Introduction

- Typically hearing (TH) listeners can identify the speech target in the presence of interfering maskers with greater accuracy and less perceived effort when are spatially separated.1,2
- Under experimental conditions, behavioral measures like percent correct scores are not designed to also provide insight into the momentary development of the decision-making process that produces the ultimate perceptual judgement.2,4

For this study, we adapted the “visual world paradigm” (VWP) wherein participants’ eye-gaze was monitored as they searched for the image representing the speech target among four on-screen alternatives following presentation of the auditory stimuli.5,6

Objective

To establish normative ranges for the effects of target-masker configuration and signal-to-noise ratio (SNR) on the events unfolding over the time course of target determination. Additionally, we will examine whether these effects vary with trial uncertainty by randomly varying the congruence between auditory and visual stimuli.

Materials and Methods

- **Speech stimuli:** Coordinate Response Measure (CRM) sentences. Figure 1A illustrates sentence syntax.
- **Signal-to-noise ratio (SNR):** target only (i.e., speech in quiet), 0 dB SNR, or -5 dB SNR
- **Target-Masker Configurations** included quiet (target-only), co-located (target and masker presented from the same loudspeaker), and separated (target and masker presented from different loudspeakers).
- Loudspeakers were located at +/- 90° azimuth and 43.5° from the participant’s head.
- **Visual Stimuli** depicting the target and masker number-color combinations were displayed on a computer monitor immediately following auditory stimulation (as described in figure 1B).
- The congruence between the side of the monitor on which the target visual stimulus was drawn and the side of the loudspeaker that produced the target auditory stimulus was counterbalanced across trials.
- Eye gaze was monitored using an SR Research Eyelink 1000 eye tracker with a 1000 Hz sampling rate.
- Each trial consisted of auditory stimulation (figure 1A) followed by visual stimuli presentation (figure 1B) and concluded when the participant clicked the image. They were instructed to select the image corresponding to the call sign “Baron.”

Table 1: Four visual stimulus combinations are shown during each trial

<table>
<thead>
<tr>
<th>Visual Stimulus</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target</td>
<td>Target number drawn in the target color</td>
</tr>
<tr>
<td>Masker</td>
<td>Masker number drawn in the masker color</td>
</tr>
<tr>
<td>Partially Congruent Foil</td>
<td>Target number drawn in the masker color</td>
</tr>
<tr>
<td>Partially Congruent Target</td>
<td>Masker number drawn in the target color</td>
</tr>
</tbody>
</table>

Figure 1A (B): example trial in which visual and auditory stimuli are incongruent

Table: Results

**Figure 2:** Differences in mean target accuracy by target-masker configuration, SNR, and trial congruence for all trials

Accuracy was greatest when targets were presented in quiet but poorest when they were presented with a co-located speech masker.

Spatial separation of the target and masker facilitated greater target accuracy (i.e., spatial release from masking).

**Figure 3:** Differences in proportion of looks to target by target-masker configuration, SNR, and trial congruence for correct trials

The influence of SNR on proportion of looks to target was most noticeable in the co-located target-masker configuration. However, the way SNR affected visual fixation behavior was inconsistent across listeners.

- For this study, we adapted the visual world paradigm to glean insight into the decision-making processes associated with listening for a speech target in a context designed to have auditory information and visual cues be unpredictably associated.
- Preliminary analysis of the trial accuracy results demonstrated the canonical finding of spatial release from masking.
- Findings related to time to decision and looks to the target varied by participant, suggesting individual differences in how participants approach the listening task.
- We will explore these differences through further analyses and additional experimental conditions.

References


Acknowledgements

The authors would like to thank Shelly Godar and Stephanie Seliner for assistance with recruiting and screening participants. This work was supported by NIH-NIDCD grant R01DC003083 to RYL and in part by a core grant from the NIH-NICHD (U54 HD090256) to the Waisman Center.