Title: Phonetic cue competition within multiple phonological contrasts: Perception of Seoul Korean sibilants

Authors:

1. Jessamyn Schertz*
   Department of Language Studies, University of Toronto Mississauga
   Department of Linguistics, University of Toronto

   Email: jessamyn.schertz@utoronto.ca
   Mailing address: Department of Language Studies, UTM; Suite 301, Erindale Hall; 3359 Mississauga Road North; Mississauga, ON L5L 1C6; Canada
   Fax: 416-971-2688

2. Yoonjung Kang
   Centre for French and Linguistics, University of Toronto Scarborough
   Department of Linguistics, University of Toronto

   Email: yoonjung.kang@utoronto.ca
   Mailing address: Centre for French and Linguistics, UTSC; 1265 Military Trail, HW314; Toronto, ON M1C 1A4; Canada
   Fax: 416-971-2688

* corresponding author

Abstract

This work examines Seoul Korean listeners' perception of the five Korean sibilants: affricates /c', c, c^h/ and fricatives /s', s/. Natural productions of the consonants were manipulated to vary orthogonally along several phonetic parameters relevant to the place/manner contrast ((denti)alveolar fricative vs. (palato)alveolar affricate) and the laryngeal contrast (fortis vs. lenis vs. aspirated). Of particular interest was listeners' representation of /s/, whose laryngeal status is ambiguous. All manipulated parameters (baseline consonant and vowel affiliation, fundamental frequency at vowel onset, frication duration, and aspiration duration) influenced categorization, with consonant and vowel spectral information playing the primary role in distinguishing most sibilants. However, f0, a laryngeal cue, trumped place and manner cues in affricate vs. fricative classification, highlighting the increasing importance of f0 in Korean segmental phonology.

1. Introduction

Listeners draw on a rich array of acoustic details to identify which sounds they hear. Most experimental work quantifies the use of acoustic dimensions in binary contrast classification, for example by examining laryngeal contrasts like /b/ vs. /p/ and place of articulation contrasts like /b/ vs. /d/ separately. However, in everyday online perception, listeners must consider these multiple contrasts simultaneously. This work investigates the interplay of laryngeal, place, and manner cues to the perception of Korean sibilants: three affricates /c', c, c^h/ and two fricatives /s', s/. Traditionally, the fricatives and affricates are described as contrasting in place as well as manner, with the fricatives produced as (denti-) alveolar (e.g. Cho et al. 2002) and affricates as palatal (e.g. Kim-Renaud 1974). However, articulatory work indicates that the affricates are articulated in the alveolar region (e.g. Kim

2001, Anderson et al. 2004), with dialect- and age-related variation found in acoustic work (Kong et al. 2014, Kang et al. 2014).

From both a phonological and phonetic point of view, the affricates fit in unambiguously with the three-way laryngeal contrast characterizing Korean stops (fortis /c'/, lenis /c/, aspirated /c^h/) (Kim 2004, Jang 2011), differing in terms of aspiration, f0 (fundamental frequency at vowel onset, the acoustic correlate of pitch), and voice quality (e.g. Cho et al. 2002, Lee & Jongman 2012). Although traditionally described as a three-way aspiration distinction (*Fortis < Lenis < Aspirated*; e.g. Lisker and Abramson 1964), recent work has shown that the aspiration durations of the lenis and aspirated categories are undergoing a merger in word-initial stops (Silva 2006), and f0 is becoming a more primary cue for younger speakers (*Lenis < Fortis < Aspirated*; e.g. Kang & Guion 2008, Kang 2014). Voice quality also plays a role, with vowels following fortis stops showing more "pressed" or laryngealized voice quality than the "breathier" vowels following lenis and aspirated stops (e.g. Cho et al. 2002).

On the other hand, the laryngeal status of the fricative contrast is ambiguous. Fortis /s'/ patterns phonetically and phonologically with the other fortis consonants, but "nonfortis" /s/ shares characteristics of both the lenis and aspirated series. Like both lenis and aspirated stops, it has phonetic aspiration and relatively breathy voice. However, it patterns phonologically with lenis stops (including a loss of aspiration in medial position), while its relatively high pitch suggests an association with aspirated, and not lenis, stops (e.g. Iverson 1983, Chang 2013, Kang 2000, Kang et al. 2009).

Studies focusing on perception of the two-way fricative contrast have converged to show that both consonantal and vocalic information influence categorization. Chang (2013)

found that frication duration (shorter frication = more /s/ responses), aspiration duration (longer aspiration = more /s/ responses), fundamental frequency (lower f0 = more /s/ responses), and vowel affiliation (vowels spliced from /s/ = more /s/ responses than vowels spliced from /s'/) all play a role, with the vocalic cues dominating perception, especially in the context of the vowel /a/ (see also Yoon 1999, Holliday 2010, Jang 2011, Holliday 2014 Lee & Jongman 2014). There has been less work on perception of the fricative-affricate manner contrast; however, Park et al. (1998) examined the influence of frication duration and rise time on perception of the /s/-/c/ contrast, and found a role for both dimensions, with frication duration serving as the dominant cue.

Along with place and manner, the /s/-/c/ contrast may also differ in laryngeal status, given the ambiguous laryngeal affiliation of /s/. Since previous experiments targeting listeners' perception of the /s/-/c/ contrast have been independent from those focusing on the fricative laryngeal contrast /s/-/s'/, the *relative* importance of laryngeal vs. place/manner cues to the /s/-/c/ contrast has not been examined. Recall that /s/ shares characteristics of both lenis and aspirated consonants. Discrepancies in the use of f0 between perception and production of /s/ add to this ambiguity: Chang (2013) found no difference between /s/ and /s'/ in production, with /s/ showing the same high pitch as the fortis series, but found that the same participants used pitch to distinguish the two in perception, with /s/ being cued by *lower* pitch than /s'/. Given the overall fuzziness surrounding the laryngeal affiliation of /s/, it is reasonable to expect that listeners may place less weight on the ambiguous laryngeal component of the /s/-/c/ contrast, relying more heavily on the unambiguous place/manner contrast when cues are in competition.

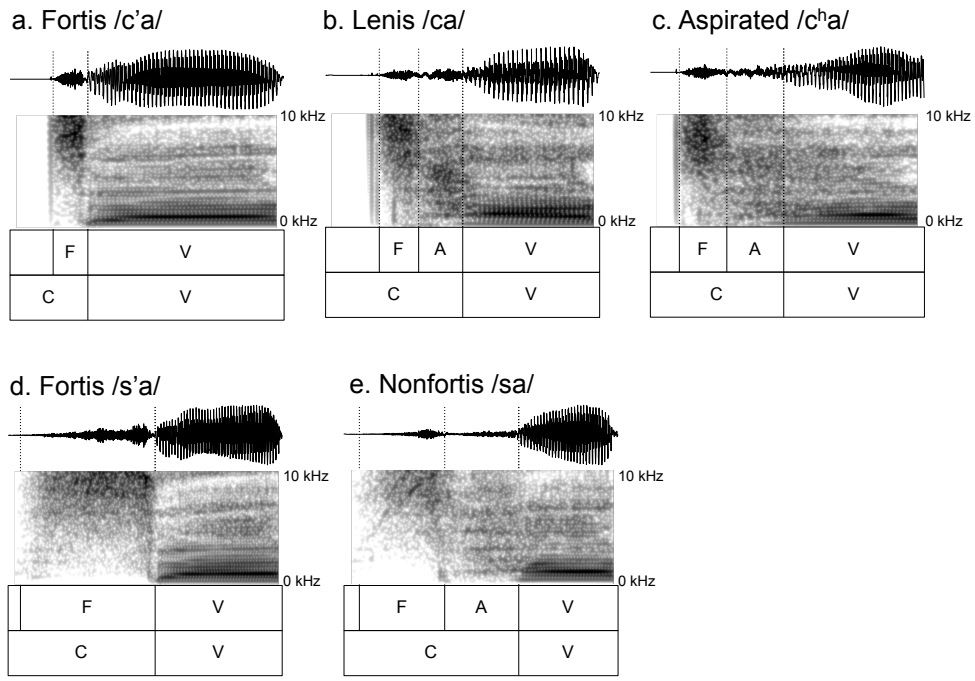
The current study maps the sibilant perception of native Seoul Korean listeners currently living in Canada, within a broad acoustic stimulus space encompassing all five sounds. By pitting laryngeal and place/manner cues against each other within the same experiment, we aim to address listeners' representation of phonological *classes*, in addition to individual phonetic categories. In particular, we test the prediction that place/manner cues will play a stronger role than laryngeal cues in listeners' perception of /s/-/c/ contrast, stemming from the ambiguous laryngeal status of /s/. We also explore how the increasing importance of f0 in the stop contrast might extend its influence to perception of the sibilant system more generally.

2. Methods

24 native Korean listeners (12 F, 12 M, 20-28 years old in 2015) participated in the experiment. All grew up in Seoul or the surrounding area and were currently residing in Toronto (mean length of residence 5.4 years).

Stimuli: Sibilant-initial monosyllables /c'a/, /ca/, /c^{h}a/, /s'a/, and /sa/ were recorded by a young female speaker from Seoul. The vowel /a/ was chosen based on production data (Kang et al. 2009) showing that some relevant acoustic cues (aspiration duration, COG) were most reliably distinguished in this vowel context. Waveforms and spectrograms of these target syllables, which served as the baseline for stimulus manipulations, are shown in Figure 1. These prototypical tokens differ in the various acoustic dimensions relevant to the various phonological contrasts within the sibilant system, summarized in Table 1. The acoustic values of the baseline stimuli for all dimensions are given in Table 4 in the Appendix.

Figure 1: Waveforms and spectrograms of natural recordings, labeled with annotations used for stimulus manipulations (F = frication, A = aspiration, V = Vowel, C = Consonant).



First, the affricates (a-c) differ from the fricatives (d-e) in **manner of articulation**. A visible burst in the spectrogram distinguishes affricates from fricatives, as does a faster intensity rise time (the waveform reaches its maximum amplitude faster in affricates than fricatives). Furthermore, the frication duration itself is longer in fricatives than affricates. Affricates and fricatives also differ in **place of articulation**, with affricates showing on average lower frequency frication noise (corresponding to backer articulation) than fricatives. The acoustics also differ based on the **laryngeal** status of the consonants. While fortis fricatives and affricates are unaspirated, all other consonants, including the fricative /s/, include a substantial period of aspiration between release of the fricative constriction and vowel onset. Several other properties of the laryngeal contrast, while less evident from visual inspection of the waveform and spectrogram, also characterize these stimuli. Vowels

following fortis consonants have more laryngealized or creaky voice quality, while those following lenis and aspirated stops are breathier. Although less discussed, duration of the following vowel may also differ as a function of laryngeal category: vowels following fortis segments tend to be longer than those of lenis or aspirated segments, at least before /a/ (Broersma 2010). Finally, f0 at vowel onset robustly distinguishes the laryngeal contrast in affricates: lenis /c/ is produced with lower f0 in the following vowel than its fortis and aspirated counterparts, while the fricatives /s/ and /s'/ are characterized by similar f0 trajectories in the following vowel.

Table 1: Parameters manipulated from natural recordings. The specific acoustic values for the stimuli used in the current work are given in the Appendix (Table 4).

Parameter	**Levels**	**Relevance to Place, Manner, and/or Laryngeal contrasts**
Consonant Affiliation (C-Affil)	/c'a/, /ca/, /c^{h}a/, /s'a/, /sa/	**Place contrast:** fricative constriction is fronter than affricate constriction. **Manner contrast**: affricates have bursts and overall shorter rise times than fricatives.
Vowel Affiliation (V-Affil)	/c'a/, /ca/, /c^{h}a/, /s'a/, /sa/	**Laryngeal contrast:** vowels following fortis consonants /c'/ and /s'/ have lower H1-H2 (creakier voice quality) and longer durations than those following nonfortis consonants.
Fundamental frequency (f0)	211, 273, 335 Hz (onset)	**Laryngeal contrast:** lenis consonants (/c/) have lower f0 in following vowel than fortis and aspirated counterparts (/c', c^h/); /s/ and /s'/ have similar f0.
Frication duration (FricDur)	25, 75, 155 ms	**Manner contrast**: frication duration is longer in fricatives than affricates.
Aspiration duration (AspDur)	0, 45, 90 ms for C-Affil /ca, sa, c^{h}a/; 0 ms for C-Affil /c'a, s'a/	**Laryngeal contrast**: fortis consonants (/c', s'/) are unaspirated; all others (/c, c^h, s/) are aspirated.

The baseline stimuli shown in Figure 1 were then manipulated along several parameters expected to influence perception of the place, manner, and laryngeal contrasts (Table 1) to create an acoustic space encompassing all five sibilants, then embedded in the carrier sentence __-라고 했어요 (__-lako hayssta, '(I) said __').

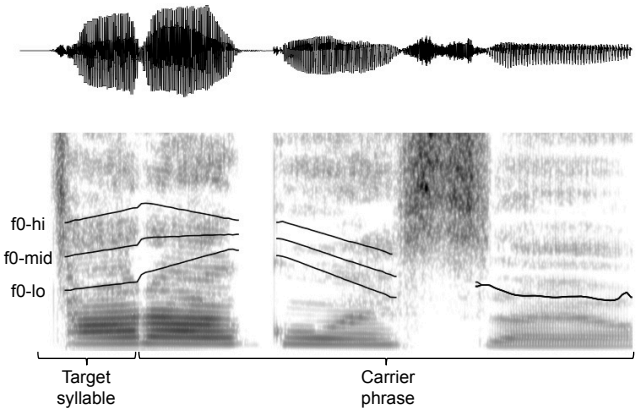
To create the stimuli, the natural productions were manually annotated for frication, aspiration (if any), and vocalic portions of each sibilant (cf. Figure 1). The onset of frication was marked at the beginning of visible frication in the waveform. The boundary between frication and aspiration (if any) was placed at the transition between high-frequency frication and more diffuse aspiration noise in the spectrogram. The boundary between the end of frication or aspiration and vowel onset was marked as the onset of periodicity of the following vowel. We define the "consonantal" portion as everything before vowel onset, such that it includes the burst (if any), frication, and aspiration (if any), as shown in the second tier of the annotations in Figure 1.

25 baseline tokens were created by cross-splicing the five consonantal portions of the natural productions of each sibilant (C-Affil) with the five vocalic portions (V-Affil). The cross-splicing procedure allowed us to independently vary the consonantal cues (including differences in COG and rise time targeted in previous work, as well as any other spectral differences between consonants) and vocalic cues (including voice quality and vowel duration). Each baseline syllable was spliced into the carrier sentence.

These cross-spliced baseline tokens were manipulated to vary in equally-spaced steps of f0, frication duration (FricDur), and aspiration duration (AspDur). f0 and duration manipulations were both performed using the PSOLA algorithm (Moulines & Charpentier 1990) as implemented in Praat (Boersma and Weenink, 2011). The endpoints of 25 and 155

ms for frication duration were based on the minimum (from /c'a/) and maximum (from /s'a/) frication durations of the speakers' natural productions. Aspiration was manipulated for baseline nonfortis consonants only (i.e., those baseline stimuli containing aspiration: /c/, /c^h/, or /s/), because splicing aspiration onto fortis consonants resulted in unnatural-sounding tokens. We slightly extended the natural range of aspiration duration in our speaker's productions (ranging from 0 ms for fortis /c'a/ and /s'a/ to 70 ms for /sa/), for a range of 0 to 90 ms. Finally, for f0, although the largest differences in production are found at vowel onset, f0 perturbation stemming from laryngeal differences extends across multiple syllables (e.g. Kim 2000), and this proved to be the case in our speaker's natural productions as well. In order to account for this long-distance perturbation in our manipulations, we extracted our speaker's natural pitch contours across the target syllable and three subsequent syllables. We then created stylized f0 contours based on the lowest (/ca/) and highest (/c^{h}a/) of these contours, as well as intermediate level falling between the two, as shown in Figure 2.

Figure 2: f0 contours across the target syllable and carrier phrase



Procedure: Participants listened to each stimulus and chose which of the 5 syllables (presented in Korean orthography) they heard, with the option of "none of the above." The full set of 495 stimuli were presented using PsychoPy (Peirce 2007), divided into 5 blocks. The order of presentation was randomized separately for each listener. The experiment was conducted in a sound-attenuated booth at the University of Toronto, and the task took about 20 minutes; listeners were paid for their participation.

2.1. Statistical analyses

Comparisons of overall patterns: We analyzed how listeners' responses were influenced by each acoustic cue with mixed-effects binary logistic regression using the lme4 package v. 1.1-7 in R (Bates et al. 2015). We built a separate model for each response choice, pitting each against the four other choices pooled together (e.g. /s/ vs. all other choices, cf. Lee et al. 2013). Trials in which listeners chose "none of the above" (3.8% of responses) were excluded prior to analysis. Each model estimated the likelihood of a given response choice (e.g. probability of choosing /s/ vs. all other consonants combined) and included a fixed effect for each manipulated parameter, with by-subject random intercepts and random slopes for fixed parameters. C-Affil and V-Affil were each collapsed into binary factors: C-Affil was collapsed into fricative (from baseline /sa/ and /s'a/) vs. affricate (from baseline /ca/, /c'a/, and /c^{h}a/), while V-Affil was collapsed into fortis (from baseline /s'a/ and /c'a/) vs. nonfortis (from baseline /sa/, /ca/, and /c^{h}a/). Both C-Affil and V-Affil were sum-coded; the three numerical factors (f0, FricDur, and AspDur) were converted to z-scores prior to analysis. These analyses were used as a metric for how much each parameter contributed to the choice of a given sibilant within the context of the entire system.

Targeted two-way comparisons: We also built models exploring listeners' use of cues in perception of two binary contrasts of particular interest: the fortis vs. nonfortis fricative contrast (/s'/ vs. /s/) and the fricative-affricate manner contrast (/s/ vs. /c/). For each of these two-way contrasts, we built a logistic regression model identical to those described above but only considering the subset of the data where listeners had chosen one of the two target responses. The stimulus set was restricted to those sounds created from baseline fricatives (C-Affil = fricative) for the /s/-/s'/ comparison, and those sounds created from baseline nonfortis tokens (V-Affil = nonfortis) for the /s/ vs. /c/ contrast, to facilitate comparison with previous work. We used coefficients from these analyses to quantify listeners' relative use of various cues which have been shown to influence perception in previous work but which have not been examined simultaneously.

3. Results

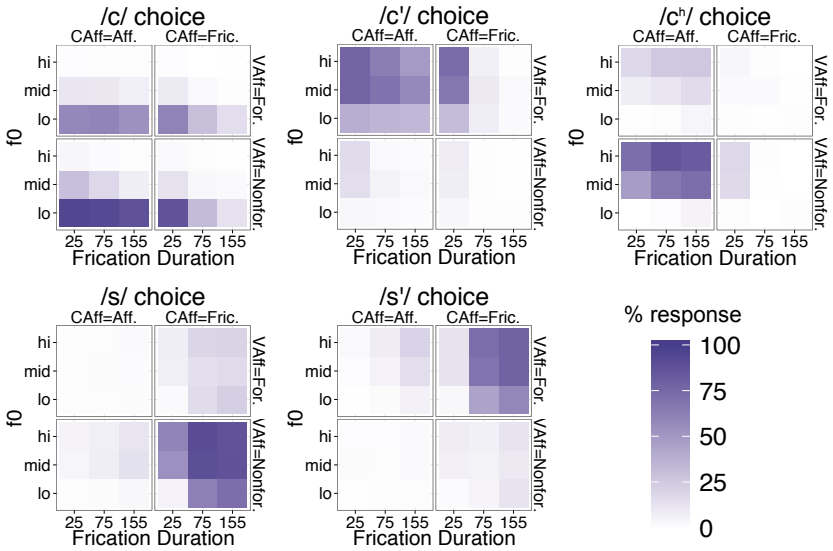
3.1. Overall Results

The heat plots in Figure 3 map the experimental stimulus space in terms of the five parameters described above four of the five parameters described above (f0, FricDur, C-Affil, and V-Affil in 3a, and AspDur in Figure 3b). Individual cells represent stimuli with different values of each parameter, and the darkness of each cell gives a measure of listeners' responses, with darker cells representing a higher percentage for a given response. From the graphs, it is clear that C-Affil and V-Affil, which provide spectral information about the consonant and vowel respectively, play relatively deterministic roles for most sounds. For example, in the fricatives /s/ and /s'/, the dark portions of the graphs (i.e. high response rates) are localized to the stimuli created from baseline fricative consonants (C-Affil = Fric.,

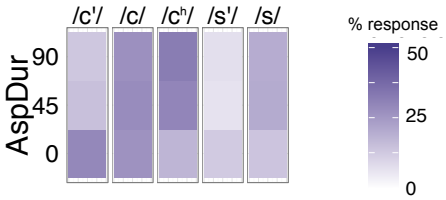
the right half of each graphs) and the corresponding baseline vowels (e.g. V-Affil = Fortis for /s'/, upper half of the graphs). Similarly, listeners heard the aspirated affricate /c^{h}/ primarily when the baseline consonant was an affricate and the baseline vowel was nonfortis.

Figure 3: Darkness represents the percentage of time listeners chose a given response as a function of different manipulated parameters. (a) Parameters: f0, FricDur, C-Affil, V-Affil. (b) Parameter: AspDur.

a.



b.



The other two affricates, /c/ and /c'/, are somewhat less constrained by baseline consonant and vowel affiliation. Although listeners required a fortis V-Affil to choose the fortis affricate /c'/, C-Affil was less important. Examining the combined patterns of the /c'/ and /s'/ responses (i.e., the fortis manner contrast), a trading relationship between frication

duration and C-Affil emerges: there is a high proportion of /c'/ responses even when the baseline consonant is a fricative (C-Affil = Fric.), as long as the frication duration is short, and there is an increase in /s'/ responses to stimuli based on affricates (C-Affil = Aff.) as the frication duration increases. The Lenis affricate /c/ shows the most divergence from vowel and consonant affiliation, with fairly high response rates seen in all quadrants, as long as f0 is low. Along with these general trends, smaller effects of individual factors are discussed below in the statistical results.

Figure 3b shows participants' responses as a function of AspDur (which was only manipulated from baseline nonfortis consonants). Listeners were more likely to choose /s/ and /c^h/ when there was aspiration, and more likely to choose /c'/ and /s'/ when there was not, reflecting production patterns. Interestingly, aspiration appears to have little to no influence on listeners' categorization of the lenis affricate /c/, even though it is clearly aspirated in production.

Table 2 shows results from the five regression analyses. The beta-coefficients give the predicted increase (for positive) or decrease (for negative) in log odds of the given response choice when each parameter is manipulated. For the numerical factors (f0, FricDur, and AspDur), a positive coefficient shows that listeners were more likely to choose the given consonant for stimuli with higher values of a given factor. For example, the positive beta-coefficient for f0 in the model predicting /c'/ choice shows that listeners were more likely to choose /c'/ as f0 increased. For each categorical factor (V-Affil and C-Affil), the beta-coefficients show the difference in log odds between the two levels of the factor. Positive beta-coefficients for V-Affil represent higher response rates for nonfortis (compared with fortis) baseline vowels, while positive coefficients for C-Affil represent higher response rates

for stimuli created from fricative (vs. affricate) baseline consonants. The z-scores in Table 2 show the effect normalized by its standard error, and z-scores with absolute values greater than two can generally be considered significant at $p < .05$.

Table 2: β-coefficients and z-scores from mixed-effects logistic regression analyses predicting each response choice as a function of all manipulated parameters.

	/c'/ choice		/c/ choice		/c^h/ choice		/s'/ choice		/s/ choice	
	β	z	β	z	β	z	β	z	β	z
(Intercept)	-2.90	-13.92	-3.43	-12.12	-3.92	-19.03	-3.26	-20.07	-2.76	-13.62
f0	0.81	5.70	-3.99	-16.48	2.23	25.11	0.55	6.20	0.67	8.74
FricDur	-1.14	-13.25	-0.84	-15.90	0.25	3.86	1.14	14.60	0.88	11.93
AspDur	-0.98	-8.79	*0.10*	*1.53*	1.08	18.32	-0.64	-7.97	0.13	2.05
V-Affil: nonfortis (vs. fortis)	-3.73	-21.71	1.40	5.22	3.47	10.54	-2.84	-15.28	3.35	12.33
C-Affil: fricative (vs. affricate)	-1.97	-17.33	-2.09	-12.29	-4.43	-21.77	2.88	14.28	3.99	23.71

Along with the strong influence of both C-Affil and V-Affil, seen in the graphs above, the results demonstrate that *all* of the manipulated parameters, with the exception of AspDur for /c/ (in gray in Table 2, $p > .1$), exerted a significant influence on *all* response choices (effect of AspDur on /s/ choice $p = .04$, all others $p < .001$). In particular, more fortis affricate /c'/ responses were elicited by shorter FricDur, shorter AspDur, and higher f0 (in order of relative influence). More lenis affricate /c/ responses were elicited by lower f0 and shorter FricDur, with no significant effect of AspDur. More aspirated affricate /c^h/ responses were elicited by higher f0, longer AspDur, and longer FricDur. More fricative responses were elicited with longer FricDur and higher f0. Finally, longer AspDur elicited more nonfortis /s/ responses, while shorter AspDur elicited more fortis /s'/.

The overall prominence of f0 in predicting /c/ responses, discussed above, is also reflected in the statistical results. The β-coefficient of f0 (-3.99) is much larger for /c/ than for all other responses, while the β-coefficients for V-Affil (1.40) and C-Affil (-2.09) are

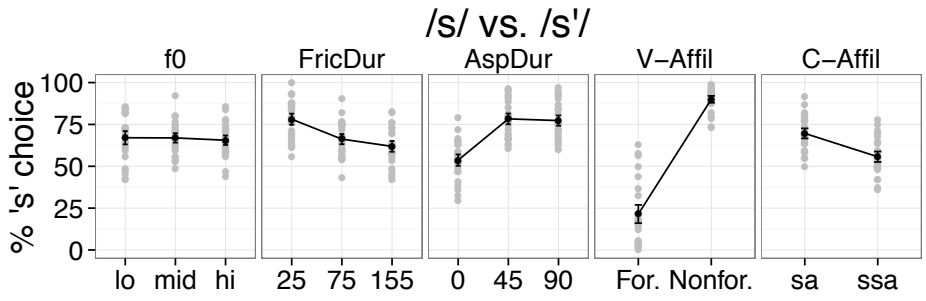
smaller, demonstrating the primacy of f0, compared to other cues, in influencing response patterns for /c/.

3.2. Targeted two-way comparisons

Fricative laryngeal contrast: /s/ vs. /s'/: Graphs and statistics showing listeners' use of the five acoustic parameters in the fortis vs. nonfortis fricative contrast are shown in Figure 4a and Table 3 (left). For the fricative contrast, both aspiration and laryngeal status of the vocalic portion influenced categorization. As indicated by the graph in Figure 4a and the large beta-coefficient, V-Affil plays the primary role in the fricative laryngeal contrast: fewer than 25% of stimuli with fortis baseline vowels were heard as /s/, compared to greater than 80% of stimuli from baseline nonfortis vowels. The presence of aspiration also elicited significantly more /s/ responses, although as shown in Figure 4a, longer (90 ms) aspiration duration did not elicit more /s/ responses than shorter (45 ms) duration. On the other hand, f0, FricDur, and C-Affil (/s/ vs. /s'/) did not significantly influence fricative classification[1].

Figure 4: Response patterns (percentage /s/ response) for (a) the fortis vs. nonfortis fricative contrast and (b) the nonfortis fricative vs. affricate contrast.

a.



[1] AspDur was only manipulated from baseline nonfortis consonants, introducing a confound in the analysis. We therefore investigated the effect of AspDur on the subset of data where C-Affil = /s/. The pattern of results was the same (51%, 79%, and 78% /s/ response for AspDur = 0, 45, and 90 ms), and a regression model with AspDur as the predictor variable retained significance (p<.001), suggesting an independent effect for AspDur.

b.

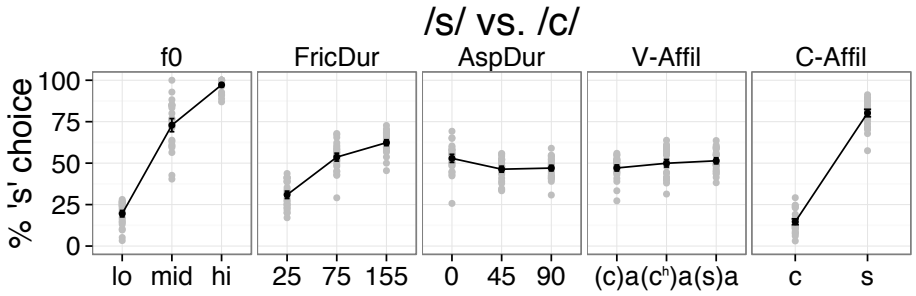


Table 3: β-coefficients, z-scores and p-values from mixed regression analyses of two-way comparisons of the fricative contrast /s/ vs. /s'/ (reference level) and the manner contrast /s/ vs. /c/ (reference level). Higher beta-coefficients represent higher rates of /s/ choice.

	/s/ vs. /s'/ (/s'/ = ref. level)				**/s/ vs. /c/ (/c/ = ref. level)**		
	β	z	p		β	z	p
(Intercept)	**0.784**	**2.966**	**.003**	**(Intercept)**	**-2.394**	**-9.349**	**< .001**
f0	-0.085	-0.735	> .1	**f0 (binary)**	**8.228**	**9.745**	**< .001**
FricDur	-0.001	-0.006	> .1	**FricDur**	**2.054**	**9.332**	**< .001**
AspDur	**1.955**	**8.066**	**< .001**	AspDur	0.213	1.769	.08
V-Affil nonfor (vs. for)	**6.012**	**10.575**	**< .001**	V-Affil = (c^h)a (vs. (c)a)	-0.042	-0.179	> .1
C-Affil /s'/ (vs. /sa/)	-0.039	-0.177	> .1	V-Affil = (s)a (vs. (c)a)	0.057	0.259	> .1
				C-Affil = fric (vs. aff.)	**5.32**	**10.492**	**< .001**

Manner contrast: /s/ vs. /c/: Listeners' use of the five acoustic parameters in the /s/ vs. /c/ manner contrast are shown in Figure 4b and Table 3 (right). The graphs reflect the primary importance, discussed above, of f0 and C-Affil to this contrast. In order to be able to directly compare the relative role of these two cues in the logistic regression analysis, we converted f0 to a binary factor, only including low and high f0 tokens (omitting mid-f0 tokens), such that it matched the number of levels of the C-Affil factor. Comparing the beta-coefficients of these two factors suggests that f0 plays a larger role than C-Affil in cuing the /c/-/s/ contrast. Nevertheless, C-Affil also plays an important role, and frication duration further contributes to listeners' responses. Effects of AspDur and V-Affil were not significant.

4. Discussion

Overall, our results reinforce the findings of previous work demonstrating that consonantal and vocalic spectral cues play important roles in listeners' categorization of Korean sibilants. Of the five manipulated parameters, baseline consonant affiliation (providing place and manner cues) and baseline vowel affiliation (providing laryngeal cues) are of primary importance for categorization of four out of the five sibilants. However, the lenis affricate /c/ was an exception to this generalization, in that it was defined above all by low f0. Beyond these primary patterns, almost all parameters influenced classification of all sibilants, albeit to a lesser degree. In particular, longer frication duration elicited more /c^h/, /s'/, and /s/ responses (and fewer /c'/ and /c/ responses), while longer aspiration duration elicited more /c^h/ and /s/ responses (and fewer /s'/ and /c'/ responses, with no significant effect on /c/ responses). Additionally, a trading relation between frication duration and consonant affiliation defined the perception of the fortis fricative-affricate contrast (/s'/ vs. /c'/).

The two-way fortis vs. nonfortis fricative contrast (/s/-/s'/) was characterized primarily by vowel affiliation, with a secondary effect of aspiration duration, reflecting results of Chang (2013) and Lee & Jongman (2014). Further, aspiration appears to play a categorical rather than gradient role with these listeners, with its presence vs. absence, as opposed to duration, driving the effect. Although the natural recordings of the two consonantal baselines differed in their centre of gravity of frication (9317 Hz for /s/ vs. 10300 Hz for /s'/), listeners did not appear to use this as a cue to perception (as in Chang 2013, but see Holliday 2010). f0 also failed to play a role, diverging from the (unexpected) finding of Chang (2013), in which higher f0 elicited more /s'/ responses. This discrepancy

may arise from the fact that participants in Chang (2013) chose between /s/ and /s'/ only, as opposed to the full set of sibilants in the current work. Since our findings suggest that f0 is above all a cue to /c/, stimuli with lower f0 which were categorized as /s/ by the participants in Chang (2013) may have been better categorized as /c/. This highlights the importance of considering the options given to the listener when interpreting results of forced-choice cue weighting tasks.

Given the ambiguous laryngeal status of /s/, we expected listeners to rely more heavily on place or manner cues than laryngeal cues when distinguishing the /s/-/c/ contrast. Unexpectedly, our data suggest the opposite: while the place and manner cues in C-Affil do play an important role, and frication duration plays a minor role, listeners distinguish the /s/-/c/ contrast primarily using f0, a laryngeal cue. This appears to diverge from previous work on perception of the contrast, in which frication duration was found to provide the dominant cue to the manner contrast (Park et al. 1998); however, f0 was not manipulated in that work. This prominence of f0 in cuing the fricative-affricate "manner" contrast therefore highlights the overall importance of f0 in the laryngeal system of contemporary Korean listeners; in a sense, the use of f0 as a laryngeal cue is also permeating the manner contrast more generally.

The current study only examined sibilants preceding /a/, the context in which vocalic information is most critical to fricative identification, as compared to other vowel contexts (Chang 2013, Lee & Jongman 2014). Nevertheless, we predict that our finding of prominent use of f0, even for the /s/-/c/ "manner" contrast, would generalize to other vowel contexts, given that a carefully controlled study (Chang 2013) revealed equal or greater effects of f0 in the contexts of /i/, /u/, and /ɯ/, as compared to /a/. Whether our findings would generalize

to cue use in other prosodic positions such as word-medial position, where f0 is a less reliable cue to the laryngeal contrast in production (e.g. Kim 2000), are less clear, and future work should explore to what extent this primary use of f0 generalizes to other environments.

Overall, Korean listeners use a combination of place, manner, and laryngeal cues to distinguish sibilants. These multiple cues, integrated with higher-level (lexical, syntactic, and pragmatic) information, presumably aid the listener in the face of a noisy and unpredictable signal. At the same time, taking the perspective of Winter (2014), the presence of many functionally redundant cues for a single sound contrast provides the variation necessary for language evolution. Under this view, it is perhaps not surprising that a particularly rich and interconnected system of contrasts, the set of Korean sibilants examined in the current work, is undergoing widespread dialectal variation and sound change.

Acknowledgments

This work was funded by SSHRC Grant #435-2013-2092 to Yoonjung Kang. We would like to thank Sungwoo Han and Hyeseon Maeng from Inha University for their assistance with stimulus preparation.

Appendix

Table 4: Acoustic parameters of baseline tokens used for manipulations.

	Centre of Gravity (Hz)	Rise time (ms)	Vowel duration	H1-H2 (dB)	Frication duration (ms)	Aspiration duration (ms)	f0, onset (Hz)	f0, midpoint (Hz)
/c/	7367	31	126	12	36	40	222	231
/c'/	9331	24	150	-2	27	0	317	320
/c^h/	7770	21	108	8	37	44	315	350
/s/	9317	66	92	6	79	69	337	350
/s'/	10300	90	142	-5	155	0	302	311

References

Anderson, Victoria, Insung Ko, & William O'Grady. 2004. Palatographic investigation of place of articulation in Korean coronal obstruents. *Korean Linguistics* 12.1–24.

Bates, Douglas, Martin Maechler, Benjamin M. Bolker, & Steven C. Walker. 2015. Lme4: Linear Mixed-Effects Models using Eigen and S4. R package version 1.1-8, http://CRAN.R-project.org/package=lme4.

Boersma, Paul, & David Weenink. 2011. Praat: Doing phonetics by computer. Version 5.3, http://www.praat.org.

Broersma, M. 2010. Korean lenis, fortis, and aspirated stops: Effect of place of articulation on acoustic realization. *INTERSPEECH 2010*: 941-944.

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.

Cho, Taehong, Sun-Ah Jun, & Peter Ladefoged. 2002. Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics* 30:2.193–228.

Holliday, Jeffrey J. 2010. An acoustic study of L2 perceptual acquisition of Korean fricatives. *Harvard Studies in Korean Linguistics* 13.17–32.

Holliday, Jeffrey J. 2014. The perception of Seoul Korean fricatives by listeners from five different native dialect and language groups. *Korean Linguistics* 16:2.93-110.

Iverson, Gregory K. 1983. Korean S. *Journal of Phonetics* 11.191–200.

Jang, Hyejin. 2011. 국어 어두 장애음의 음향적 특성과 지각 단서. PhD thesis.

Kang, Kyoung-Ho, & Susan Guion. 2008. Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. *The Journal of the Acoustical Society of America* 124:6.3909–17.

Kang, Kyung-Shim. 2000. On Korean fricatives. *Eumseonggwahag* 7:3.53–68.

Kang, Yoonjung. 2014. Voice Onset Time merger and development of tonal contrast in Seoul Korean stops: A Corpus Study. *Journal of Phonetics* 45.76–90.

Kang, Yoonjung, Sungwoo Han, Alexei Kochetov, & Eun Jong Kong. 2014. Dialectal variation in affricate place of articulation in Korean. *The 168th Meeting of the Acoustical Society of America* 136:4.2145.

Kang, Yoonjung, Alexei Kochetov, & David Go. 2009. The acoustics of Korean fricatives. *CRC-Sponsored Summer Phonetics/Phonology Workshop, Department of Linguistics, University of Toronto*.

Kim, Hyunsoon. 2001. The place of articulation of the Korean plain affricate in intervocalic position: An articulatory and acoustic study. *Journal of the International Phonetic Association* 31:2.229–57.

Kim, Midam. 2004. Correlation between VOT and F0 in the perception of Korean stops and affricates. *Interspeech 2004* 8.

Kim, Mi-Ryoung. 2000. Segmental and tonal interactions in English and Korean: A phonetic and phonological study. Ph.D. thesis, University of Michigan.

Kim-Renaud, Young-Key. 1974. Korean consonantal phonology. PhD thesis, University of Hawaii, Manoa.

Kong, Eun Jong, Soyoung Kang & Misun Seo. 2014. Gender difference in the affricate productions of young Seoul Korean speakers. *The Journal of the Acoustical Society of America* 136:4.EL329–L335.

Lee, Goun & Allard Jongman. 2014. Cue integration in the perception of fricative-vowel coarticulation in Korean. *The Journal of the Acoustical Society of America* 136:4.2173.

Lee, Hyunjung & Allard Jongman. 2012. Effects of tone on the three-way laryngeal distinction in Korean: An acoustic and aerodynamic comparison of the Seoul and South Kyungsang dialects. *Journal of the International Phonetic Association* 42:2.145–69.

Lee, Hyunjung, Stephen Politzer-Ahles, & Allard Jongman. 2013. Speakers of tonal and non-tonal Korean dialects use different cue weightings in the perception of the three-way laryngeal stop contrast. *Journal of Phonetics* 41:2.117–32.

Lisker, Leigh & Arthur Abramson. 1964. A cross-language study of voicing in initial stops: acoustical measurements. *Word* 20.384–422.

Moulines, Eric & Francis Charpentier. 1990. Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communication* 9:5.453-467.

Park, Soonboak, Bong-won Yi, Jiyoung Shin, & Kee-Ho Kim. 1998. Perceptual cues for Korean affricate vs. fricative distinction. *Speech Sciences* 4:1.47–58.

Peirce J. W. 2007. PsychoPy - Psychophysics software in Python. *Journal of Neuroscience Methods* 162:1-2.8–13.

Silva, David. 2006. Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology* 23:2.287–308.

Winter, Bodo. 2014. Spoken language achieves robustness and evolvability by exploiting degeneracy and neutrality. *Bioessays* 36.960–67.

Yoon, Kyuchul. 1999. A Study of Korean Alveolar fricatives: An acoustic analysis, synthesis, and perception experiment. *Proceedings of the 34th Mid-America Linguistics Conference*, 549–63.