Effects of clear speech and language background on multimodal perception of English fricatives

Sylvia Cho and Yue Wang
Department of Linguistics, Simon Fraser University, Canada

Allard Jongman and Joan A. Sereno
Department of Linguistics, University of Kansas

Background

- Clearly produced English fricatives involve acoustic modifications that facilitate the auditory intelligibility for native (L1) and non-native (L2) perceivers. Visual–facial cues also contribute to fricative perception. However, clear-speech effects on the visual perception of fricatives have received less attention.
- A comparison of auditory and visual fricative perception can uniquely test how these two input modalities differently affect clear-speech benefits, since sibilant fricatives in English are more auditorily salient than non-sibilants, while non-sibilants are more visually salient.
- Previous research on non-native audio-visual perception suggests ineffective auditory and visual weighting when perceiving fricatives that do not exist in the L1. Such findings motivate examination of how clear-speech effects are associated with intelligibility gains or detractors as a function of L1 background.

Present Study

- Examines the audio-visual perception of clearly versus plainly produced English sibilant and non-sibilant fricatives by native English perceivers, and non-native (Mandarin and Korean) perceivers differing in L1 fricative inventories.

Research Questions

1. Are clear (relative to plain) fricatives more intelligible when visual information is available?
2. Are greater clear-speech benefits found with certain fricatives (i.e., sibilants vs. non-sibilants) as a function of auditory & visual saliency?
3. How does L1 phoneme inventory affect audio-visual clear fricative perception?

Methods

Perceivers:
- 21 native English
- 17 native Mandarin & 30 Korean
- Late intermediate-level learners of English

Stimuli:
- English /f, v, θ, s, z, j, ʒ/ followed by /a/
- Produced by 4 native English speakers (2F) in plain and clear speech (Fig. 1)
- Presented in audio-visual (AV), audio-only (AO), and visual-only (VO) modes (Fig. 2)
- Audio embedded in multi-talker babble noise (sibilants: 9dB SNR; non-sibilants: 0dB SNR)

Comparisons with Mandarin & Korean fricatives:

L1 group comparisons:
- English > Mandarin & Korean in clear AV & AO; English = Mandarin = Korean in clear VO
- Non-natives less accurate than natives in auditory perception but native-like visual perception in clear VO

Clear speech benefits:
- Visual clear-speech benefit for all but sibilants for English group
- Across groups, clear speech benefits sibilants in AO and non-sibilants in VO
- Across fricatives, natives (English) benefit more in AO, non-natives (Mandarin, Korean) benefit more in VO

Modality comparisons:
- AV > AO > VO for all, except clear sibilants: AO > AV > VO in English perceivers, AO = AV > VO in Mandarin & Korean perceivers

Fig. 1. Elicitation of plain and clear speech

Fig. 2. Stimulus presentation

Task:
- 8-alternative forced choice (8AFC)

Acknowledgements

We thank members of the Language and Brain Lab at Simon Fraser University for their assistance in stimulus development and data collection. This project was supported by a research grant from the Social Sciences and Humanities Research Council of Canada (SSHRC Insight Grant 435-2012-1641).

Results

Results (Accuracy)

<table>
<thead>
<tr>
<th>Identification Accuracy (%)</th>
<th>English</th>
<th>Mandarin</th>
<th>Korean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sibilants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plain</td>
<td></td>
<td>!</td>
<td>!</td>
</tr>
<tr>
<td>Clear</td>
<td>!*</td>
<td>!*</td>
<td>!</td>
</tr>
<tr>
<td>AO</td>
<td>!*</td>
<td>!*</td>
<td>!</td>
</tr>
<tr>
<td>VO</td>
<td>!*</td>
<td>!*</td>
<td>!</td>
</tr>
</tbody>
</table>

Results (Place Accuracy)

Accuracy based on place of articulation across voicing

<table>
<thead>
<tr>
<th>Identification Accuracy (%)</th>
<th>English</th>
<th>Mandarin</th>
<th>Korean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sibilants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plain</td>
<td></td>
<td>!</td>
<td>!</td>
</tr>
<tr>
<td>Clear</td>
<td>!*</td>
<td>!*</td>
<td>!</td>
</tr>
<tr>
<td>AO</td>
<td>!*</td>
<td>!*</td>
<td>!</td>
</tr>
<tr>
<td>VO</td>
<td>!*</td>
<td>!*</td>
<td>!</td>
</tr>
</tbody>
</table>

Summary

- Visual clear-speech benefit for all but sibilants for English group
- Across groups, clear speech benefits sibilants in AO and non-sibilants in VO
- Across fricatives, natives (English) benefit more in AO, non-natives (Mandarin, Korean) benefit more in VO

Fig. 3. English, Mandarin, and Korean perceivers’ mean percent accuracy, as a function of Modality, Style, and Sibilance (* = p < 0.05). Error bars indicate standard error.

Fig. 4. English, Mandarin, and Korean perceivers’ mean percent place accuracy (pooled across voicing), as a function of Modality, Style, and Sibilance (* = p < 0.05). Error bars indicate standard error.

Discussion & Conclusions

- Current study extends previous findings, showing visual clear-speech benefit in English fricative perception.
- Results reveal clear-speech effects as a function of audio-visual saliency. Clear speech benefits perception of acoustically salient (sibilant) fricatives in AO & perception of visually salient (non-sibilant) fricatives in VO.
- Clear fricative perception is also modulated by L1 background, where non-natives benefit more from visual clear speech, particularly in receiving fricatives non-existent in their L1 (e.g., labiodentals for Koreans).

References

We thank members of the Language and Brain Lab at Simon Fraser University for their assistance in stimulus development and data collection. This project was supported by a research grant from the Social Sciences and Humanities Research Council of Canada (SSHRC Insight Grant 435-2012-1641).