Kyungsang Korean compared to Seoul both in production and perception. Under the notion of a phonetic trading relation, when one cue is weakened, other cues compensate for the weakened cue to maintain phonetic equivalence. Thus, Kyungsang speakers may try to maximize the VOT difference to maintain the three-way stop contrast. By maximizing the VOT difference as a consequence of the weakened F0 cue for stops, Kyungsang speakers can successfully make the three-way laryngeal contrast as well as the initial high and low lexical pitch accent contrast.

[3]Kenstowicz and Park (2006) considered the lenis stop and the fortis/aspirated stops as underlyingly voiced and voiceless, respectively, following Kim and Duanmu (2004).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Another explanation for the dialectal difference in the perception of the Korean stop can be found in Seoul Korean. Seoul listeners' strong dependency on F0 to distinguish the lenis from the aspirated stop may be a consequence of the weakening VOT distinction between the lenis and aspirated stops. The ongoing diachronic change in the stop has been demonstrated by many researchers (e.g., Silva, 2006; Wright, 2007; Kang & Guion, 2008; Perkins & Lee, 2010), reporting that the VOT difference between the lenis and aspirated stops in Seoul Korean has decreased over decades. Silva (2006) argues that the underlying contrast between the lenis and aspirated stops is maintained by a lower F0 in lenis than aspirated stops, that is, phonetic change occurs without a phonemic shift despite the neutralized VOT. Therefore, it may be the case that the Seoul listeners' primary use of F0 for the lenis and aspirated distinction is triggered by the loss of the VOT distinction, which is opposite from Kyungsang Korean where the emphasized VOT difference is triggered by the reduced F0 cue due to the presence of lexical tone.

5. Conclusions

The current study tested the perception of the three-way distinction among Korean voiceless stops in the non-tonal Seoul and tonal Kyungsang dialects, examining the functions of VOT and F0 between the two dialects with different phonologies.

Inspection of heat plots revealed that Seoul and Kyungsang listeners exhibit a different phonetic trading relation between VOT and F0, and this dialectal difference reflects the different VOT ranges between the two dialects. The logistic regression analysis confirmed that the two dialects use the perceptual cues of VOT and F0 differently, particularly for the lenis and aspirated stop. While Seoul listeners rely primarily on F0 for making lenis responses and on VOT and F0 for aspirated responses, F0 plays a less important role in modulating both lenis and aspirated responses for Kyungsang listeners than it does for Seoul. Our findings suggest that Seoul listeners use F0 as a primary and VOT as a secondary cue to the perception of the lenis and aspirated stops, whereas Kyungsang listeners primarily use VOT, and F0 secondarily. Based on the acoustic and perceptual correspondence, we suggested that the presence of lexical tone in Kyungsang Korean made F0 differences a less reliable cue to the laryngeal distinction than in Seoul, and the on-going diachronic change in the aspirated stop in Seoul Korean made VOT differences less salient than in Kyungsang. Despite the weakened cue in each dialect, the three-way laryngeal contrast is maintained by strengthening the other cue for each dialect. Therefore, it is concluded that although the difference in phonology between the two dialects influences the way that phonemes are perceived, the phonetic trade-off between two perceptual cues allows each dialect to maintain the phonemic distinction in its own way.

In future study, testing the effect of initial F0 contours as well as overall F0 could add to our understanding of the role of F0 in the three-way laryngeal contrast in Korean. In his cross-language study, Oglesbee (2008) showed that Japanese listeners tended to hear more voiceless stops when presented with high-falling initial F0 contours compared to manipulated overall F0 levels. That is, the change in F0 in the first 100 ms of the following vowel could be another factor affecting stop perception. Future research examining these kinds of potential variables will enable us to gain a more comprehensive account of the relationship between segment and tone in general as well as a more fine-grained understanding of the Korean stop.

Acknowledgments

The editor Kenneth de Jong and three anonymous reviewers provided us insightful comments and suggestions that helped us improve the quality of this work. We would like to thank Joan Sereno and Jie Zhang for their helpful comments and critique at all stages of this work. We are grateful to Unha Kim who helped us collect the data in Korea. Our thanks also go to our fellow participants in the Empirical Research seminar in the KU Linguistics Department for their valuable feedback and encouragement. We thank the University of Kansas dissertation fellowship, and KU linguistics department for funding support. An earlier version of this work was presented at the 164th meeting of the Acoustical Society of America in Kansas City, Missouri.

References

Abramson, AS.; Lisker, L. Relative power of cues: F0 shift versus voice timing. In: Fromkin, VA., editor. Phonetic Linguistics: essays in honor of Peter Ladefoged. New York: Academic Press; 1985. p. 25-33.

Beddor PS, Strange W. Cross-language study of perception of the oral-nasal distinction. Journal of the Acoustical Society of America. 1982; 71:1551–1561. [PubMed: 7108030]

Benkí JR. Place of articulation and first formant transition pattern both affect perception of voicing in English. Journal of Phonetics. 2001; 29:1–22.

Best CT, Strange W. Effects of language-specific phonological and phonetic factors on cross-language perception of approximants. Journal of Phonetics. 1992; 20:305–30.

Boersma, Paul; Weenink, David. Praat: Doing phonetics by computer (Version 4.3.22). 2007. [Computer software]. Available from http://www.Praat.org/

Chen Y, Robb MP, Gilbert HR, Lerman JW. A study of sentence stress production in Mandarin speakers of American English. Journal of the Acoustical Society of America. 2001; 109:1681–1690. [PubMed: 11325136]

Cho, T. MA Thesis. University of Texas; Arlington: 1996. Vowel correlates to consonant phonation: An acoustic-perceptual study of Korean obstruents.

Cho T, Jun SA, Ladefoged P. Acoustic and aerodynamic correlates of Korean stops and fricatives. Journal of Phonetics. 2002; 30:198–228.

Francis AL, Nusbaum HC. Selective attention and the acquisition of new phonetic categories. Journal of Experimental Psychology: Human Perception and Performance. 2002; 28:349–366. [PubMed: 11999859]

Francis AL, Ciocca V, Wong VKM, Chan JKL. Is fundamental frequency a cue to aspiration in initial stops? Journal of the Acoustical Society of America. 2006; 120:2884–2895. [PubMed: 17139746]

Gottfried TL. Effects of consonant context on the perception of French vowels. Journal of Phonetics. 1984; 12:91–114.

Gandour J. Consonant types and tone in Siamese. Journal of Phonetics. 1974; 2:337–350.

Halle M, Stevens KN. A note on laryngeal features. Quarterly Progress Report (Research Laboratory of Electronics, MIT). 1971; 101:198–212.

Han MS, Weitzman RS. Acoustic features of Korean /P, T, K/, /p, t, k/ and /p^h, t^h, k^h/. Phonetica. 1970; 22:112–128.

Hombert, JM. Consonant types, vowel quality, and tone. In: Fromkin, VA., editor. Tone: A Linguistic Survey. New York: Academic Press; 1978. p. 77-111.

Kang KH, Guion SG. Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. Journal of the Acoustical Society of America. 2008; 124:3909–3917. [PubMed: 19206816]

Kenstowicz M, Park C. Laryngeal features and tone in Kyungsang Korean: A phonetic study. Studies in Phonetics, Phonology and Morphology. 2006; 12:247–264.

Kenstowicz M, Cho H, Kim J. A note on contrasts, mergers and acquisitions in Kyungsang accents. Toronto Working Papers in Linguistics. 2008; 28:107–122.

Kim CW. On the autonomy of the tensity feature in stop classification. Word. 1965; 21:339–359.

Kim, M-R. PhD Dissertation. University of Wisconsin; Madison: 1994. Acoustic characteristics of Korean stops and perception of English stop consonants.

Kim MR, Beddor PS, Horrocks J. The contribution of consonantal and vocalic information to the perception of Korean initial stops. Journal of Phonetics. 2002; 30:77–100.

Kim MR, Duanmu S. "Tense" and "lax" stops in Korean. Journal of East Asian Linguistics. 2004; 13:59–104.

Kim, M. Correlation between VOT and F0 in the perception of Korean stops and affricates. Proceedings of the 8th International Conference on Spoken Language Processing; 2004. p. 49-52.

Kingston J, Diehl RL. Phonetic knowledge. Language. 1994; 70:419–494.

Kingston, J.; Diehl, RL. Intermediate properties in the perception of distinctive feature values. In: Connell, B.; Arvaniti, A., editors. Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV. Cambridge: Cambridge University Press; 1995. p. 7-27.

Lai, Yuwen. MA Thesis. University of Kansas; 2004. The effect of aspiration on the fundamental frequency of the following tone in Taiwanese.

Lee H, Jongman A. 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. 2012; 42:145–169.

Lisker L, Abramson AS. Cross-language study of voicing in initial stops: Acoustical measurements. Word. 1964; 20:384–422.

Lisker L. "Voicing" in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees. Language and Speech. 1986; 29:3–11. [PubMed: 3657346]

Lorch RF, Myers JL. Regression analyses of repeated measures data in cognitive research. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1990; 16:149–157.

Löfqvist A, Baer T, McGarr NS, Seider Story R. The cricothyroid muscle in voicing control. Journal of the Acoustical Society of America. 1989; 85:1314–1321. [PubMed: 2708673]

Miyawaki K, Strange W, Verbrugge RR, Liberman AM, Jenkins JJ, Fujimura O. An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. Perception & Psychophysics. 1975; 18:331–340.

Oglesbee, EN. PhD Dissertation. Indiana University; 2008. Multidimensional stop categorization in English, Spanish, Korean, Japanese, and Canadian French.

Parker EM, Diehl RL, Kluender KR. Trading relations in speech and nonspeech. Perception & Psychophysics. 1986; 39:129–142. [PubMed: 3725537]

Perkins, J.; Lee, SJ. Korean Affricates and Consonant-Tone Interaction. Proceedings of the 6th Workshop on Altaic Formal Linguistics. MIT Working Paper of Linguistics; 2010. p. 277-286.

Ramsey, SR. PhD Dissertation. Yale University; 1975. Accent and morphology in Korean dialects: A descriptive and historical study.

Repp BH. Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. Psychological Bulletin. 1982; 92:81–110. [PubMed: 7134330]

Repp BH. Trading relations among acoustic cues in speech perception are largely a result of phonetic categorization. Speech Communication. 1983; 2:341–361.

Silva DJ. Acoustic evidence for the emergence of tonal contrast in contemporary Korean. Phonology. 2006; 23:287–308.

Summerfield Q, Haggard M. On the dissociation of spectral and temporal cues to the voicing distinction in initial stop consonants. Journal of the Acoustical Society of America. 1977; 62:435–448. [PubMed: 886081]

Strange, W. Cross-language studies of speech perception: A historical review. In: Strange, W., editor. Speech Perception and Linguistic Experience: Issues in cross-language speech research. Timonium, MD: York Press; 1995. p. 3-45.

Tagliaferri, B. Perception Research Systems. KS: University of Kansas; 2005. Paradigm (Version 1.2.1) [Computer software]. Available from http://www.paradigmexperiments.com

Whalen DH, Abramson AS, Lisker L, Mody M. F0 gives voicing information even with unambiguous voice onset times. Journal of the Acoustical Society of America. 1993; 47:36–49.

Williams L. The modification of speech perception and production in second-language learning. Perception & Psychophysics. 1979; 26:95–104. [PubMed: 537858]

Wright, J. PhD Dissertation. University of Pennsylvania; 2007. The phonetic contrast of Korean obstruents.

Xu Y, Gandour JT, Francis AL. Effects of language experience and stimulus complexity on the categorical perception of pitch direction. Journal of the Acoustical Society of America. 2006; 120:1063–1074. [PubMed: 16938992]

Yamada RA, Tohkura Y. The effects of experimental variables on the perception of American English /r/ and /l/ by Japanese listeners. Perception & Psychophysics. 1992; 52:376–392. [PubMed: 1437471]

Yang BG. A comparative study of American English and Korean vowels produced by male and female speakers. Journal of Phonetics. 1996; 24:245–261.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

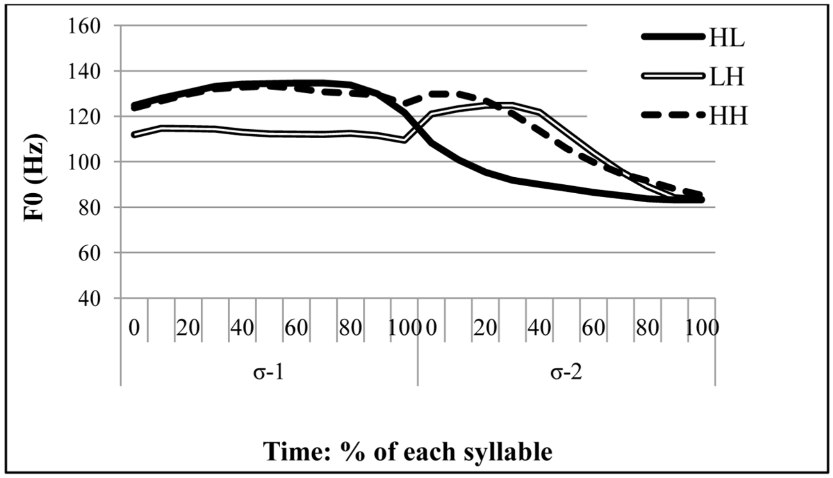

Figure 1.
Pitch tracks of disyllabic HL, LH and HH words. The pitch contours were extracted from a male Kyungsang speaker's production of /kori/ (HL) 'ring', /salam/ (LH) 'person' and /moki/ (HH) 'mosquito'.

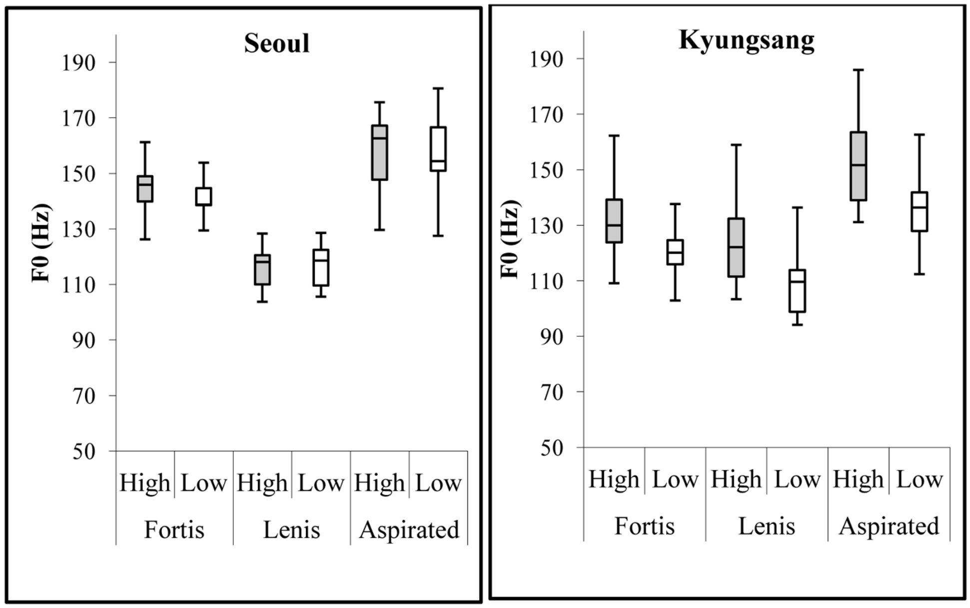

Figure 2.
F0 distribution for Seoul and Kyungsang Korean as a function of laryngeal type and (Kyungsang) pitch accent (adapted from Lee & Jongman, 2012).

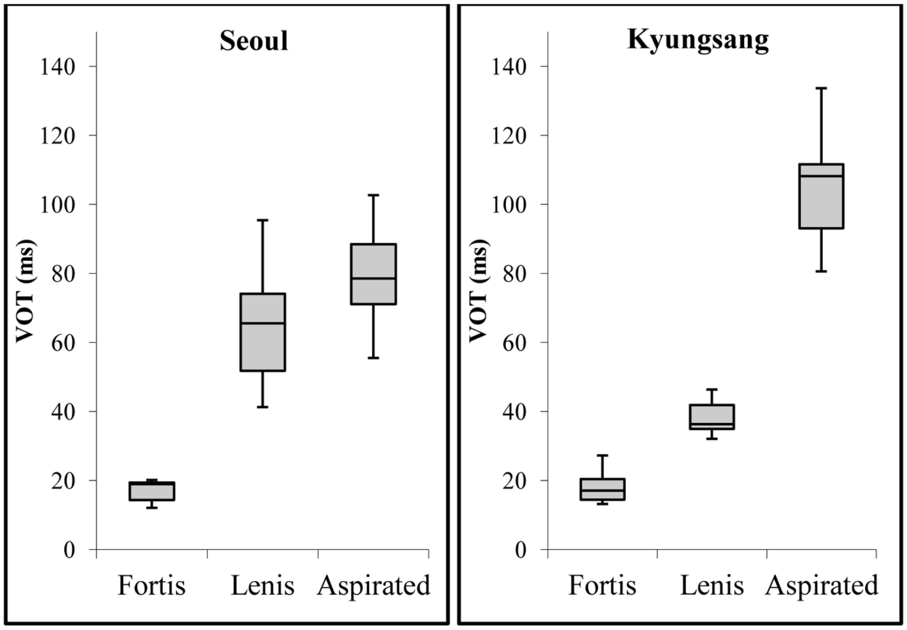

Figure 3.
VOT distribution for Seoul and Kyungsang Korean as a function of laryngeal type (adapted from Lee & Jongman, 2012).

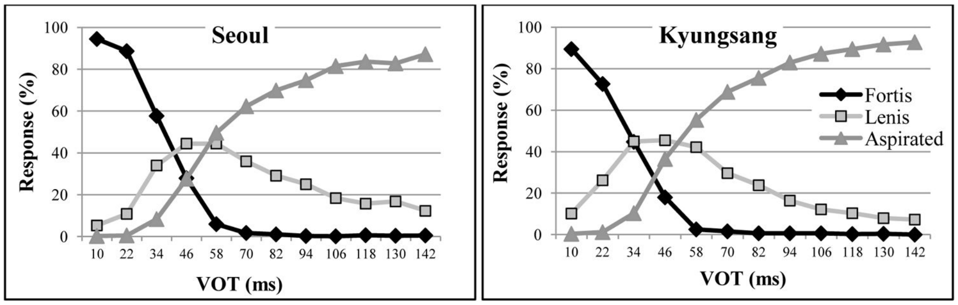

Figure 4.
Identification rate (%) for the three-way stop distinction as a function of VOT for Seoul (left) and Kyungsang (right), pooled across F0 values.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

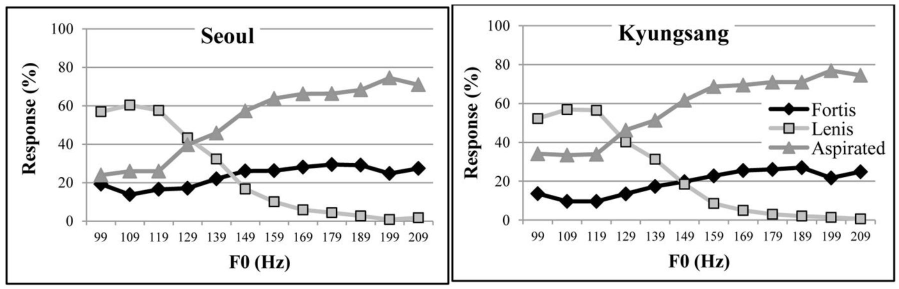

Figure 5.
Identification rate (%) for the three-way stop distinction as a function of F0 for Seoul (left) and Kyungsang (right), pooled across VOT values.

Identification %	0 – 19	20 – 39	40 – 59	60 – 79	80 – 100
Color					

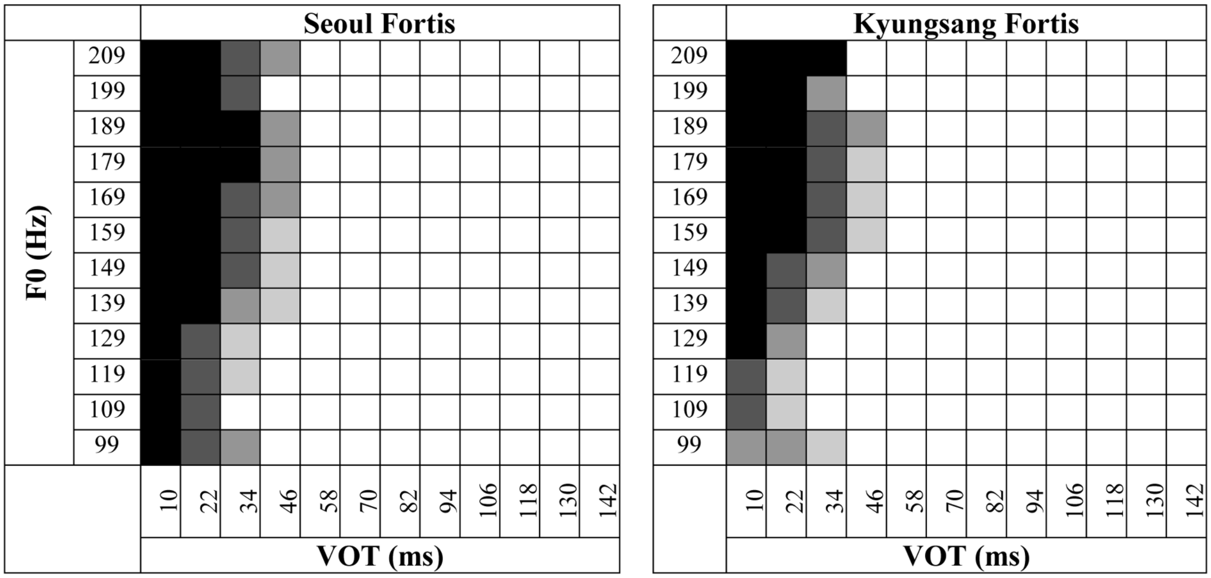

Figure 6.

Identification rate (%) for the fortis stop for each listener group as a function of VOT and F0.

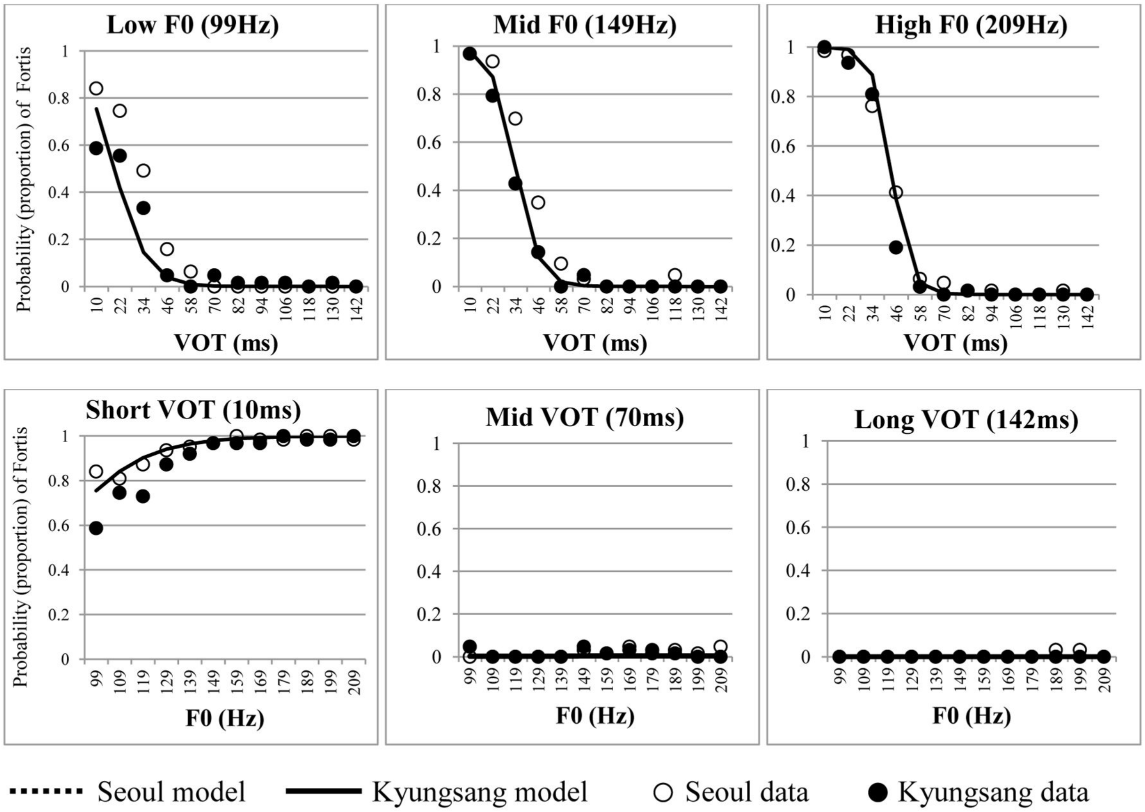

Figure 7.

Regression curves from averaged coefficients in the population-averaged model for Fortis as a function of VOT at low, mid and high F0 (top) and F0 at short, mid, and long VOT (bottom). The curve labelled “Kyungsang model” is based on coefficients averaged across all subjects from both dialects, because the differences between dialects were not significant.

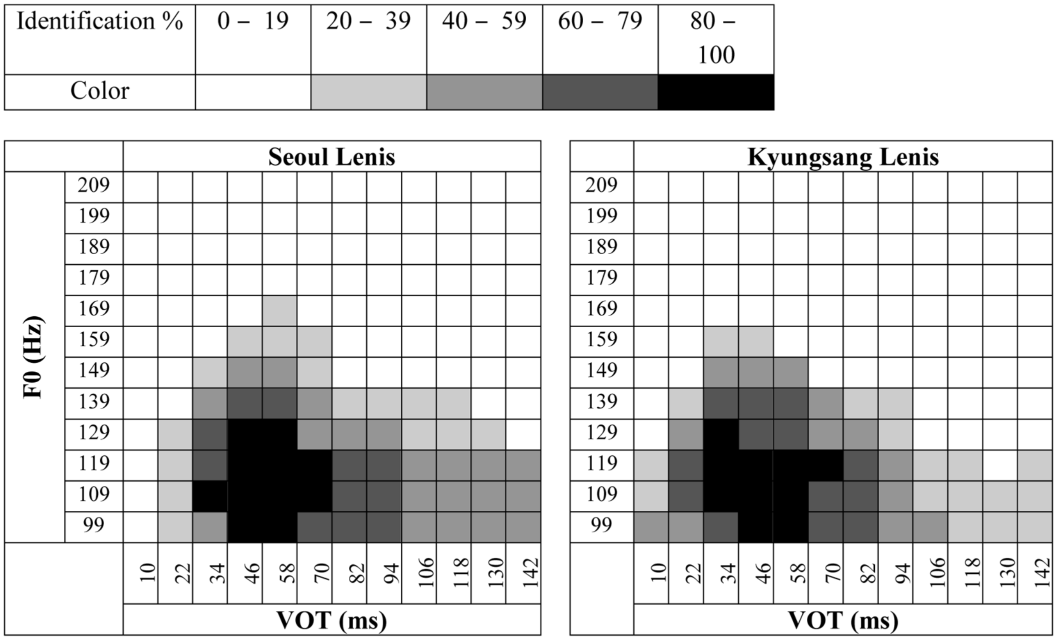

Figure 8.
Identification rate (%) for the lenis stop for each listener group as a function of VOT and F0.

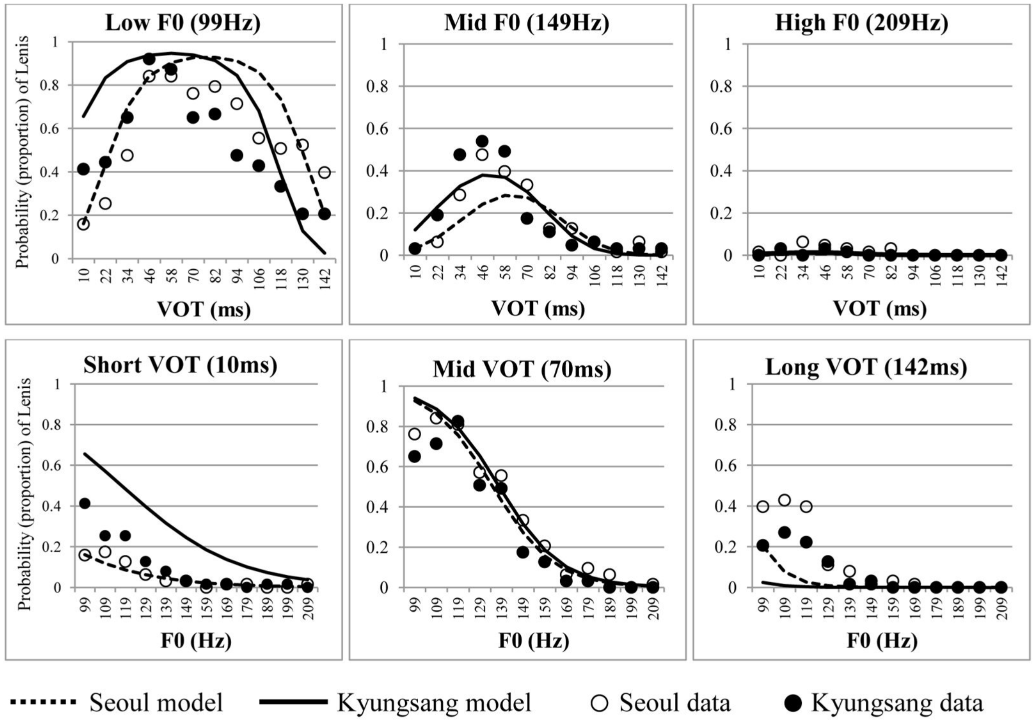

Figure 9.

Regression curves from averaged coefficients in the population-averaged model for Lenis as a function of VOT at low, mid and high F0 (top) and F0 at short, mid, and long VOT (bottom).

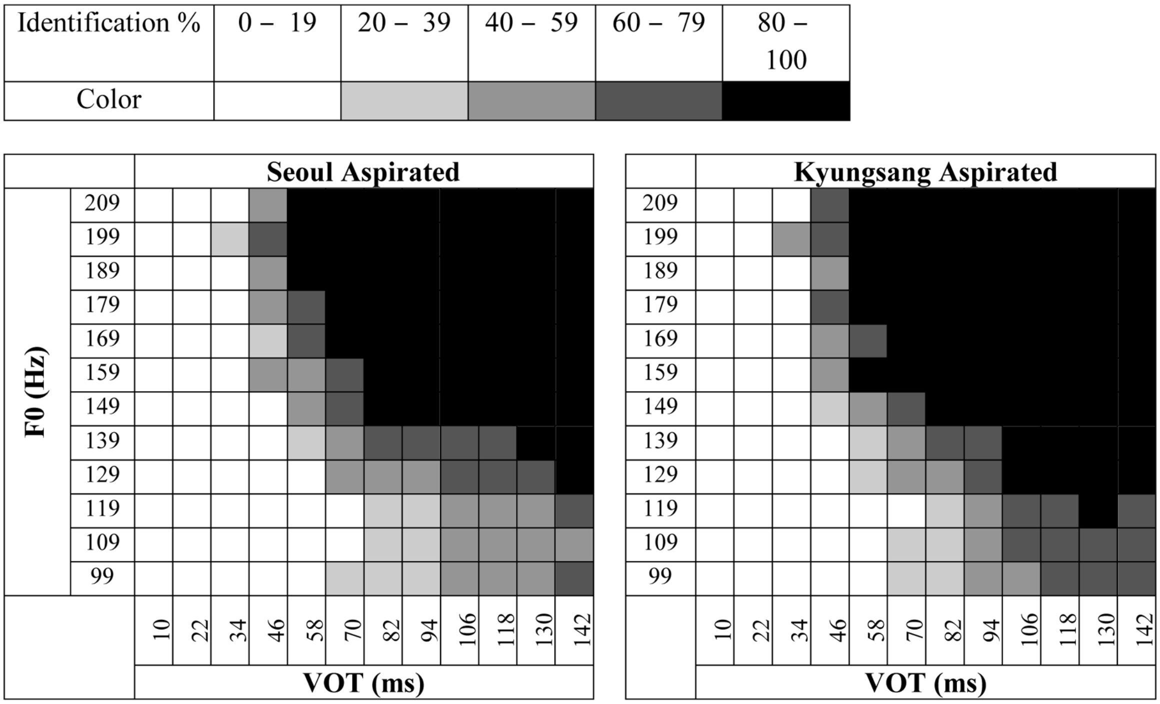

Figure 10.
Identification rate (%) for the aspirated stop for each listener group as a function of VOT and F0.

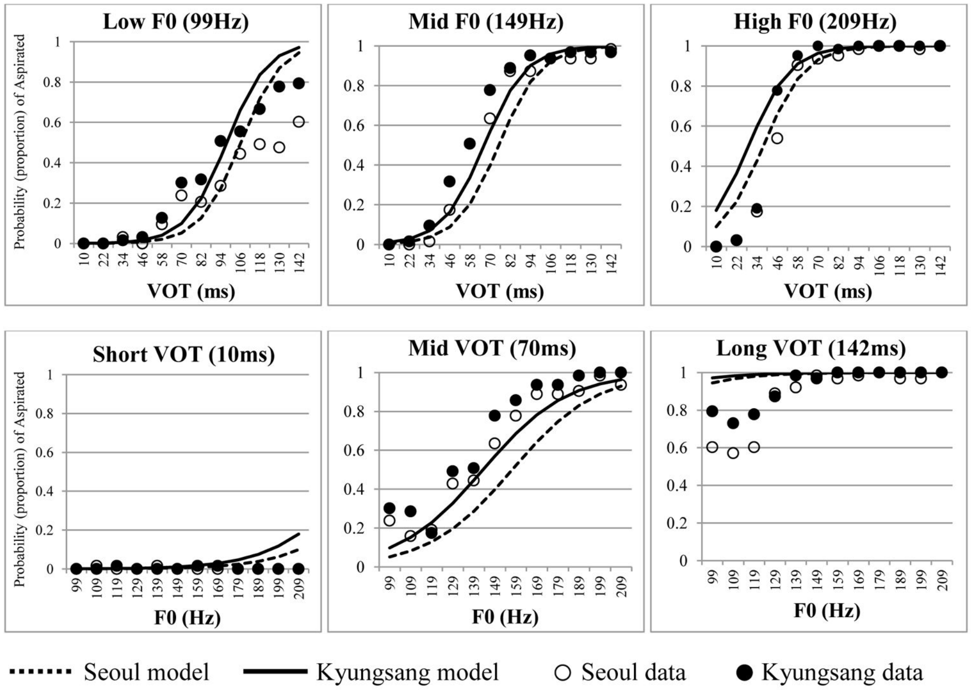

Figure 11.

Regression curves from averaged coefficients in the population-averaged model for Aspirated as a function of VOT at low, mid and high F0 (top) and F0 at short, mid, and long VOT (bottom).

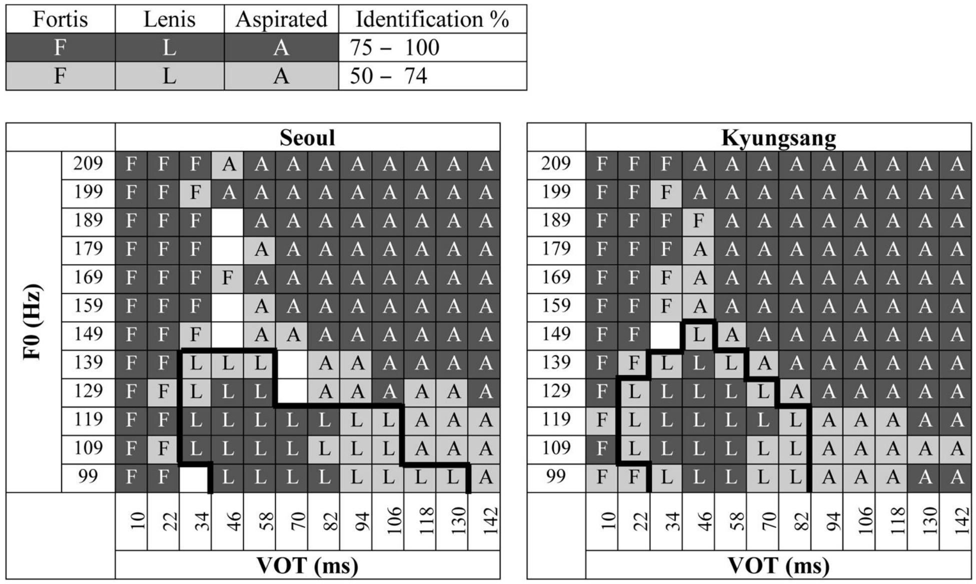

Figure 12.

Identification rate (%) between Seoul and Kyungsang Korean for the fortis, lenis and aspirated stops as a function of VOT and F0.

Table 1

Population-averaged logistic regression model for Fortis across Seoul and Kyungsang groups by one-sample *t*-test.

	β	t (df = 41)	Sig. (2-tailed)	95% CI (lower, upper)	
Intercept	−6.8324	−11.231	0.000	−8.0610	−5.6038
VOT	−0.1661	−12.810**	0.000	−0.1923	−0.1399
F0	−0.0030	−0.467	0.643	−0.0158	0.0099
VOT by F0	−0.0008	−5.767**	0.000	−0.0011	−0.0005

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 2

Population-averaged logistic regression model for Fortis between Seoul and Kyungsang groups by independent-samples *t*-test.

	β (Seoul, Kyungsang)		Mean difference	t (df = 40)	Sig. (2-tailed)	95% CI (lower, upper)	
Intercept	−7.1717	−6.4932	0.6785	0.553	0.583	−1.8016	3.1587
VOT	−0.1821	−0.1502	0.0318	1.237	0.223	−0.0202	0.0839
F0	0.0028	−0.0088	−0.0115	−0.910	0.368	−0.0373	0.0141
VOT by F0	−0.0006	−0.0010	−0.004	−1.531	0.134	−0.0010	0.0001

Table 3

Population-averaged logistic regression model for Lenis across Seoul and Kyungsang groups by one-sample *t*-test.

	β	t (df = 41)	Sig. (2-tailed)	95% CI (lower, upper)	
Intercept	−1.4796	−9.565	0.000	−1.7920	−1.1672
VOT	−0.0402	−7.916**	0.000	−0.0504	−0.0299
F0	−0.0739	−15.433**	0.000	−0.0836	−0.0642
VOT^2	−0.0009	−12.537**	0.000	−0.0011	−0.0008
VOT by F0	−0.0005	−6.404**	0.000	−0.0006	−0.0003

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 4

Population-averaged logistic regression model for Lenis between Seoul and Kyungsang groups by independent-samples *t*-test.

	β (Seoul, Kyungsang)		Mean difference	t (df = 40)	Sig. (2-tailed)	95% CI (lower, upper)	
Intercept	−1.3046	−1.6546	0.3500	1.135	0.263	−0.2730	0.9731
VOT	−0.0299	−0.0504	0.0206*	2.110	0.041	0.0009	0.0403
F0	−0.0751	−0.0727	−0.0024	−0.243	0.809	−0.0219	0.0172
VOT^2	−0.0009	−0.0010	0.0001	0.318	0.752	−0.0003	0.0004
VOT by F0	−0.0006	−0.0003	−0.0003*	−1.902	0.064	−0.0005	0.0000

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 5

Population-averaged logistic regression model for Aspirated across Seoul and Kyungsang groups by one-sample *t*-test.

	β	t (df = 41)	Sig. (2-tailed)	95% CI (lower, upper)	
Intercept	1.5142	7.581	0.000	1.1108	1.9176
VOT	0.0942	20.949**	0.000	0.0851	0.1033
F0	0.0689	23.474**	0.000	0.0630	0.0749
VOT by F0	0.0009	14.571**	0.000	0.0008	0.0010

Table 6

Population-averaged logistic regression model for Aspirated between Seoul and Kyungsang groups by independent-samples *t*-test.

	β (Seoul, Kyungsang)		Mean difference	t (df = 40)	Sig. (2-tailed)	95% CI (lower, upper)	
Intercept	0.3242	1.0184	0.6941*	2.141	0.038	0.0389	1.3494
VOT	0.0733	0.0830	0.0097	1.433	0.160	−0.0040	0.0235
F0	0.0545	0.0533	−0.0012	−0.229	0.820	−0.0119	0.0095