THE EFFECT OF CONSONANT-VOWEL BOUNDARY TO SPEECH PERCEPTION IN COCHLEAR IMPLANTS

Nirmal Kumar Srinivasan¹, Hussnain Ali¹, Emily A. Tobey², and Philipos C. Loizou¹
Department of Electrical Engineering¹, Dallas Callier Advanced Hearing Research Center²: The University of Texas at Dallas, Richardson, TX.

INTRODUCTION

- Cochlear Implant (CI) users report difficulty in understanding speech in noisy environments
  - No masking release from modulated maskers [5].
  - Cannot take advantage of temporal gaps
  - Can perform auditory stream segregation tasks [2].
- CI listeners have difficulty in successfully fusing interrupted speech signals into a coherent speech stream [6].
- Speech sounds are broadly classified into vowels and consonants.
- Consonants carry more information than vowels while reading.
- Vowels contribute more to speech intelligibility than consonants
  - Vowel-only sentences led to a 2.1 intelligibility advantage over consonant-only sentences regardless of the type of segmental replacement in normal hearing listeners [1].

METHODS

Listeners: 8 adult CI users participated in this study.

Speech Corpus: Speech stimuli were taken from the AZTIMIT sentences [3]. All lists were equi-intelligible with a mean intelligibility of 73% [4].

LIST CHARACTERISTICS:
- average of 128 words per list (range = 115 to 136 words)
- 18-20 unique talkers per list
- At least 6 male and 6 female talkers per list
- Talkers were from four dialectal regions (New England, Northern, North Midland, and Western)

Speech Processing strategies:
- Segment replacement paradigm on the AZRIMIT speech corpus was used to evaluate the effect of consonant-vowel boundary to speech perception.
- Two processing strategies were created to emphasize the duration of vowels and consonants by presenting different amounts (0%, 40%, 60%, 80%, and 90%) of consonants and vowels.
- First strategy (FVXC) preserved full vowels and presented different amounts of consonants by replacing the consonant centers with either silence or speech shaped noise.
- Second strategy (FCXV) preserved full consonants and presented different amounts of vowels by replacing the vowel centers with either silence or speech shaped noise.
- Clean speech and interrupted speech (3Hz interruption rate, 50% duty cycle) were also presented.
- PDA based research platform was used to present speech stimuli.

RESULTS – Average Speech Intelligibility

<table>
<thead>
<tr>
<th>FV0%V</th>
<th>FV40%V</th>
<th>FV60%V</th>
<th>FV80%V</th>
<th>FV100%V</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%V</td>
<td>40%V</td>
<td>60%V</td>
<td>80%V</td>
<td>100%V</td>
</tr>
<tr>
<td>FULL VOWEL PERCENT CONSONANT</td>
<td>FULL CONSONANT PERCENT VOWEL</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1. Schematic of replacement paradigm depicted for a single VCV (panel a) or CVC (panel b). Horizontal bars indicate vowels, vertical bars indicate consonants and stippled bars indicate replaced portions (silence or speech shaped noise replacement).

RESULTS – Average Speech Intelligibility

<table>
<thead>
<tr>
<th>FVXV_SSN</th>
<th>FVXV_SIL</th>
<th>FVxC_SSN</th>
<th>FVxC_SIL</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%V</td>
<td>40%V</td>
<td>60%V</td>
<td>80%V</td>
</tr>
<tr>
<td>FULL CONSONANT</td>
<td>FULL VOWEL</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2. Mean speech intelligibility for all experimental conditions. The original TIMIT C-V boundary is at 0%V and 0%C. Error bars display standard error of mean.

REFERENCES


CONCLUSIONS

- CI listeners did not tolerate periodic interruptions in continuous speech signal
- Showed evidence to fuse interrupted speech signals into a coherent speech stream
- Vowels contribute more to speech intelligibility as compared to consonants for CI listeners

ACKNOWLEDGEMENTS

Research supported by NIH-NIDCD (R01 DC010494).