Consonant and Tone Interaction in Cantonese

Holman Tse (Department of Linguistics)
hbt3@pitt.edu
University of Pittsburgh
Dietrich School of Arts & Sciences Grad Expo
March 22, 2012
Introduction

- Two subfields of Linguistics
  1. Phonetics: the study of the universal properties of human speech
  2. Phonology: the study of how sounds are systematically organized in individual languages
- Often thought of as distinct subfields, but the interrelationship between them is itself an active research area
Project Goals

- This project explores this interrelationship by addressing how consonants and tone interact with each other.
- Do they do so in a universally similar way or does it depend on the specific language?
Specific Research Questions:

1. Does tone affect the length of stop consonants in Cantonese, a language with 6 tones (4 of which are relevant)?
2. If so, how?
3. Is this purely a phonetic effect (based on universal constraints) or is it also influenced by (language-specific) phonological contrasts?
Stop Consonants

- Sounds produced by a closure somewhere in the mouth. Ex: /p/, /b/, /t/, /d/, /k/, and /g/.
- Put a piece of paper in front of your mouth.
- Say the words <pan> and <span>.
- The /p/ in <pan> is aspirated, but the /p/ in <span> is not.
- **Phonetically** different but not **phonologically** different for English speakers.
- But in some languages, they ARE **phonologically** different.
Voice Onset Time (VOT)

- Acoustic measurement of stops (roughly equivalent to how long they are produced)
- Three types of stops defined by VOT:
  - Voiced: VOT < 0
    - Ex: /b/
  - Voiceless Unaspirated (short-lag): small positive VOT
    - Ex: /p/
  - Voiceless Aspirated (long-lag): large positive VOT
    - Ex: /pʰ/
- VOT values (phonetic) can vary without changing stop categories (phonological)
F0 vs Pitch vs Tone

- F0 (fundamental frequency) is phonetic
- Pitch is perceptual
- Tone is phonological
  - not based on absolute F0 values but rather on relative F0 values that vary between speakers
  - not measured directly
## F0 (phonetic) vs Tone (phonological) Illustrated

<table>
<thead>
<tr>
<th>Fo Contour</th>
<th>English</th>
<th>English Meaning</th>
<th>Tone</th>
<th>Mandarin Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>High level</td>
<td>Ma.</td>
<td>Declarative</td>
<td>55</td>
<td>‘mother’</td>
</tr>
<tr>
<td>Rising</td>
<td>Ma?</td>
<td>Questioning /uncertainty</td>
<td>25</td>
<td>‘hemp’</td>
</tr>
<tr>
<td>Falling</td>
<td>Ma!</td>
<td>Anger</td>
<td>51</td>
<td>‘to scold’</td>
</tr>
</tbody>
</table>

- English and Mandarin pronunciations are **phonetically** identical, but only in Mandarin is this considered tonal and **phonological**.
Methodology

6 subjects (5 male, 1 female)
- native Hong Kong Cantonese speakers in their early 20’s
- Less than 4 years in the U.S. at the time of recording.

Recordings
- made with solid state recorder in sound-proof booth
- 20 words spoken in a phrase
- 10 repetitions each, 200 samples per subject, but only 4 words (40 samples/speaker) used for present study, the rest used for different experiments/distraction
Sample VOT segmentation

vertical striation marks stop burst

aspiration begins

regular periodicity begins

onset of voicing begins
Sample Textgrid labels

52.395481 0.055312 (18.079 / s) 52.450793

0.5967

0.47

5000 Hz

3525 Hz

0 Hz

1

2

3

4

5

6

52.370792 52.370792

Visible part 0.349890 seconds

Total duration 557.392000 seconds

Ref

Code

Word

Comments

VOT

VL

pha55-01

na

VOT

VL

(34/401)

(401)
Results
### Aspirated Stops

<table>
<thead>
<tr>
<th>Tone</th>
<th>Normalized VOT Avg</th>
<th>Std. Dev. of Normalized VOT</th>
<th>VOT average (in seconds)</th>
<th>Std. Dev. of VOT</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>55</td>
<td>0.2487</td>
<td>0.06648</td>
<td>0.0627</td>
<td>0.19569</td>
<td>60</td>
</tr>
<tr>
<td>33</td>
<td>0.2619</td>
<td>0.06912</td>
<td>0.0657</td>
<td>0.02152</td>
<td>60</td>
</tr>
<tr>
<td>25</td>
<td>0.3002</td>
<td>0.06110</td>
<td>0.0771</td>
<td>0.01932</td>
<td>60</td>
</tr>
<tr>
<td>21</td>
<td>0.3240</td>
<td>0.07782</td>
<td>0.0785</td>
<td>0.02346</td>
<td>60</td>
</tr>
<tr>
<td>All Tones</td>
<td>0.2837</td>
<td>0.07476</td>
<td>0.0701</td>
<td>0.02202</td>
<td>240</td>
</tr>
</tbody>
</table>

Repeated Measure ANOVA test for tone and Normalized VOT for all tokens:  $p < 0.001$, significant
ANOVA test for tone and actual VOT:  $p < 0.001$, significant
### Pair-wise analysis: aspirated stops

<table>
<thead>
<tr>
<th>Tonal Pairs</th>
<th>p =</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>55 &amp; 25</td>
<td>0.1118</td>
<td>n.s.</td>
</tr>
<tr>
<td>55 &amp; 33</td>
<td>0.9816</td>
<td>n.s.</td>
</tr>
<tr>
<td>55 &amp; 21</td>
<td>0.0046</td>
<td>**</td>
</tr>
<tr>
<td>25 &amp; 33</td>
<td>0.1680</td>
<td>n.s.</td>
</tr>
<tr>
<td>25 &amp; 21</td>
<td>0.8133</td>
<td>n.s</td>
</tr>
<tr>
<td>33 &amp; 21</td>
<td>0.0219</td>
<td>*</td>
</tr>
</tbody>
</table>

**General grouping:** 55, 33, (25) < (25), 21
# Tone vs F0 correlations

<table>
<thead>
<tr>
<th>Subject</th>
<th>ANOVA p value</th>
<th>ANOVA significance</th>
<th>Pearson Coefficient</th>
<th>p value of correlation</th>
<th>Correlation Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (male)</td>
<td>&lt; 0.000</td>
<td>**</td>
<td>-0.053</td>
<td>0.744</td>
<td>n.s.</td>
</tr>
<tr>
<td>2 (male)</td>
<td>0.001</td>
<td>**</td>
<td>-0.326</td>
<td>0.040</td>
<td>*</td>
</tr>
<tr>
<td>3 (male)</td>
<td>&lt;0.000</td>
<td>**</td>
<td>-0.350</td>
<td>0.027</td>
<td>*</td>
</tr>
<tr>
<td>4 (male)</td>
<td>0.003</td>
<td>**</td>
<td>-0.001</td>
<td>0.997</td>
<td>n.s.</td>
</tr>
<tr>
<td>5 (male)</td>
<td>0.02</td>
<td>*</td>
<td>-0.313</td>
<td>0.049</td>
<td>*</td>
</tr>
<tr>
<td>6 (female)</td>
<td>0.469</td>
<td>n.s.</td>
<td>0.129</td>
<td>0.426</td>
<td>n.s.</td>
</tr>
<tr>
<td>All Subjects</td>
<td>&lt;0.000</td>
<td>**</td>
<td>-0.210</td>
<td>0.001</td>
<td>*</td>
</tr>
</tbody>
</table>

Correlation test of actual VOT for all subjects: p = 0.249, n.s.
## Effects of VOT on Tone

<table>
<thead>
<tr>
<th>Language</th>
<th>Stop categories</th>
<th>Reported VOT effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kera (Chadic)</td>
<td>Voiceless (historically voiced/voiceless)</td>
<td>L &lt; M &lt; H</td>
</tr>
<tr>
<td>Mazatec</td>
<td>Pre-nasalized, short-lag, long-lag</td>
<td>L &lt; H</td>
</tr>
<tr>
<td>Shanghainese</td>
<td>Voiced, short-lag, long-lag</td>
<td>LM, MM &lt; H, HL</td>
</tr>
<tr>
<td>Korean (younger speakers)</td>
<td>Short-lag, medium-lag, long-lag</td>
<td>L &lt; H</td>
</tr>
<tr>
<td>Mandarin</td>
<td>Short-lag, long-lag</td>
<td>213, 25 &gt; 55, 51</td>
</tr>
<tr>
<td>Hakka</td>
<td>Short-lag, long-lag</td>
<td>Short tones &lt; long tones</td>
</tr>
<tr>
<td>Taiwanese</td>
<td>(Voiced?), Short-lag, long-lag</td>
<td>LL, LR, LF &gt; HL, HF</td>
</tr>
<tr>
<td>Cantonese</td>
<td>Short-lag, long-lag</td>
<td>21 &gt; 55, 33</td>
</tr>
</tbody>
</table>
Discussion/Conclusion

- Tone has a significant effect on VOT for aspirated stops (w/o a loss in phonological contrasts)
- The effect is roughly (but not exactly) inversely correlated with Fo
- Tone is a stronger predictor of VOT than Fo
- A comparison with other languages suggests that this effect is language-specific
- VOT differences corresponding to tonal differences appear to enhance phonological contrasts between tonal categories