The perception of Seoul Korean fricatives by listeners from five different native dialect and language groups

Jeffrey J. Holliday
Indiana University

Previous studies have investigated the acoustic properties of Korean sibilant fricatives, with some attention given to effects of native dialect and language on both their production and perception. The current study investigates the effects of native dialect and language on the perception of Seoul Korean fricatives by testing the identification of fricative-initial CVs by native Korean speakers from Seoul, Daegu, and Jeju, as well as native Mandarin and Japanese second language learners of Korean. The results show that although native Korean listeners are far more accurate than non-native listeners, there was no significant variation within the native and non-native groups themselves. The results also show an inverse relationship between identification accuracy and vowel height that was consistent across both native and non-native listeners. This finding is in line with previous studies showing that the cues to the contrast are stronger in low vowel contexts than in high vowel contexts.

Keywords: fricatives, perception, dialects, second language acquisition

1. Introduction

The aim of this paper is to investigate the effect of native dialect and language on the identification of Korean sibilant fricatives. In recent years there have been several thorough studies of the acoustic properties of Korean sibilant fricatives (Pyo et al. 1999, Yoon 1999, Cho et al. 2002, Chang 2007, Chang 2013, Kim et al. 2010), but fewer studies have investigated how listeners perceive them. The two sibilant fricatives, which will be notated in this paper as /s^h/ and /s*/, form a typologically unusual contrast that not only presents a challenge for second language (L2)

Korean Linguistics 16:2 (2014), **93–110**. DOI 10.1075/kl.16.2.01hol
ISSN 0257–3784 / E-ISSN 2212–9731

learners but is also widely believed to be phonetically neutralized in Gyeongsang[1] varieties of Korean. Accordingly, in this study we compared not only L2 listeners with native (L1) listeners from Seoul, but also with L1 listeners from Daegu and Jeju. In the following sections we will briefly review the phonetic properties of /s^{h}/ and /s*/, claims regarding their dialectal variation, and the results of previous perceptual studies.

1.1 Articulatory, acoustic, and distributional properties

Imaging studies of the laryngeal properties of Korean sibilant fricatives (Kagaya 1974, Kim et al. 2011) have found that glottal width is generally wider during the frication portion of /s^{h}/ than in that of /s*/. After the consonantal portion in both fricatives, including any aspiration, the glottis closes rapidly, and in the case of /s*/ the vowel may even be preceded by a brief glottal closure (Kagaya 1974). Differences between /s^{h}/ and /s*/ in lingual articulation are less clear. While the place of articulation is generally agreed on as denti-alveolar (Anderson et al. 2004), EPG data have suggested that /s*/ may have more linguopalatal contact (S. Kim 2001), a narrower groove width, and a smaller front cavity (Baik 1998) than /s^{h}/.

Several acoustic cues to the contrast in word-initial position have been suggested. First, the duration of frication noise has been shown to be longer in /s*/ than in /s^{h}/ (Yoon 1999, Cho et al. 2002, Chang 2007, Kim et al. 2010, Holliday 2012, Chang 2013), although this difference was greatly reduced before /i/ and /u/. The frication period is generally followed by a period of aspiration, although it has been shown that the duration of aspiration is much longer in /s^{h}/ than in /s*/ (Kagaya 1974, Yoon 1999, Cho et al. 2002, Chang 2007, Kim et al. 2010, Holliday 2012, Chang 2013). As in frication duration, the difference in aspiration duration between /s^{h}/ and /s*/ is much smaller in high vowel contexts, and especially before /i/.

Second, differences in tongue posture and front cavity size between /s^{h}/ and /s*/, as reported in Baik (1998) and S. Kim (2001), suggest that centroid or spectral peak frequency could cue the contrast in a way similar to the English /s/-/ʃ/ contrast (e.g., Jongman et al. 2000). Yoon (1999) reported that although the peak frequency was usually higher in /s*/ than in /s^{h}/ when averaged across 20 speakers, the individual data did not present a consistent pattern. Cho et al. (2002) found

1. Place names in this paper have been transliterated using Revised Romanization (RR). The Yale Romanization (Y) and McCune-Reischauer Romanization (MR) for the place names used in this paper are as follows: Busan (RR), Pwusan (Y), Pusan (MR); Daegu (RR), Taykwu (Y), Taegu (MR); Gyeongju (RR), Kyengcwu (Y), Kyŏngju (MR); Gyeongsang (RR), Kyengsang (Y), Kyŏngsang (MR); Jeju (RR), Ceycwu (Y), Cheju (MR); Nakdong (RR), Naktong (Y), Naktong (MR); Uljin (RR), Wulcin (Y), Uljin (MR); Yeongcheon (RR), Yengchen (Y), Yŏngch'ŏn (MR).

that the centroid frequency in /s*/ was higher than in /s^{h}/ (approximately 6600 Hz versus 6200 Hz), but their data were limited in that they only reflect the productions of four speakers, and only in the /a/ environment. Other studies, however, have shown a consistent difference between /s^{h}/ and /s*/ in the centroid frequency of the frication noise (Chang 2013, Lee 2013).

Lastly, the different laryngeal settings in /s^{h}/ and /s*/ may also trigger spectral differences in the following vowel. Such spectral differences have typically been captured by spectral tilt, which is the degree to which the amplitudes of progressively higher harmonics taper off in the spectrum (see D. Klatt & L. Klatt 1990, Hanson 1997), and can be quantified as the difference in amplitude between the first (H1) and second (H2) harmonics. Cho et al. (2002) found that /s^{h}/ consistently had a higher H1-H2 than /s*/, suggesting a more breathy phonation in /s^{h}/ and a more pressed voice quality in /s*/. Chang (2007) also found that H1-H2 was much higher in /s^{h}a/ than in /s*a/, but that the difference was greatly reduced in /s^{h}u/ and /s*u/.

Both /s^{h}/ and /s*/ may occur in any vowel context, but /s^{h}/ (and sometimes /s*/) is typically palatalized before /i/ or /j/ (Martin 1992). Because palatalization is conditioned by the following vowel, although [s^{h}] and [ɕ^{h}] (and [s*] and [ɕ*]) are acoustically distinct, many L1 Korean speakers consider them to be vowel contrasts and not consonant contrasts. Lastly, /s^{h}/ and /s*/ are neutralized in coda position, with both being realized as unreleased [t̚].

1.2 Dialectal variation

It is widely believed that Gyeongsang varieties of Korean lack the word-initial /s^{h}/-/s*/ contrast due to /s*/ being acoustically realized as [s^{h}]. Although some sources report this neutralization without any accompanying empirical evidence or citations to support the claim (Martin 1992: 28, Lee & Ramsey 2000: 63), studies that include empirical evidence offer a nuanced, if not ambiguous, account. Drawing on fieldwork transcriptions, Gim (1963: 61–62) claimed that speakers in the southern and western parts of South Gyeongsang province maintained the /s^{h}/-/s*/ contrast, and that only speakers in the northeastern part of the province, roughly northeast of the Nakdong River, neutralized the contrast. Nearly 30 years later, T. Kim (1991: 16–21), conducted more fieldwork and reached the same conclusion.

Although T. Kim (1991) only surveyed speakers from South Gyeongsang province, he hypothesized that this /s^{h}/-/s*/ neutralization isogloss might run along the historic Gyeongsang left-right boundary, a line that ran along the Nakdong River and separated Gyeongsang province into "left" and "right" regions before the modern distinction between North and South Gyeongsang came into

being. If the Left-Right boundary did indeed separate the neutralizing and non-neutralizing dialects, the current North Gyeongsang cities of Daegu, Gyeongju, and Yeongcheon should be in the /s^{h}/-/s*/ neutralizing region (Kim 1991:21). It is not clear how far this neutralizing region would have stretched, however, as Y. Kim and B. Kim (2013) reported on two elderly speakers from Uljin, the most northeastern county of North Gyeongsang province, who did maintain the /s^{h}/-/s*/ contrast. Further complicating our interpretation of these findings, none of the aforementioned studies employed any sort of acoustic analysis and were based entirely on impressionistic transcriptions. Because /s^{h}/-/s*/ neutralization is such a salient stereotype of Gyeongsang speech, transcription data suffers from potential bias on the part of the transcriber.

More recent studies of the /s^{h}/-/s*/ contrast in Gyeongsang varieties have used acoustic data. Kenstowicz and Park (2006) analyzed the fricative productions of seven Gyeongsang speakers, with three from near Busan and four from near Daegu. They found that these speakers clearly differentiated between /s^{h}/ and /s*/ as measured by spectral tilt. They did not find a robust f0 difference, but previous acoustic analyses of Seoul fricatives did not find a significant f0 difference between /s^{h}/ and /s*/ anyway (Kagaya 1974, Cho et al. 2002, Chang 2007). Note that although Busan lies in the region claimed to maintain the /s^{h}/-/s*/ contrast (Gim 1963, T. Kim 1991), Daegu is in the region suspected to neutralize the contrast (T. Kim 1991) and so this finding from Kenstowicz and Park (2006) suggests that younger Daegu speakers produce a clear contrast even if older Daegu speakers do not. Holliday (2012) analyzed the fricative productions of 12 speakers from Seoul and 13 speakers from Daegu, all born after 1980, and also found no significant differences between the groups in terms of five different acoustic cues. Lastly, Lee (2013) conducted an acoustic analysis of the fricative productions from 10 younger and 10 older speakers from the South Gyeongsang region, near Busan, and compared them to productions from the same number of older and younger speakers from Seoul. She found that the contrast between the two fricatives was as distinct among the younger Gyeongsang speakers as it was for Seoul speakers, but less distinct among the older Gyeongsang speakers, indicating a generational change from a less distinct to a more distinct contrast. It should be noted, however, that the older Gyeongsang speakers did not exhibit total neutralization of /s^{h}/ and /s*/, as the popular stereotype suggests they would.

One of the few studies to investigate the sibilant fricatives of a non-Seoul, non-Gyeongsang dialect was Cho et al. (2002), who compared the fricative productions of four Seoul speakers and eight Jeju speakers, all of whom were at least 50 years old at the time of recording. While several acoustic correlates of the sibilant fricative contrast were compared between dialects, the only significant difference was that the centroid frequency of frication noise was significantly higher for Seoul

speakers than for Jeju speakers. Because no other dialect differences were found and the number of informants was rather small, it may be premature to conclude that Jeju fricatives differ in any meaningful way from Seoul fricatives. Moreover, due to the advanced age of the speakers, it remains unknown how or whether the fricative productions of young speakers from both dialects differ.

1.3 Perception of Korean fricatives

One of the earliest perceptual studies of Korean fricatives was Yoon (1999). Using synthetically manipulated stimuli, it was found that L1 Korean listeners were sensitive to the period of aspiration between frication and vowel onset. In an identification task, the 20 participants, which included 17 from the Seoul region and 3 from Gyeongsang province, perceived fricative-initial CV tokens with longer periods of aspiration as /s^{h}/, and tokens with less or no aspiration as /s*/. If aspiration is such a salient cue in Korean fricative identification, these results predict that listeners should be less accurate at identifying naturally produced fricatives in high vowel contexts, such as /i/ and /u/, in which the aspiration in /s^{h}/ is greatly reduced or altogether absent.

Chang (2007) also conducted an identification task with synthetically manipulated stimuli, but with several acoustic cues manipulated simultaneously: segmental duration (i.e., frication plus aspiration), aspiration duration, f0 onset of the following vowel, and a combined cue of F1 onset, intensity buildup, and spectral tilt. The participants, who consisted mostly of listeners from the Seoul region but included some listeners from other non-Seoul, non-Gyeongsang regions, were asked to identify each token as either /s^{h}/ or /s*/ in both the /a/ and /u/ vowel contexts. Overall, it was found that listeners paid most attention to aspiration duration and vocalic cues, such as spectral tilt, but that listeners' reliance on these cues depended on vowel context. For example, listeners were less reliant on spectral tilt in the /u/ context, which could have been because F1 onset was not manipulated in the /u/ context and so listeners had to rely on a different cue, such as aspiration duration. In a later study using a different set of manipulated fricative-initial CV stimuli, Chang (2013) again found that vocalic cues played a significant role in fricative perception but that such cues are weaker in high vowel contexts, such as /i/ and /u/, in which listeners must rely more on fricative-internal cues.

This relative weighting of vocalic cues over consonantal cues was also found by Lee and Jongman (2012), who tested identification with naturally produced fricative-initial CVs, the isolated fricative and vowel portions of naturally produced CVs, and fricative-vowel cross-spliced CV stimuli. In the perception of the naturally produced stimuli it was found that listeners were more accurate in the

/a/ context than in the /i/ context (roughly 100% vs. 75% accuracy, respectively), as predicted they would be, by Yoon (1999) and Chang (2007).

In one of the few studies of the perception of Korean fricatives by L2 listeners, Cheon (2005) tested the ability of naïve L1 Korean and English listeners to discriminate between Korean /s^{h}/ and /s*/ and English /s/ and /ʃ/ in the vowel contexts /a/ and /i/. While the L1 Korean listeners discriminated between Korean /s^{h}a/ and /s*a/ with 100% accuracy, the L1 English listeners only performed with 60% accuracy. The /i/ context was more difficult, with L1 Korean listeners at 95% accuracy and L1 English listeners at 36% accuracy.

Lastly, Holliday (2013, 2014) tested the perceptual assimilation of Korean obstruents, including the sibilant fricatives, by naïve L1 Mandarin and L1 Japanese listeners. It was found that although L1 Japanese listeners consistently perceived both /s^{h}/ and /s*/ as the same Japanese category (/s/ before /a, u/ and /ɕ/ before /i/), L1 Mandarin listeners perceived them as either an affricate or a sibilant fricative, depending on vowel context. Before /a/, L1 Mandarin listeners perceived /s^{h}/ as an affricate 73% of the time and /s*/ almost always as a sibilant fricative. Before /i/, /s^{h}/ was almost always perceived as a sibilant fricative and /s*/ was perceived as an affricate 40% of the time. Before /u/, however, both /s^{h}/ and /s*/ were perceived as sibilant fricatives 80% and 96% of the time, respectively. Although perceptual assimilation by naïve listeners is substantially different from identification by L2 learners, the fact that naïve L1 Mandarin listeners (but not L1 Japanese listeners) perceived a difference between /s^{h}/ and /s*/ in the /a/ and /i/ contexts suggests that L1 Mandarin learners of Korean may be more accurate than L1 Japanese learners at identifying Korean sibilant fricatives.

1.4 The current study

There are three main motivations for the current study. First, Yoon (1999), Chang (2007, 2013), and Lee and Jongman (2012) all tested L1 Korean listeners' identification of Korean /s^{h}/ and /s*/, but, with the exception of part of one experiment in Lee and Jongman (2012), these studies all used synthetically manipulated stimuli. Cheon (2005) used all naturally produced stimuli, but the participants in that study were only asked to discriminate between fricatives, which is ostensibly easier than identifying them. It thus remains to be seen how well L1 Korean listeners can identify naturally produced sibilant fricatives across different vowel contexts.

Second, while some older Gyeongsang speakers may not differentiate between /s^{h}/ and /s*/ (Gim 1963, T. Kim 1991), more recent experimental studies have shown that younger Gyeongsang speakers maintain a robust /s^{h}/-/s*/ contrast (Kenstowicz & Park 2006, Holliday 2012, Lee 2013). Although the results of these studies suggest that Gyeongsang speakers should have no difficulty identifying

Seoul Korean fricatives, it has never been explicitly tested. For other dialect regions, such as Jeju, there has been little investigation of production and no investigation of perception. An investigation of Gyeongsang and Jeju listeners' perception of Seoul fricatives would thus contribute to our understanding of Korean dialectal variation.

Third, the results of Holliday (2013, 2014) suggest that L1 Mandarin and Japanese learners of Korean may perceive /s^{h}/ and /s*/ differently. Specifically, the results predict that L1 Mandarin listeners should perceive the difference between /s^{h}/ and /s*/ more clearly in the /a/ context than in other contexts, and may be altogether more accurate than L1 Japanese listeners.

Accordingly, the current study will investigate the perception of Seoul Korean fricatives by L1 listeners from Seoul, Daegu, and Jeju, as well as native Mandarin and Japanese L2 learners of Korean. Because identifying members of a contrast is generally harder than discriminating between them, and because perception of CVs excised from words is also generally harder than CVs produced in isolation, the current study will test the identification of CVs excised from words. This decision was made to avoid ceiling effects that could wash out differences between listener groups.

For the L1 listeners, it is predicted that the Daegu listeners should be as accurate as the Seoul listeners. While Jeju listeners are not predicted to be any less accurate than Seoul or Daegu listeners, their presence in the current study serves as a control, since if Daegu listeners are less accurate than Seoul listeners it would be unclear whether it was due to the purported neutralization in their native dialect or simply due to the stimuli being from a different dialect (i.e., whether it is a "Gyeongsang effect" or a "non-Seoul effect"). For the L2 listeners, it is predicted that the native Mandarin listeners should be more accurate than the native Japanese listeners, although perhaps not in the /u/ context.

2. Methods

2.1 Participants

Data[2] were collected from a total of 100 participants, consisting of 20 participants from each of five different groups: L1 Korean speakers from Seoul, Daegu, and Jeju, and native speakers of Mandarin and Japanese.

2. The data from the Seoul, Mandarin, and Japanese listeners were included as part of the author's Ph.D. dissertation in 2012. The remaining data, from the Daegu and Jeju listeners, have not been presented elsewhere.

Native Korean participants
The Seoul listeners were either born and raised in Seoul or had moved to Seoul from a different region at a young age. Their year of birth ranged from 1975 to 1989 with a median of 1981. The Daegu and Jeju listeners were all born and raised in their respective cities. The Daegu listeners were born between 1985 and 1991 with a median year of birth of 1990, and the Jeju listeners were born between 1984 and 1992 with a median of 1990. All of the native participants were recruited and tested in Seoul, Daegu, or Jeju.

Non-native participants
The non-native participants were all novice L2 learners enrolled in intensive Korean language programs at one of three universities in Seoul, South Korea. The intensive Korean language programs at these schools all consist of six levels, with Level 1 assuming no knowledge of Korean whatsoever and graduates of Level 6 being functionally fluent. Students receive classroom instruction for 4 hours a day, 5 days a week, with each level lasting 10 weeks. The number of participants from each level was balanced across native language groups, with 14 from Level 1, five from Level 2, and one from Level 3. Each native language group thus comprised L2 proficiency levels ranging from very novice to low intermediate. All of the non-native participants were recruited and tested in Seoul.

2.2 Stimuli

Participants completed a two-alternative forced choice perception task consisting of two parts: a minimal pair test and a CV Test.

Minimal pair (MP) test
The purpose of the MP test was to familiarize the participants with the experimental format and to test whether listeners could accurately identify the Korean fricatives in the context of known words, which we hypothesized would be easier than in isolated CVs. The stimuli consisted of six Korean words produced by three native Korean speakers (1 female and 2 male), yielding a total of 18 stimuli. The words were the minimal pairs /s^{h}a.ta/ 'buy' vs. /s*a.ta/ 'cheap', /s^{h}ʌt.ta/ 'stood' vs. /s*ʌt.ta/ 'used', and /s^{h}i/ 'poem' vs. /s*i/ 'seed'.

CV test
The CV stimuli consisted of 324 word-initial /s^{h}/- or /s*/-initial syllables that were excised from 18 real-word productions of 9 male and 9 female native Seoul Korean speakers produced in a picture-prompted word repetition task (Kong et al. 2011). The CVs were spread across five different vowel contexts, with the number of the

CVs from each context shown in Table 1. The number of correct trials that were needed to be significantly above chance level for each vowel context was calculated using a binomial test (n = number of stimuli for that vowel context, p = .5, α = .05).

Table 1. Number of stimuli by vowel context and percentage of correct trials needed to be above chance level

	/a/	/e/	/o/	/u/	/i/	Total
Number of stimuli	108	72	36	36	72	324
% correct to be above chance	59.3	61.1	66.7	66.7	61.1	54.9

2.3 Procedure

The experiment was run on the experiment presentation software E-Prime (Schneider et al. 2002) on a laptop computer with stimuli presented over headphones. The experiment took place in a quiet room at a location convenient for the participant. Participants completed a production task (the results of which are not discussed here), the perception task, and a language background questionnaire. Participants were paid either cash or a cash-equivalent for their participation.

Minimal Pair Test

Participants heard the 18 stimuli one at a time, with each stimulus accompanied by the target word and its minimal pair presented on the screen in Hangul as visual prompts. The orientation of the prompts (the /s^{h}/-initial word on the right and the /s*/-initial word on the left, or vice versa) was balanced across participants. After hearing each stimulus, the participant indicated his or her choice by pressing either "1" for the word on the left side of the screen or "0" for the word on the right. The stimuli were randomized and presented as a single block, which took participants approximately 1 minute to complete.

CV Test

The CV Test procedure was nearly identical to that of the MP Test. Participants heard the 324 stimuli described above, with each stimulus accompanied by the Hangul characters for /s^{h}/ and /s*/ as visual prompts. The orientation of the prompts (/s^{h}/ on the right and /s*/ on the left, or vice versa) was balanced across participants, but the orientation for the MP Test and CV Test were always the same for each participant. After hearing each stimulus, the participant indicated his or her choice by pressing a button on the keyboard: "1" for the fricative on the left side of the screen and "0" for the fricative on the right. The stimuli were randomized and presented in 3 blocks of 108 tokens each, with optional breaks between the blocks. The CV test lasted between 20 and 30 minutes for most participants.

3. Results

3.1 Minimal Pair Test

The identification accuracy of each listener group in the MP Test is shown in Table 2. There was both a clear difference between the native and non-native listener groups, with native listeners being nearly at ceiling and non-native listeners being near chance level, and also very little variation among native and non-native groups themselves. That is, a listener's identification accuracy seemed to depend more on native Korean status rather than the listener's particular native dialect or language.

Table 2. Mean and standard deviation of by-listener identification accuracy for the MP Test

	Seoul	Daegu	Jeju	Mandarin	Japanese
/a/	97.5 (15.7)	98.3 (12.9)	99.2 (9.1)	63.3 (48.3)	69.2 (46.4)
/ʌ/	96.7 (18.0)	98.3 (12.9)	98.3 (12.9)	52.5 (50.1)	47.5 (50.1)
/i/	88.3 (32.2)	90.8 (30.0)	90.0 (30.1)	59.2 (49.4)	52.5 (50.1)
Overall	94.2 (23.5)	95.8 (20.0)	95.8 (20.0)	58.3 (49.4)	56.4 (49.7)

This trend was confirmed through a repeated measures ANOVA with a between-subjects factor of listener group and a within-subjects factor of vowel context. There were significant main effects of listener group ($F(4,95) = 103.9$, $p < .001$) and vowel context ($F(2,190) = 12.37$, $p < .001$), as well as a significant interaction between the two factors ($F(8,190) = 2.97$, $p = .004$). Because the differences between the L1 and L2 listeners are so clear, instead of testing between individual listener groups we ran two more repeated measures ANOVAs with the same structure as above, with one for only the L1 groups and the other for only the L2 groups. The L1 model showed a significant main effect of vowel context ($F(2,114) = 20.02$, $p < .001$) and no significant effect of listener group ($p = .483$) or interaction ($p = .992$). The L2 model also showed a significant effect of vowel context ($F(2,76) = 7.38$, $p = .001$) and no significant effect of listener group ($p = .64$) or interaction ($p = .291$). The number of stimuli in the MP Test was quite small, however, and investigating the effect of vowel context further was not expected to be meaningful. Accordingly, we now turn to the results of the CV Test, from which we can draw more concrete conclusions about the effect of vowel context.

3.2 CV Test

The results of the CV Test, just as those of the MP Test, show a clear difference between L1 and L2 listeners. The distribution of mean accuracy scores is shown in Figure 1. A repeated measures ANOVA showed main effects of listener group

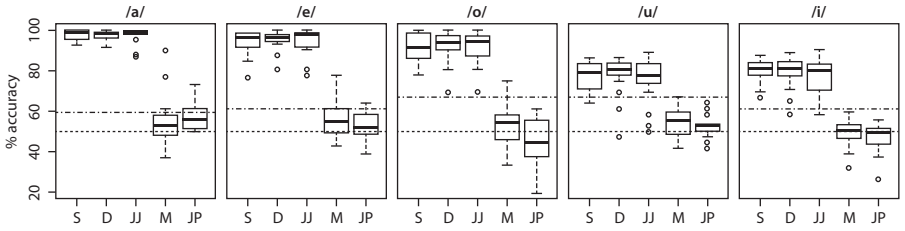

Figure 1. Mean accuracy in the CV Test across vowel context and listener group: Seoul (S), Daegu (D), Jeju (JJ), Mandarin (M), and Japanese (JP). The dotted line is at 50% and the dashed line is at the percentage of correct responses needed to be above chance level for that vowel context.

($F(4,95) = 277.4$, $p < .001$), vowel context ($F(4,380) = 110.9$, $p < .001$), and a significant interaction between the two main effects ($F(16,380) = 12.1$, $p < .001$). To confirm these L1/L2 effects we ran two-sided t-tests for each vowel context, comparing the individual mean accuracy of the L1 listeners and the L2 listeners. All comparisons were significant with an alpha value corrected for multiple comparisons ($p < .001$), further demonstrating the clear difference between L1 and L2 listeners shown in Figure 1.

3.2.1 *Accuracy rates*

Having confirmed that the L1 and L2 listeners form completely separate populations, we analyzed the L1 and L2 groups separately. Turning first to the L2 listeners, a repeated measures ANOVA returned a main effect of vowel context ($F(4,152) = 8.62$, $p < .001$) with a significant interaction between vowel context and native language ($F(4,152) = 2.93$, $p = .023$), but no main effect of native language ($p = .112$). The mean accuracy rates across vowel contexts for the Mandarin listeners were /a/ 54.6%, /e/ 55.8%, /o/, 52.9%, /u/ 54.2%, and /i/ 49.5%, and for the Japanese listeners were /a/ 57.6%, /e/ 52.2%, /o/ 44.7%, /u/ 51.7%, and /i/ 46.7%. To investigate whether any of these differences were significant, we ran multiple paired t-tests with a Bonferroni corrected α value of .0015, as we made additional comparisons with the L1 listeners below. None of the vowel context comparisons for the Mandarin listeners were significant, but for the Japanese listeners accuracy was higher in the /a/ context than in the other four vowel contexts. Thus although L2 listeners' identification accuracy varied across vowel contexts, most of these differences were not statistically significant.

A repeated measures ANOVA of just the responses from the L1 listeners also returned a main effect of vowel context ($F(4,228) = 201.04$, $p < .001$) but no main effect of dialect ($p = .826$) and no significant interaction ($p = .76$). Because dialect did not seem to play any role in perception, we pooled the three dialect groups together and compared mean accuracy across vowel contexts: /a/ 97.5%, /e/ 95.0%,

/o/ 91.6%, /u/ 76.9%, and /i/ 79.1%. To investigate whether any of these differences were significant, we pooled the Seoul, Daegu, and Jeju listeners together and ran multiple paired t-tests (again with Bonferroni corrected $\alpha = .0025$). The results indicated that all differences between vowel contexts were significant except for between /i/ and /u/.

3.2.2 *Response bias*

From the mean accuracy results we can conclude that L2 listeners are quite poor at identifying Korean fricatives from isolated CVs, but this finding could be due to either response bias or random guessing. To explore these possibilities we calculated the percentage of /s*/ responses, shown in Figure 2, to see whether some listeners were always choosing the same response. Overall, there seems to be substantial variation both between talkers and across vowel contexts. To test whether there was any effect of native language or vowel context on /s*/ response rate we ran another repeated measures ANOVA, which returned a main effect of vowel context ($F(4,152) = 11.22$, $p < .001$) but not native language ($p = .666$), and no significant interaction ($p = .102$). Paired t-tests indicated ($p < .0025$) that the mean /s*/ response rate for /i/ (63.6%) was higher than for any other vowel, and that the rate for /e/ (52.2%) was also higher than the rate for /o/ (44.1%). It is not clear why the L2 listeners chose /s*/ so much more frequently in the /i/ context.

Judging from Figure 2, the L1 listeners showed overall less variability than the L2 listeners in proportion of /s*/ responses, although it should be noted that this difference is a necessary consequence of the L1 listeners being more accurate. Because the stimuli were balanced between /s^{h}/ and /s*/, the more accurate a listener is the closer the listener's /s*/ response rate should be to 0.5. The L1 listeners were significantly less accurate in the /u/ and /i/ contexts, however, and so there remains some possibility that listeners were biased towards one response. The mean /s*/ response rates across vowel contexts were /a/ 48.7%, /e/ 48.7%, /o/ 45.2%, /u/ 51.6%, and /i/ 42.7%. To test whether this trend was significant we first ran the same ANOVA that was run for the L2 listeners, which also returned a

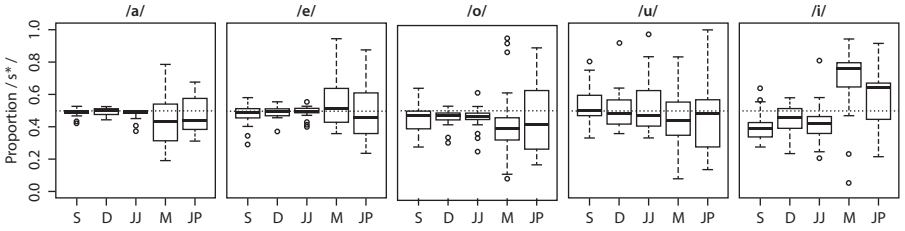

Figure 2. Proportion of /s*/ responses across vowel context and listener group: Seoul (S), Daegu (D), Jeju (JJ), Mandarin (M), and Japanese (JP).

significant effect of vowel context ($F(4,228) = 11.39$, $p < .001$) but no effect of dialect ($p = .83$) and no significant interaction ($p = .936$). Paired t-tests revealed that the /s*/ response rate was significantly higher in /a/, /e/, and /u/ than in /o/ or /i/ ($p < .0025$). No other comparisons were significant. As the /s*/ response rates for /a/ and /e/ were already below 50%, this result suggests that the L1 listeners were slightly biased towards /s^{h}/ for the /o/ and /i/ contexts.

The preference for /s^{h}/ in the /i/ context could reflect an asymmetry between the acoustic correlates of /s*/ in production and perception. Whereas /s*/ has been claimed to palatalize just like /s^{h}/ (Yoon 1999), less frequently than /s^{h}/ (Martin 1992: 28), or not at all (Baik 1998), it has often been claimed that /s^{h}/ always palatalizes before /i/. If listeners believe that palatalization is a cue to /s^{h}/ but not /s*/, then to the extent that any tokens of /s*i/ are palatalized the listeners would exhibit bias towards /s^{h}/. We offer no explanation for why listeners might be biased towards /s^{h}/ in the /o/ context, however.

It may also be illustrative to look at the /s*/-response rates for individual listeners across all stimuli, shown in the top panel of Figure 3, as the number of stimuli was not balanced across vowel contexts. While there were a few Mandarin and Japanese listeners who were heavily biased towards /s*/, most of them were in the .35 to .65 range that all of the L1 Korean listeners fell in. Two-sided Kolmogorov-Smirnov tests revealed no significant differences between the distribution of /s*/ response rates between the L1 listener groups ($p \geq .32$) and between the Mandarin and Japanese groups ($p = .56$). The extremely low (and sometimes even negative) *d'*

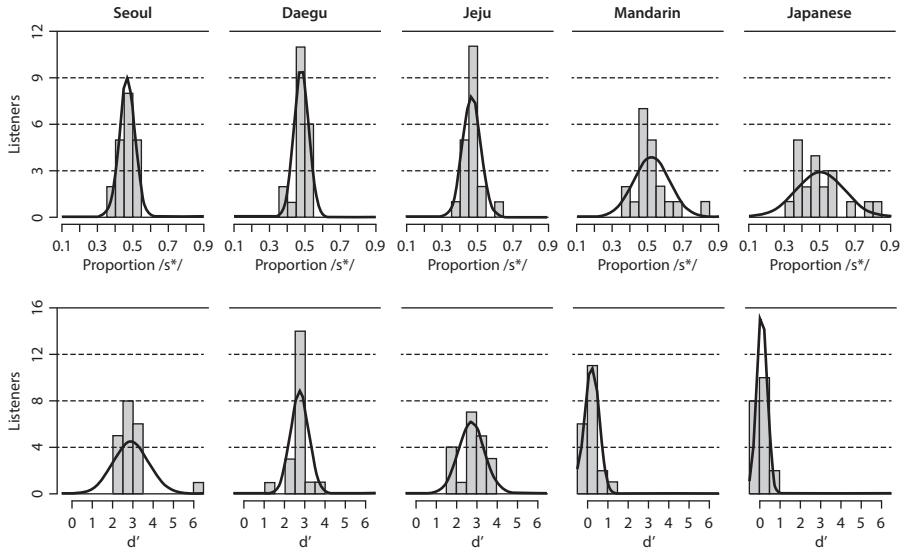

Figure 3. The distribution of /s*/ response rates and *d'* for all CV Test stimuli across all listeners.

values shown in the bottom panel of Figure 3 further confirm that the L2 learners did not do poorly because they were always choosing the same response — rather, they appear to have been guessing randomly. The L1 listeners, on the other hand, were not obviously biased towards either category and the high *d'* values indicate that listeners were very sensitive to differences between the categories. Two-sided Kolmogorov-Smirnov tests revealed no significant differences between the *d'* distributions of the L1 listener groups' ($p \geq .32$), or between the Mandarin and Japanese groups ($p = .98$).

4. Discussion

4.1 Native dialect effects

The aim of this study was to investigate the ability of different listener groups to identify Seoul Korean fricatives. It was found that there was no meaningful variation between the different L1 dialect groups of Seoul, Daegu, and Jeju listeners in either overall accuracy or in any individual vowel context. This result has implications for our understanding of the status of the /s^{h}/-/s*/ contrast in Gyeongsang dialects. Although it has been claimed that speakers in some areas of the Gyeongsang region, including Daegu (T. Kim 1991), neutralize the /s^{h}/-/s*/ contrast, young Gyeongsang speakers from both Daegu (Kenstowicz & Park 2006, Holliday 2012) and Busan (Kenstowicz & Park 2006, Lee 2013) clearly distinguish between /s^{h}/ and /s*/ in production. The results of the current study show that young listeners from Daegu can also accurately identify the /s^{h}/ and /s*/ productions of Seoul speakers just as accurately as Seoul listeners can. While the Jeju listeners were intended to serve as a control group to help interpret the difference between the Seoul and Daegu listeners as either a "Gyeongsang effect" or a "non-Seoul effect," the fact that all three dialect groups performed with equal accuracy eliminated the need for any such interpretation.

It is unlikely that the failure to find dialectal differences is due to the task being too easy, as the task used in the current study was intended to be quite difficult: the stimuli were word-initial CVs in five different vowel contexts excised from real word productions of 18 different speakers and presented in a completely randomized fashion. Nevertheless, it is still possible that other tasks could reveal some type of dialectal variation that the identification task in the current study could not. For example, an identification or discrimination task based on a synthetic continuum from /s^{h}/ to /s*/ could reveal different category boundaries across dialects. But again, because the Daegu listeners' performance was indistinguishable from the Seoul and Jeju listeners in the current study, it seems unlikely that such different

category boundaries really do exist and further unlikely that they would be able to impact the perception of naturally produced fricatives in any meaningful way.

4.2 Native language effects

In this study it was also found that both native Mandarin and native Japanese novice L2 learners of Korean are very poor at identifying Korean fricatives, and that there was no overall significant difference between these two native language groups. It was expected that the native Mandarin listeners would be more accurate than the native Japanese listeners at least in the /a/ context, if not in any others. In this case, the failure to find any difference between the two groups could have been due to the task being too difficult. Mean accuracy rates were somewhat higher in the Minimal Pair Test than in the CV Test, and so a more in-depth study of L2 perception of Korean fricatives in chunks larger than isolated CVs could reduce floor effects and reveal differences between different native language listener groups. In this study, because naïve native Mandarin listeners have been shown to assimilate /s^{h}a/ and /s*a/ to different Mandarin categories (Holliday 2014), it was predicted that novice native Mandarin L2 learners of Korean should at least be able to discriminate between /s^{h}a/ and /s*a/. In this study, their failure to accurately identify them could be due to their lack of experience with the L2: they might be sensitive to the acoustic difference between the two sounds, but have not yet learned to map these acoustic differences onto category labels. Further research is needed to understand in more detail how these sounds are acquired by L2 learners.

4.3 Vowel height effects

Despite the substantial difference in identification accuracy between the L1 and L2 listener groups, there was an effect of vowel context consistent across all groups, in which listeners were significantly more accurate in low vowel contexts than in high vowel contexts. This effect was predicted by Yoon (1999) and Chang (2007, 2013), who found that the acoustic cues relevant to the /s^{h}/-/s*/ are weaker in the /u/ and /i/ contexts. The current study also replicates and expands on Lee and Jongman (2012), who also found an effect of vowel height on identification accuracy, but because their study only tested identification in the /a/ and /i/ contexts it was unclear whether the lower identification accuracy for /i/ was due to palatalization or vowel height. In the current study, the poor accuracy of the L1 listeners in only the /u/ and /i/ contexts suggests that vowel height may be the best predictor of identification accuracy. This effect of vowel height for the L1 listeners separates not only /u/ and /i/ from /a, e, o/, but also distinguishes among /a, e, o/. These three vowels have been shown to differ in vowel height, with /a/ being the lowest and

/o/ being the highest (Yang 1996), and it was found here that the mean accuracy rates of the L1 listeners in these three vowel contexts (97.5%, 95.0%, and 91.6%, respectively) were significantly different from each other at a significance level of $\alpha = .0025$. The relationship between vowel height and identification accuracy is clearly not linear, however, as there was a steep drop in accuracy for /u/ and /i/ even though they are not much higher than /o/.

4.4 Suggestions for future work

Although no differences were found between the different dialect groups, it is still a commonly held stereotype that Gyeongsang speakers cannot correctly produce /s*/. Studies investigating the perception of Gyeongsang fricatives by Seoul listeners could potentially show whether this stereotype is purely due to perceptual bias on the part of the listener, or whether Gyeongsang fricatives really are less intelligible to Seoul listeners.

In terms of the L2 perception and acquisition of the /s^{h}/-/s*/ contrast there is still much that is unknown. Because the L2 listeners in the current study were relatively novice learners, it remains to be seen whether and how advanced L2 learners come to accurately identify the members of the contrast. It is particularly puzzling that naïve native Mandarin listeners seem to perceive a clear difference between /s^{h}a/ and /s*a/, and yet the native Mandarin listeners in the current study were not any more accurate than the native Japanese listeners. It is predicted that native Mandarin listeners should at least be better at discriminating between /s^{h}/ and /s*/, and so a comparison of the performance of L2 learners on several different perception tasks could help us understand why this contrast is so difficult.

References

Anderson, Victoria, Insung Ko & William O'Grady. 2004. A Palatographic Investigation of Place of Articulation in Korean Coronal Obstruents. *Korean Linguistics* 12: 1–24. DOI: 10.1075/kl.12.01va

Baik, Woonil. 1998. On Tensity of Korean Fricatives (Electropalatographic Study). *Speech Sciences* 4.1: 135–145.

Chang, Charles B. 2007. Korean Fricatives: Production, Perception, and Laryngeal Typology. *UC Berkeley Phonology Lab Annual Report* 20–70.

Chang, Charles B. 2013. The Production and Perception of Coronal Fricatives in Seoul Korean: The Case for a Fourth Laryngeal Category. *Korean Linguistics* 15.1: 7–49. DOI: 10.1075/kl.15.1.02cha

Cheon, Sang Yee. 2005. *Production and Perception of Phonological Contrasts in Second Language Acquisition: Korean and English Fricatives*. Ph. D. dissertation, University of Hawai'i at Mānoa.

Cho, Taehong, Sun-Ah Jun & Peter Ladefoged. 2002. Acoustic and Aerodynamic Correlates of Korean Stops and Fricatives. *Journal of Phonetics* 30.2: 193–228. DOI: 10.1006/jpho.2001.0153

Gim, Yeong Song. 1963. Kyengnampangenuy umwun [The phonology of the South Gyeongsang dialect]. *Kwukekwukmwun Hakci* [The Journal of Korean Language and Literature] 4: 51–64.

Hanson, Helen M. 1997. Glottal Characteristics of Female Speakers: Acoustic Correlates. *Journal of the Acoustical Society of America* 101.1: 466–481. DOI: 10.1121/1.417991

Holliday, Jeffrey J. 2012. The Acoustic Realization of the Korean Sibilant Fricative Contrast in Seoul and Daegu. *Journal of the Korean Society of Speech Sciences* 4.1: 67–74. DOI: 10.13064/KSSS.2012.4.1.067

Holliday, Jeffrey J. 2013. The Perceptual Assimilation of Korean Obstruents by Native Japanese Listeners. *Proceedings of the 2013 Spring Meeting of the Korean Society of Speech Sciences*, 124–126.

Holliday, Jeffrey J. 2014. The Perceptual Assimilation of Korean Obstruents by Native Mandarin Listeners. *Journal of the Acoustical Society of America* 135.3: 1585–1595. DOI: 10.1121/1.4863653

Jongman, Allard, Ratree Wayland & Serena Wong. 2000. Acoustic Characteristics of English Fricatives. *Journal of the Acoustical Society of America* 108.3: 1252–1263. DOI: 10.1121/1.1288413

Kagaya, Ryohei. 1974. A Fiberscopic and Acoustic Study of the Korean Stops, Affricates and Fricatives. *Journal of Phonetics* 2: 161–180.

Kenstowicz, Michael & Chiyoun Park. 2006. Laryngeal Features and Tone in Kyungsang Korean: A Phonetic Study. *Studies in Phonetics, Phonology and Morphology* 12.2: 247–264.

Kim, Hyunsoon, Shinji Maeda, Kiyoshi Honda & Stephane Hans. 2010. The Laryngeal Characterization of Korean Fricatives: Acoustic and Aerodynamic Data. *Turbulent Sounds: An Interdisciplinary Guide* ed. by Susanne Fuchs, Martine Toda & Marzena Zygis, 143–166. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110226584.143

Kim, Hyunsoon, Shinji Maeda & Kiyoshi Honda. 2011. The Laryngeal Characterization of Korean Fricatives: Stroboscopic cine-MRI Data. *Journal of Phonetics* 39: 626–641. DOI: 10.1016/j.wocn.2011.06.001

Kim, Sahyang. 2001. *The Interaction between Prosodic Domain and Segmental Properties: Domain Initial Strengthening of Fricatives and Post Obstruent Tensing Rule in Korean*. M. A. thesis, University of California, Los Angeles.

Kim, Taek-ku. 1991. *A Study of Geographical Differentiations in Kyongsangnam-do Dialects*. Ph. D. dissertation, Konkuk University.

Kim, Young-Seon & Baek-Hee Kim. 2013. A Study on the Phonological System of Uljin Dialect and Samchuck Dialect. *Wulimal Yenkwu* [Korean Language Research] 34: 141–162.

Klatt, Dennis H. & Laura C. Klatt. 1990. Analysis, Synthesis, and Perception of Voice Quality Variations among Female and Male Talkers. *Journal of the Acoustical Society of America* 87.2: 820–857. DOI: 10.1121/1.398894

Kong, Eun Jong, Mary E. Beckman & Jan Edwards. 2011. Why are Korean Tense Stops Acquired so Early?: The Role of Acoustic Properties. *Journal of Phonetics* 39: 196–211. DOI: 10.1016/j.wocn.2011.02.002

Lee, Goun & Allard Jongman. 2012. *Perceptual cues in Korean Fricatives*. Poster presented at the 164th Meeting of the Acoustical Society of America. Kansas City, Kansas.
Lee, Hyunjung. 2013. *Evidence of Diachronic Sound Change: A Comparative Acoustic Study of Seoul and Kyungsang Korean*. Ph. D. dissertation, University of Kansas.
Lee, Iksop & S. Robert Ramsey. 2000. *The Korean Language*. Albany: SUNY.
Martin, Samuel E. 1992. *A Reference Grammar of Korean*. Singapore: Tuttle.
Pyo, Hwa-Young, Ju-Hwan Lee, Seong-Hee Choi, Hyun-Sub Sim & Hong-Shik Choi. 1999. An Acoustic and Aerodynamic Study of Korean Fricatives and Affricates. *Eumseong Gwahak* [The Korean Society of Speech Sciences] 6: 145–161.
Schneider, Walter, Amy Eschman & Anthony Zuccolotto. 2002 *E-Prime User's Guide*. Pittsburgh: Psychology Software Tools, Inc.
Yang, Byunggon. 1996. A Comparative Study of American English and Korean Vowels Produced by Male and Female Speakers. *Journal of Phonetics* 24: 245–261. DOI: 10.1006/jpho.1996.0013
Yoon, Kyuchul. 1999. *A Study of Korean Alveolar Fricatives: An Acoustic Analysis, Synthesis, and Perception Experiment*. M. A. thesis, University of Kansas.

Author's address

Jeffrey J. Holliday
Department of Second Language Studies
Indiana University
1021 E. Third St.
Bloomington, IN 47405, USA

jjhollid@indiana.edu

[Received 14 March 2014;
revision received 10 May 2014;
accepted 21 June 2014]