Do non-speech contexts mask or contrast with speech targets?

John Kingston, Michael Key, and Sarah Watsky

Linguistics Department, University of Massachusetts, Amherst

0 Introduction

What is perceived?
- Gestures? (Fowler, 1986).

1 Background

Non-speech = speech when:
- Perceived as speech (Best, et al., 1981).
- Psychoacoustically similar (D & K, 89).

2 Context effects

More “ga” / [al]__ than [ar]__:
- Non-speech = speech when:
  - Glide and pure tone than lower.
  - More “ga” responses after higher FM.
- What is perceived?
  - Glide and pure tone than lower.
  - Auditory qualities?
  - Gestures?

3 Masking?

FM & pure tones RMS = [al, ar]
  - b. F3+: F4 closer to F3; F3 20 dB > F1.

4 Forward masking? Not!

Pure tones mask signal at nearby frequencies (Moore, 1997), L&K’s effect at distant frequencies.

5 Exp 1 Match (Kidd, et al., 2005)

Method

Target [da–ga] and Context [al] vs [ar]:
- Bandpass filtered, 1 erb wide.
- Subset of bands; context bands:
  - Matched or mismatched targets.
  - Context bands vocoded > Non-speech.
- Modulate pure tone @ CF
- Context RMS amplitude:
  - +6, 0, -6 dB relative to target.
- Silent gap between context and target:
  - 50 or 150 ms.
- If masking, context effects:
  - Match > mismatch
  - +6 > 0 > -6 dB
  - 50 > 150 ms.

Stimuli

<table>
<thead>
<tr>
<th>Stimuli</th>
<th>Follow me</th>
<th>Match</th>
<th>Mismatch</th>
</tr>
</thead>
<tbody>
<tr>
<td>alga &gt;</td>
<td>match</td>
<td>mismatch</td>
<td></td>
</tr>
<tr>
<td>alda &gt;</td>
<td>mismatch</td>
<td>match</td>
<td></td>
</tr>
</tbody>
</table>

Results

<table>
<thead>
<tr>
<th>l:r x</th>
<th>Est.</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Match</td>
<td>-0.295</td>
<td>-4.409</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Level</td>
<td>-0.019</td>
<td>-2.360</td>
<td>0.018</td>
</tr>
<tr>
<td>Gap</td>
<td>-0.277</td>
<td>-2.838</td>
<td>0.005</td>
</tr>
</tbody>
</table>

6 Exp 2 Off-frequency

Proximity of tone to target endpoints (Viswanathan, 2009):

Method

Move both tones away from endpoints, Context RMS = Target or 2x Target.

Results

<table>
<thead>
<tr>
<th>l:r x</th>
<th>Est.</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Match</td>
<td>-0.144</td>
<td>-2.650</td>
<td>0.009</td>
</tr>
<tr>
<td>Level</td>
<td>-0.144</td>
<td>-0.404</td>
<td>0.687</td>
</tr>
<tr>
<td>Gap</td>
<td>0.528</td>
<td>2.386</td>
<td>0.017</td>
</tr>
</tbody>
</table>

7 Conclusions

Masking hypothesis not supported:
- Exp 1: l:r Match < Mismatch
- Exp 2: Failed to replicate V (09).

Contrast:
- Exp 1: Gross energy distribution.
- Exp 2: Streaming.

References

- More “ga” / [al]__ than [ar]__:
  - Compensation for coarticulation (Mann, 1980):
    - Stop pulled forward by [l], percept pulled back.
    - Spectral contrast (Lotto & Kluender, 1998):
      - Ambiguous F3 perceived as lower after [l]’s high F3 than [r]’s low F3.

4 Forward masking? Not!

Pure tones mask signal at nearby frequencies (Moore, 1997), L&K’s effect at distant frequencies.

3 Masking?

FM & pure tones RMS = [al, ar]
  - b. F3+: F4 closer to F3; F3 20 dB > F1.

More “ga” responses after higher FM glide and pure tone than lower.

Follow me