The Role of Voice Quality in Shanghai Tone Perception

Jiayin Gao\textsuperscript{1}, Pierre Hallé\textsuperscript{1,2}
1. Laboratoire de Phonétique et Phonologie (Paris 3–CNRS);
2. Laboratoire Mémoire et Cognition (Paris 5–INSERM)
Tones in City Shanghai Chinese

**Tone contour**

<table>
<thead>
<tr>
<th>Tone register</th>
<th>falling</th>
<th>rising</th>
<th>short and glottalized</th>
</tr>
</thead>
<tbody>
<tr>
<td>yin (high)</td>
<td>T1</td>
<td>T2</td>
<td>T4</td>
</tr>
<tr>
<td>yang (low)</td>
<td>T3</td>
<td>T5</td>
<td></td>
</tr>
</tbody>
</table>

\[\text{¢34 vs. ¢23:} \]

胆 ‘gallbladder’ vs. 蛋 ‘egg’

\[\text{yin (high) — yang (low)}\]

○ short and glottalized — ● long

F0 contours of the five lexical tones from a male speaker aged 24.
Tones in City Shanghai Chinese

F0 contours of the five lexical tones from a male speaker aged 24.

<table>
<thead>
<tr>
<th>Tone register</th>
<th>falling</th>
<th>rising</th>
<th>short and glottalized</th>
</tr>
</thead>
<tbody>
<tr>
<td>yin (high)</td>
<td>T1</td>
<td>T2</td>
<td>T4</td>
</tr>
<tr>
<td>yang (low)</td>
<td></td>
<td>T3</td>
<td>T5</td>
</tr>
</tbody>
</table>

\(\text{te}34\) vs. \(\text{te}23\):
胆 ‘gallbladder’ vs. 蛋 ‘egg’
Low tone and breathy voice: production

- Tone is multidimensional: pitch, intensity, duration, voice quality, etc.
- Onsets in low tone syllables are described as breathy:
  - impressionistic descriptions (Karlgren, 1915–1926: 260; Liu, 1923; Chao, 1928).
  - experimental investigations (Cao & Maddieson, 1992; Chen, 2011; Ren, 1988, but see Gao et al., 2010).
- Cross-age and cross-gender difference
  - The phonation difference: elderly > young; male > female (Gao & Hallé, 2013a)
  - It suggests a trend towards loss of breathiness in production.
Perception of Shanghai tones

- Rarely investigated in the literature (see Cao, 1987; Ren, 1992; Gao & Hallé, 2013b)

- Perhaps the only study on the role of breathiness in tone perception (Ren, 1992):
  - main findings: breathy voice is used as a perceptual cue to low tone identity
  - some methodological shortcomings

- Shanghai Chinese has been evolving at a fast rate
Goal of the study:
perceptual aspect of the breathiness

Questions:

- Is phonation difference perceived as a secondary cue to tone identity?, i.e., Does breathiness bias tone perception towards the low tone category?

- If a redundant cue tends to disappear in production, is it supposed to have already lost its perceptual function?
Method: identification test

- **Prediction**
  - If voice quality is perceived as a secondary cue to tone identity, breathy stimuli should bias tone perception toward the low tone category.

- **Identification (2AFC)**
  - along high (T2) – low (T3) tone continua
  - modal and breathy stimuli
  - between two T2–T3 minimal pair choices illustrated by two Chinese characters, e.g., 胆 (T2) ‘gallbladder’ vs. 蛋 (T3) ‘egg’
    - Block 1: 160 trials with synthesized stimuli constructed with VocalTractLab 2.1 (Birkholz, 2012)
    - Block 2: 192 trials with modified natural stimuli
Method: material

- **Stimuli**
  - Voice quality: *modal*, *breathy*
  - Onset: zero, p, t, f, s; (plus /m/ for natural stimuli)
  - Rime: ɛ
  - Tone: T2–T3 continua (8 F0–equidistant steps)

<table>
<thead>
<tr>
<th></th>
<th>zero</th>
<th>stop</th>
<th>fricative</th>
<th>nasal</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2 (Yin)</td>
<td>ɛ 爱</td>
<td>ㄆ板</td>
<td>ㄈ反, ㄕ伞</td>
<td>ㄇ美</td>
</tr>
<tr>
<td>T3 (Yang)</td>
<td>ɛ 咸</td>
<td>ㄆ办, ㄆ台</td>
<td>ㄈ烦, ㄕ笋</td>
<td>ㄇ梅</td>
</tr>
</tbody>
</table>

- **Participants**
  - 16 native speakers of Shanghai Chinese (5 M, 11 F), mean age 22 (18–26)
  - 1 male’s data on synthesized stimuli and 2 males’ data on natural stimuli were discarded
Method: tone continua

- normalized intensity (80 dB)
- normalized C and V duration between each T2–T3 pair: tone perception is affected by segmental duration (Gao & Hallé, 2013b)
- H1–H2 was measured: voice quality differences were validated

synthesized  natural

modal

breathy
Results:
Low tone responses identification curves

Low tone identification curve along the T2–T3 continua, according to the voice quality.
Results:
Low tone responses identification curves according to onset for synthesized stimuli
Results:
Low tone responses identification curves according to onset for natural stimuli

\[
\begin{align*}
\text{modal } [\varepsilon] & \quad \text{breathy } [\varepsilon] \\
\text{modal } [p\varepsilon] & \quad \text{breathy } [p\varepsilon] \\
\text{modal } [f\varepsilon] & \quad \text{breathy } [f\varepsilon] \\
\text{modal } [m\varepsilon] & \quad \text{breathy } [m\varepsilon] \\
\text{modal } [t\varepsilon] & \quad \text{breathy } [t\varepsilon] \\
\text{modal } [s\varepsilon] & \quad \text{breathy } [s\varepsilon]
\end{align*}
\]
Results:
50% low tone boundary location (Best & Strange, 1992)

<table>
<thead>
<tr>
<th>Onset</th>
<th>synthesized</th>
<th>natural</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Breathy</td>
<td>Modal</td>
</tr>
<tr>
<td>zero</td>
<td>1.63</td>
<td>1.77</td>
</tr>
<tr>
<td>p</td>
<td>2.17</td>
<td>** 3.00</td>
</tr>
<tr>
<td>t</td>
<td>2.44</td>
<td>* 3.20</td>
</tr>
<tr>
<td>f</td>
<td>4.47</td>
<td>5.11</td>
</tr>
<tr>
<td>s</td>
<td>2.40</td>
<td>** 3.07</td>
</tr>
<tr>
<td>m</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Mean</td>
<td>2.60</td>
<td>** 3.21</td>
</tr>
</tbody>
</table>

Paired t test:
* p<.05; ** p<.01
Results:
Percentage of low tone responses

GLM: “voice quality” effect on pooled binomial data over onsets, $p<.0001$ for both stimulus types; for each onset (in the figures):  * $p<.05$;  ** $p<.01$
Results: Response times

**synthesized**

ANOVA: $p = .057$

**natural**

ANOVA: $p < .01$
Is phonation difference perceived as a secondary cue to tone identity?

Yes: identification function; response rates; response times

If a redundant cue tends to disappear in production, is it supposed to have already lost its perceptual function?

Not necessarily. Voice quality is perceived as a secondary cue by young listeners, but it starts to be lost in production by young speakers.
No effect on nasal onset syllables: Why?

- Breathiness never existed on low tone nasal onset syllables. However, it did, according to the descriptions of Rose (1989, 2002) and Yip (1980)

- or, breathiness existed on nasal onsets, but had no perceptual function

- or, breathiness is losing its perceptual function, starting with the nasal onsets.

- Loss of breathiness starting with nasal onsets in production has been observed in Yue (Cantonese) dialects (Tsuji, 1977 [cited in Yip, 1980]).
Future work

- Investigate the relative perceptual weights of pitch, duration and voice quality in tone perception
- Look further into the nasal onset issue, for example by looking at other tone and voice—quality languages

Acknowledgement
- to Peter Birkholz, the author of VocalTractLab, for providing us useful information on how to use the software


