On Korean Fricatives: Production, Perception, and Laryngeal Typology

Charles B. Chang
University of California, Berkeley
charleschang@berkeley.edu

81st Annual Meeting of the Linguistic Society of America
6 January 2007

Outline
1. Background and research questions
2. Experiment 1: Production of /sʰa/ and /sʰa/
3. Experiment 2: Perception of /sʰa/ and /sʰa/
4. Experiment 3: Production of /sʰu/ and /sʰu/
5. Experiment 4: Perception of /sʰu/ and /sʰu/
6. Implications of the results
7. Conclusions

Background

• Korean: three-way laryngeal contrast in plosives and affricates
  - Series #1 = lenis (a.k.a. "lax", "weak", "plain", "slightly aspirated", "breathy")
  - Series #2 = fortis (a.k.a. "tense", "strong", "glottalized", "long", "unaspirated", "forced")
  - Series #3 = aspirated (a.k.a. "heavily aspirated", "strongly aspirated", "super aspirated")
  - all are voiceless in utterance-initial position
  - contrast corresponds to differences in:
    - linguopalatal contact
    - glottal configuration
    - subglottal and oral pressure
    - laryngeal and supralaryngeal articulatory tension
    - voice onset time (VOT)
    - fundamental frequency (f0), intensity, and voice quality of vowel onset

Questions

• What do acoustic data reveal about how the laryngeal contrast in Korean fricatives is realized?

• What are the cues that are most important in signaling the contrast in Korean fricatives?

• What does this cue hierarchy suggest about the laryngeal categories that Korean fricatives correspond to?
Experiment 1: Variables

- Dimensions examined for the fricatives:
  - duration
  - aspiration (≈ VOT)
  - f0 onset
  - F1 onset
  - intensity buildup (rate of change)
  - spectral tilt (H1-H2)
  - vowel length

Experiment 1: Methods

- 5 native Korean speakers read off a word list including /sʌ/ 'buy' and /sʰʌ/ 'cheap'
- Measurements taken in Praat 4.2.17:
  - segmental duration: from the onset of high frequency noise to the onset of periodicity in the vowel
  - aspiration duration: from the onset of a distributed spectrum with low frequency noise after the sibilant fricative to the onset of periodicity
  - f0: average over the first three pitch points in the vowel resulting from the default autocorrelation method used by Praat
  - F1: at the first visible glottal cycle
  - intensity: at the beginning of each of the first ten glottal cycles as well as across the whole vowel
  - spectral tilt: H1-H2 in a spectrum of the first four glottal cycles
  - vowel length: from the first glottal cycle to the end of visible periodicity

Experiment 1: Results

- Comparison of 15 tokens each of /sʌ/ and /sʰʌ/ reveals significant differences in…
  - segmental duration (t = -3.157, df = 4, p < .05)
  - aspiration duration (t = 4.554, df = 4, p < .05)
  - F1 onset (t = 4.682, df = 4, p < .01)
  - intensity buildup (t = 3.653, df = 8, p < .01)
  - spectral tilt (t = 5.412, df = 14, p < .001)
  …but not in f0 onset, vowel length, or average intensity (p > .05)
Outline
1. Background and research questions
2. Experiment 1: Production of /sʰa/ and /s*a/
3. Experiment 2: Perception of /sʰa/ and /s*a/
4. Experiment 3: Production of /sʰu/ and /s*u/
5. Experiment 4: Perception of /sʰu/ and /s*u/
6. Implications of the results
7. Conclusions

Experiment 2: Variables
• Concentration in the literature on the role of VOT and f0 as cues to plosive contrast (also some mention of F1 and F2, cf. Park 2002)
• Variables examined in this experiment:
  – segmental duration
  – aspiration duration
  – f0 onset
  – ‘vowel quality’ (i.e. F1 onset, intensity buildup, spectral tilt)

Experiment 2: Design
• 4 within-groups factors with levels as follows:

<table>
<thead>
<tr>
<th>Factor</th>
<th>Range</th>
<th>Levels</th>
<th>Values of Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEG. DURATION</td>
<td>125 ms ~ 250 ms</td>
<td>4</td>
<td>125 ms</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>165 ms</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>205 ms</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>250 ms</td>
</tr>
<tr>
<td>ASP. DURATION</td>
<td>10 ms ~ 60 ms</td>
<td>3</td>
<td>10 ms</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>35 ms</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>60 ms</td>
</tr>
<tr>
<td>F0 ONSET</td>
<td>145 Hz ~ 185 Hz</td>
<td>5</td>
<td>145 Hz</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>155 Hz</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>165 Hz</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>175 Hz</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>185 Hz</td>
</tr>
<tr>
<td>F1 ONSET / INTENSITY BUILDUP / SPECTRAL TILT</td>
<td>2</td>
<td>high/gradual/sharp/low/sharp/shallow</td>
<td></td>
</tr>
</tbody>
</table>

Experiment 2: Stimuli
• Speaker 2’s recordings were used for the stimuli
• 120 critical stimuli and 120 filler stimuli for a total of 240 stimuli
• Pretest stimuli: /ʤa/ ‘sleep’, /ʧ*a/ ‘salty’, /ʧʰa/ ‘tea’

Experiment 2: Methods
• Subjects: 16 native speakers of Korean with normal hearing
  – asked to take a test in which they would listen to words and identify them by clicking buttons on a computer screen
  – told that the goal of the experiment was to see how emotional speech affected intelligibility
  – could listen to each stimulus only once and did not hear the following stimulus until they made a decision about the current one

• Stimuli presented to subjects via Praat 4.2.17 on a Sony Vaio PCG-TR5L laptop computer over Direct Sound EX-29 noise reduction headphones
  – screen display contained six buttons labeled in Korean orthography
  – responses made via mouse by clicking the right button
  – stimuli arranged into a different random order for each subject and presented in four blocks of 60 stimuli, with a one-second ISI
  – experiment lasted approximately 20 minutes in all, and subjects were compensated ~USD 5 for their time
Experiment 2: Results

<table>
<thead>
<tr>
<th>Factor</th>
<th>Asymptotic significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1 ONSET/INTENSITY BUILDUP/</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>SPECTRAL TILT</td>
<td></td>
</tr>
<tr>
<td>NGD DURATION</td>
<td>p &gt; 0.4</td>
</tr>
<tr>
<td></td>
<td>p &gt; 0.5 /s/a/</td>
</tr>
<tr>
<td>ASP. DURATION</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>p &lt; 0.01 /s/a/</td>
</tr>
<tr>
<td>F0 ONSET</td>
<td>p &gt; 0.9</td>
</tr>
<tr>
<td></td>
<td>p &gt; 0.8 /s/a/</td>
</tr>
<tr>
<td></td>
<td>p &gt; 0.9 /s*a/</td>
</tr>
</tbody>
</table>

- Friedman test indicates significant effects of F1 onset/intensity buildup/spectral tilt and aspiration duration
- Interaction between F1 onset/intensity buildup/spectral tilt and aspiration duration

Remaining Questions

- Do F1 onset, intensity buildup, and spectral tilt all play a role in cuing the laryngeal status of the preceding fricative?
- There is reason to believe that F1 may be the predominant cue (cf. Kluender 1991; Benki 2001, 2005)
- What happens with a vowel where F1 does not have to travel as far (i.e. high vowels)?

Outline

1. Background and research questions
2. Experiment 1: Production of /s*a/ and /s+a/
3. Experiment 2: Perception of /s*a/ and /s+a/
4. Experiment 3: Production of /s*u/ and /s*u/
5. Experiment 4: Perception of /s*u/ and /s*u/
6. Implications of the results
7. Conclusions

Experiment 3: Variables and Methods

- Same as Experiment 1
- 3 of 5 subjects were the same
- Comparison of /s*u/ ‘number’ and /s*u/ ‘cook’

Experiment 3: Results

- Comparison of /s*u/ and /s*u/ reveals differences in:
  - segmental duration ($t = -2.395, df = 4, p = 0.075$)
  - aspiration duration ($t = 8.803, df = 4, p < 0.001$)
  - intensity buildup ($t = -3.357, df = 8, p < 0.05$)
  - spectral tilt ($t = 2.722, df = 4, p = 0.053$)
  - but not in f0 onset, F1 onset, vowel length, or average intensity ($p > .05$)
- i.e. same results as Experiment 1, except for F1

Outline

1. Background and research questions
2. Experiment 1: Production of /s*a/ and /s+a/
3. Experiment 2: Perception of /s*a/ and /s+a/
4. Experiment 3: Production of /s*u/ and /s*u/
5. Experiment 4: Perception of /s*u/ and /s*u/
6. Implications of the results
7. Conclusions
Experiment 4: Variables and Design

- Same 4 within-groups factors, with levels as follows:

<table>
<thead>
<tr>
<th>Factor</th>
<th>Range</th>
<th>Levels</th>
<th>Values of Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEG. DURATION</td>
<td>135 ms ~ 225 ms</td>
<td>4</td>
<td>135 ms, 165 ms, 195 ms, 225 ms</td>
</tr>
<tr>
<td>ASP. DURATION</td>
<td>10 ms ~ 70 ms</td>
<td>3</td>
<td>10 ms, 40 ms, 70 ms</td>
</tr>
<tr>
<td>F0 ONSET</td>
<td>180 Hz ~ 230 Hz</td>
<td>2</td>
<td>190 Hz, 230 Hz</td>
</tr>
<tr>
<td>(F2 ONSET)/INTENSITY BUILDUP/ SPECTRAL TILT</td>
<td>/s/ V ~ /s*/u/ V</td>
<td>2</td>
<td>low/gradual, steep/high/sharp/shallow</td>
</tr>
</tbody>
</table>

Experiment 4: Stimuli and Methods

- Speaker 2’s recordings used for the stimuli
- 48 critical stimuli and 75 filler stimuli for a total of 123 stimuli
- Fillers: /du/, /t*u/, /t*u/, /u/
- Same pretest stimuli as Experiment 2
- Same methods as Experiment 2
- Different subjects: 12 native speakers of Korean with normal hearing

Experiment 4: Results

<table>
<thead>
<tr>
<th>Factor</th>
<th>Asymptotic significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>(F2 ONSET)/INTENSITY BUILDUP/ SPECTRAL TILT</td>
<td>p &lt; 0.05</td>
</tr>
<tr>
<td>SEG. DURATION</td>
<td>p &gt; 0.5, p &gt; 0.8 /s/ V, p &gt; 0.05 /s*/u/ V</td>
</tr>
<tr>
<td>ASP. DURATION</td>
<td>p &lt; 0.001, p &lt; 0.001 /s/ V, p &lt; 0.001 /s*/u/ V</td>
</tr>
<tr>
<td>F0 ONSET</td>
<td>p &lt; 0.001, p &lt; 0.001 /s/ V, p &lt; 0.001 /s*/u/ V</td>
</tr>
</tbody>
</table>

- Like Experiment 2, significant effects of intensity buildup/spectral tilt and aspiration duration
- Unlike Experiment 2, significant effect of f0 onset
- Unlike Experiment 2, very high ‘error’ rate (53% vs. 5%)

Implications: Korean Phonetics

- Contra Cho et al. (2002) and others, the fortis fricative was found to be much longer than the non-fortis fricative, while no difference was found in f0 onset
- Results of previous studies regarding aspiration duration, F1 onset, intensity buildup, and voice quality were confirmed

Implications: Perception

- Adds to the results of Yoon (2002) and M.-R. Kim et al. (2002), who looked at the perception of fricatives and cross-spliced plosives
- Phonetic differences cueing a contrast are not necessarily utilized
- Vocalic cues are higher ranked than consonantal cues
- F1 trajectory is critical in the perception of phonation/laryngeal state

Outline

1. Background and research questions
2. Experiment 1: Production of /sʰa/ and /s*a/
3. Experiment 2: Perception of /sʰa/ and /s*a/
4. Experiment 3: Production of /sʰu/ and /s*u/
5. Experiment 4: Perception of /sʰu/ and /s*u/
6. Implications of the results
7. Conclusions
Implications: Laryngeal Typology

- Jansen’s (2004) laryngeal categories for plosives:
  - unaspirated prevoiced lenis (e.g. Spanish /b/)
  - unaspirated passively voiced lenis (e.g. English /b/)
  - unaspirated voiceless fortis (e.g. Spanish /p/)
  - aspirated voiceless fortis (e.g. Burmese /ph/)

- Korean’s plosives fit well into this typology:
  - Korean ‘lenis’ = unaspirated passively voiced lenis
  - Korean ‘fortis’ = unaspirated voiceless fortis
  - Korean ‘aspirated’ = aspirated voiceless fortis

Implications: Laryngeal Typology

- Jansen’s basic laryngeal categories for fricatives:
  - unaspirated voiceless fortis (e.g. English /s/)
  - unaspirated (pre)voiced lenis (e.g. English /z/)

- Jansen (2004): “aspirated fricatives only seem to occur in languages that already have distinctively voiced and plain voiceless fricatives” (cf. Burmese)
  - Korean does not fit this description

Implications: Laryngeal Typology

- Korean’s fricative contrast does not fit into Jansen’s typology
  - one fricative ➔ unaspirated voiceless fortis
  - non-fortis fricative is neither prevoiced nor passively voiced

- Either:
  - (i) the non-fortis fricative is aspirated voiceless fortis,
  - (ii) this fricative does not fit into any of Jansen’s four categories

- Analysis (ii) is superior: the non-fortis fricative is aspirated voiceless lenis

Outline

1. Background and research questions
2. Experiment 1: Production of /sʰa/ and /sʰ*ə/
3. Experiment 2: Perception of /sʰa/ and /sʰ*ə/
4. Experiment 3: Production of /sʰu/ and /sʰ*u/
5. Experiment 4: Perception of /sʰu/ and /sʰ*u/
6. Implications of the results
7. Conclusions

Conclusions

- Korean’s two sibilant fricatives differ in duration, aspiration, F1 onset, intensity buildup, and voice quality of the vowel
- F1 onset, intensity buildup, and voice quality are important cues to this contrast, outranking consonantal cues, with F1 being perhaps the most important
- Korean constitutes an exception to Jansen’s (2004) laryngeal typology in having a fricative contrast without a voiced member
- Korean contrast may be typologically unique and necessitate the addition of an aspirated voiceless lenis category to Jansen’s typology

Selected References