Mandarin Tone Speech Recognition in Binaural-Bimodal Patients

Yongxin Li, Xingmei Wei, Biao Chen, Yue Gong
Department of Otolaryngology, Head and Neck Surgery
Beijing Tongren Hospital, Capital Medical University
Backgrounds

Binaural-Bimodal

- One ear with cochlear implantation and the opposite ear with hearing aid

Indication

- One ear with cochlear implantation and the opposite ear with residual hearing
Backgrounds

Unilateral CI

- **deficiency**
  - Poorer speech recognition in noisy backgrounds
  - Poorer music recognition ability
  - Poorer sound localization

Bimodal hearing

- **advantage**
  - Binaural summation effect
  - Head shadow effect
  - Binaural suppression effect
  - Recognition for F0
Objectives

- Bimodal benefit for tone recognition
- The influence of noise for tones and vowels
- To guide the clinical work of CI and the rehabilitation
- To develop coding strategy

ā á ă ā ă
ō ō ă ō ̀
Methods

Criteria

- With cochlear implantation in one ear and hearing aid in the opposite ear
- Above 6 years old
- Communication mainly depends on hearing and oral

General condition

- Age 8→63, on average 20
- M: 8  F: 6
- 13 used HA pre-op, 1 post-op
- CI age: 3→59, on average 21
- Prelingual: 11  Post-lingual: 3
## Patients list

<table>
<thead>
<tr>
<th>Num</th>
<th>Gender</th>
<th>Age</th>
<th>Etiology</th>
<th>Pre/post-lingual</th>
<th>Deaf time pre-op (years)</th>
<th>Operation Age (years)</th>
<th>Side</th>
<th>CI time (years)</th>
<th>HA time (years)</th>
<th>Bimodal time (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M</td>
<td>20</td>
<td>LVAS</td>
<td>Pre</td>
<td>13</td>
<td>13</td>
<td>L</td>
<td>7</td>
<td>15</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>20</td>
<td>ototoxicity</td>
<td>Pre</td>
<td>13</td>
<td>15</td>
<td>R</td>
<td>5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>M</td>
<td>17</td>
<td>LVAS</td>
<td>Pre</td>
<td>12</td>
<td>15</td>
<td>L</td>
<td>1.83</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>20</td>
<td>LVAS</td>
<td>Pre</td>
<td>15</td>
<td>16</td>
<td>R</td>
<td>4</td>
<td>20</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>M</td>
<td>33</td>
<td>LVAS</td>
<td>Pre</td>
<td>28</td>
<td>30</td>
<td>R</td>
<td>2.92</td>
<td>30</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>19</td>
<td>Unknown</td>
<td>Pre</td>
<td>12</td>
<td>14</td>
<td>L</td>
<td>5</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>M</td>
<td>30</td>
<td>Unknown</td>
<td>Post</td>
<td>8</td>
<td>29</td>
<td>R</td>
<td>0.58</td>
<td>4</td>
<td>0.58</td>
</tr>
<tr>
<td>8</td>
<td>M</td>
<td>23</td>
<td>Unknown</td>
<td>Pre</td>
<td>19</td>
<td>20</td>
<td>R</td>
<td>3</td>
<td>22</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>F</td>
<td>27</td>
<td>Unknown</td>
<td>Post</td>
<td>3</td>
<td>24</td>
<td>R</td>
<td>3.33</td>
<td>8</td>
<td>3.33</td>
</tr>
<tr>
<td>10</td>
<td>F</td>
<td>19</td>
<td>LVAS</td>
<td>Pre</td>
<td>12</td>
<td>16</td>
<td>L</td>
<td>3</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>F</td>
<td>30</td>
<td>LVAS</td>
<td>Pre</td>
<td>28</td>
<td>28</td>
<td>R</td>
<td>2</td>
<td>28</td>
<td>3</td>
</tr>
<tr>
<td>12</td>
<td>M</td>
<td>8</td>
<td>LVAS</td>
<td>Pre</td>
<td>2</td>
<td>3</td>
<td>L</td>
<td>5</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>13</td>
<td>F</td>
<td>17</td>
<td>Unknown</td>
<td>Pre</td>
<td>16</td>
<td>16</td>
<td>L</td>
<td>1</td>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>F</td>
<td>63</td>
<td>Sudden SNHL</td>
<td>Post</td>
<td>15</td>
<td>59</td>
<td>R</td>
<td>4</td>
<td>9</td>
<td>4</td>
</tr>
</tbody>
</table>
Methods

- Background noise $< 25\text{dB (A)}$, sound intensity $65\text{dB (A)}$
• Pro. Qianjie Fu from UCLA, Los Angeles, Caloifornia
Methods

- **Threshold**: with CIHA、CI、HA under sound field, residual hearing of the HA side (RH)
- **Speech test**: CIHA、CI、HA
  - Quiet
  - Three SNR（10、5、0dB）
  - Two background noise: Ba1 （Ba with tone1）spoken by a man and a woman, respectively
  - Target tone are all spoken by a man
The average threshold of CIHA is lower than HA and RH, but no significant difference with CI only.

Results 1 - Threshold

- The average threshold of CIHA is lower than HA and RH, but no significant difference with CI only.
Except for quiet, there are significant bimodal benefit in all noise environment. When the target and noise are in different F0 (MF), the lower the SNR, the more significant the bimodal benefit. While, when they are the same (MM), the greater the SNR, the more significant the bimodal benefit.

Results 2 - Bimodal Benefit

<table>
<thead>
<tr>
<th>Environment</th>
<th>Percent Correct (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quiet</td>
<td>CIHA</td>
</tr>
<tr>
<td>MM10</td>
<td>90</td>
</tr>
<tr>
<td>MM5</td>
<td>60</td>
</tr>
<tr>
<td>MM0</td>
<td>40</td>
</tr>
<tr>
<td>MF10</td>
<td>10</td>
</tr>
<tr>
<td>MF5</td>
<td>0</td>
</tr>
<tr>
<td>MF0</td>
<td>0</td>
</tr>
</tbody>
</table>

Bimodal Benefit = Percent of CIHA – Percent of HA

![Graph showing bimodal benefit results]

*p-value: 0.57 0.006 0.029 0.046 0.011 0.011 0.000*
Four Tones are all affected by the background noise.

The recognition of tone 3 is best, and tone 3 is the least affected by competing talkers.

The recognition of tone 4 is lowest, and tone 4 is the most susceptible to be affected by competing talkers.
Four vowels are all affected by the background noise.
The recognition of **Ba** is best, and Ba is the least affected by competing talkers, then is Bo.
Discuss1-Threshold

• There is no significant difference between CIHA and CI

Maybe due to

- limited sample size
- Poorer residual hearing of HA side (88.3±12.2dB HL on average)

The threshold of bimodal users mainly depends on CI. Bimodal benefit is not only due to the improvement of hearing, but also because of the information of CI and HA complement each
**Discuss2-Bimodal benefit**

- **Significant bimodal benefit in noise**
  - Need more time cue in noise
  - Tone recognition need more information than vowel and consonant
  - Tone recognition mainly depends on F0, which HA can provided better

- **When target and noise are in different F0, bimodal benefit is better**
  - The lower the SNR, the greater the interference to CI, the greater role the HA plays, so the bimodal benefit better

- **When target and noise are in same F0**
  - The greater the SNR, the lower the interference to HA, so the bimodal benefit better
Discuss 3 - Four tones

- **Tone time**: Tone 3 is the longest, tone 4 is the shortest
  - The longer the tone time, the more the time cue information
- **Pitch**: Tone 3 is falling and then rising, which change biggest
- **Competing talker**: Ba with tone 1
  - The tone time and pitch is similar with tone 4, but more different with tone 3
Discuss4-Four vowels

- /a/ is located in the far end of F1, and the fundamental frequency is higher, so it’s easier to be recognized.
- The bigger the formant gap of F1 and F2, the easier to be recognized.
  - The distance between /o/ and /a/ is longer, so they are easier to be distinguished.

Luo X, Fu QJ. Speaker normalization for Chinese vowel recognition in cochlear implants. IEEE Transactions on Biomedical Engineering, 2005, 52(7): 1358-1361
Conclusion

Bimodal benefit is more significant in noise

When F0 of target and signal are different, the bimodal benefit is better

Tone 3 has the highest recognition rate and is the most stable, while tone 4 is the opposite

Ba has the highest recognition rate and is the most stable, and Bo is less affected by Ba
Thank you!

Email: entlyx@sina.com