An effect of eye position in cocktail party listening

Virginia Best, Todd Jennings, Gerald Kidd, Jr.
Department of Speech, Language & Hearing Sciences

Boston University College of Health
& Rehabilitation Sciences: Sargent College
Department of Speech, Language & Hearing Sciences

BACKGROUND

Much is known about the interaction between eye movements and auditory perception:

- eye position influences sound localization (e.g., Jones & Kabanoff 1975; Getzmann 2002)
- auditory spatial resolution is improved in the vicinity of visual fixation (Maddox et al 2014)
- fixating on a stream of tones reduces reaction times for target detection (Pomper & Chait 2017)
- people move their eyes towards the location of a talker that they are attending (e.g., Gopher 1973; Hendrikse et al 2019).

There is not much compelling evidence that eye position influences speech intelligibility (in the absence of lip reading):

- some (but not all) studies examining recall of a target talker in the presence of a distracting talker found an effect of eye position (e.g., Reisberg et al 1981; Wolters and Schiano 1989; Driver & Spence 1994; Spence et al 2000)
- surprisingly little work examining this issue for situations with a lot of masking.

AIM

To determine if there is a measurable effect of eye position in "cocktail party" listening situations.

Our primary motivation was an intuition that eye position may be especially critical in these situations, where there is substantial energetic and informational masking.

A secondary motivation was our interest in "visually guided beamforming" where the eyes are used to steer a highly directional hearing aid (e.g., Kidd 2017; Best et al 2017).

METHODS

Competing four-digit sequences were presented from five loudspeakers at -30, -15, 0, +15, +30 deg azimuth (Fig 1A).
Within a block of trials:

- red visual cue indicated where to **listen**
- green visual cue indicated where to **look** (Fig 1B).

Across blocks, tested different combinations of target position (-30, 0, +30 deg) and eye position (-30, 0, +30 deg) were tested.
The participant's head was stabilized with a headrest and their eyes were tracked continuously with an eye-tracker mounted on eyeglasses (Pupil Labs, Fig 1C).
Responses were scored as the percentage of target digits correctly reported.
Preliminary data are presented from seven young listeners with audiometrically normal hearing.

Fig 1. Experimental set up.

REFERENCES

Best, V., Roverud, E., Streeter, T., Mason, C. R., & Kidd, G. (2017). The benefit of a visually guided beamformer in a dynamic speech task. *Trends Hear., 21*, 2331216517722304.
Driver, J., & Spence, C. (1994). Spatial synergies between auditory and visual attention. In Umilta & Moscovitch (Eds.), *Attention and Performance 15: Conscious and Nonconscious Information Processing* (pp. pp. 311-331). Cambridge: The MIT Press.
Getzmann, S. (2002). The effect of eye position and background noise on vertical sound localization. *Hearing Res., 169*, 130-139.
Gopher, D. (1973). Eye-movement patterns in selective listening tasks of focused attention. *Percept. Psychophys., 14*, 259-264.
Hendrikse, M. M. E., Llorach, G., & Hohmann, V. (2019). Movement and gaze behavior in virtual audiovisual listening environments resembling everyday life *Trends Hear., 23*, 2331216519872362.
Jones, B., & Kabanoff, B. (1975). Eye movements in auditory space perception. *Percept. Psychophys., 17*, 241-245.
Kidd, G. (2017). Enhancing auditory selective attention using a visually guided hearing aid. *J. Speech Lang. Hear. Res, 60*(10), 3027–3038.
Maddox, R. K., Popisil, D. A., Stecker, G. C., & Lee, A. K. C. (2014). Directing eye gaze enhances auditory spatial cue discrimination *Curr. Biol., 24*(7), 748-752.
Pomper, U., & Chait, M. (2017). The impact of visual gaze direction on auditory object tracking. *Sci. Rep., 7*(1), 4640.
Reisberg, D., Scheiber, R., & Potemken, L. (1981). Eye position and the control of auditory attention. *J. Exp. Psych. Human Percept. Perform., 7*(2), 318-323.
Spence, C., Ranson, J., & Driver, J. (2000). Cross-modal selective attention: On the difficulty of ignoring sounds at the locus of visual attention. *Percept. Psychophys., 62*, 410-424.
Wolters, N. C. W., & Schiano, D. J. (1989). On listening where we look: The fragility of a phenomenon. *Percept. Psychophys., 45*, 184-186.

RESULTS

Speech Scores
Performance was better for lateral targets than central targets (Fig 2, compare panels).
Performance was better for on-target versus off-target fixation (Fig 2, black versus white symbols within a panel).
Effect of eye position was more consistent for central targets. Possible explanations for this:

- task difficulty is higher for central targets
- lateral talkers are acoustically more salient than central talkers
- lateral fixation may engage visual attention more strongly than "default" central fixation.

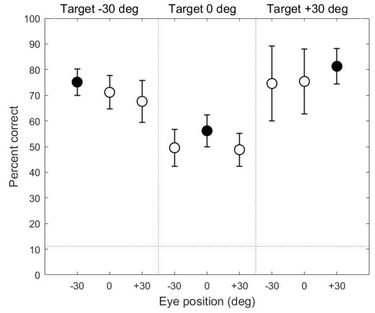

Fig 2. Mean speech scores for different combinations of target and eye position. Error bars show across-subject standard deviations.

Error Patterns
Responses overall showed a characteristic error pattern with a tendency to report digits arising from azimuths adjacent to the target (Fig 3, top row).
Additional errors caused by off-target fixation were rather non-systematic (Fig 3, bottom row). In particular, there was no evidence for a greater tendency to report digits arising from the fixation location.

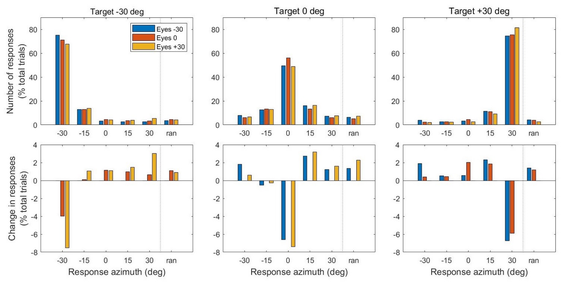

Fig 3. Top row: Distribution of responses across azimuths (ran=random responses). Bottom row: Change in response distributions relative to on-target fixation.

EYE TRACKING

Eye-tracker data suggested that fixation was generally good.
Eye position distributions were tighter for central vs. lateral fixation due to both listener and technical factors (Fig 4, compare panels).
There was little evidence for systematic eye movements to the target loudspeaker (Fig 4, compare histograms within a panel).

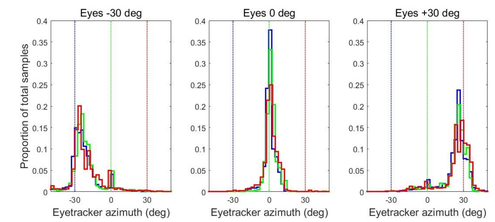

Fig 4. Pooled eye position distributions. Each panel shows one fixation position and the three histograms within a panel are for different target positions.

CONCLUSION

Preliminary results demonstrate a measurable (but modest) effect of eye position on speech intelligibility in a cocktail party mixture.

Ongoing work will determine to what extent the effect depends on:

- the spatial configuration of the competing talkers
- the relative salience of the competing talkers
- the presence of informational masking.

Work supported by NIH-NIDCD (R01 DC015760 and DC013286)