Talker head orientation discrimination using only auditory cues

David Frazier II, Brian Monson
Department of Speech and Hearing Science, College of Applied Health Sciences, University of Illinois at Urbana-Champaign

Introduction

The ability of humans to localize sound sources, including speech, has been studied extensively, however less is known about the human ability to determine the physical orientation of a given sound source. It has been proposed that determining a talker's head orientation (or facing angle) is important for ascertaining whether one is the intended recipient of a warning call or other utterance. We assessed listeners' ability to detect changes in talker head orientation.

Some studies indicate that extended high frequencies (EHF) are useful for some auditory tasks, but it is widely believed they play little to no role in speech perception. Directivity patterns of human speech and voice radiating from the mouth indicate that high-frequency radiation is increasingly directional. Thus we predicted that EHF hearing in humans improves talker head orientation discrimination ability.

Aim

To asses human listeners' ability to detect changes in talker head orientation

To examine if access to EHFs improves talker head discrimination ability

Directional Recordings

High fidelity anechoic directional recordings were obtained previously using multi-channel simultaneous recordings with microphones surrounding a subject in the horizontal plane.

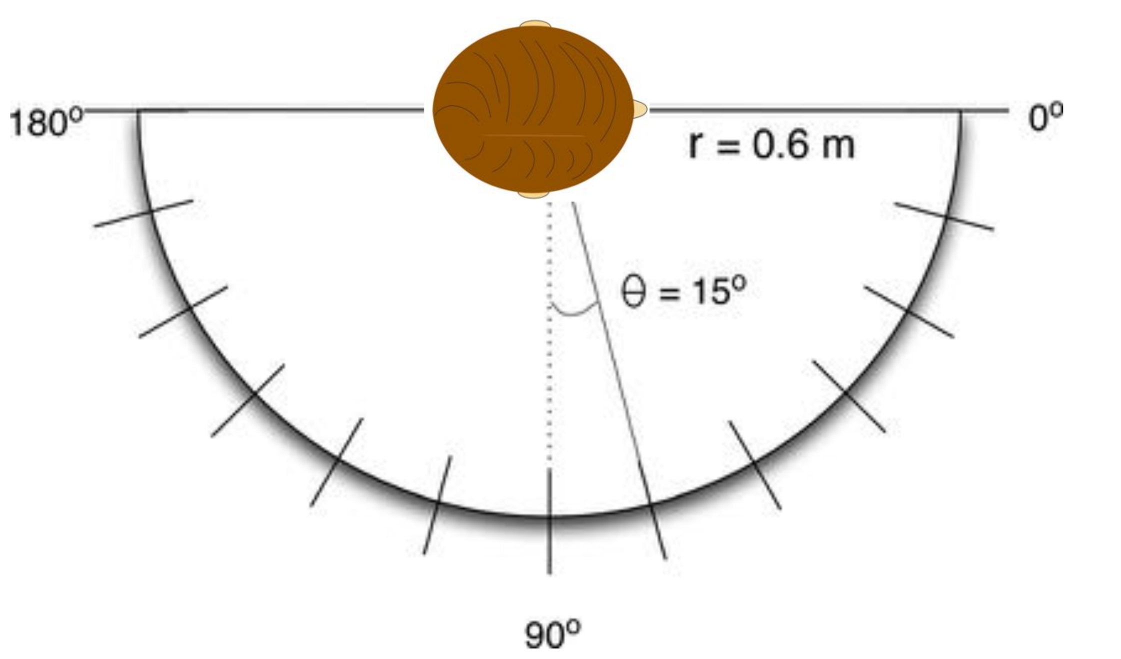

Figure 1: Depiction of the semicircular recording apparatus and microphones placed at 15° increments from directly in front of the subject (0°) to directly behind the subject (180°).

Directional Recordings (continued)

Horizontal plane directivity patterns for speech show that low frequencies radiate omnidirectionally, whereas EHFs are highly directional.

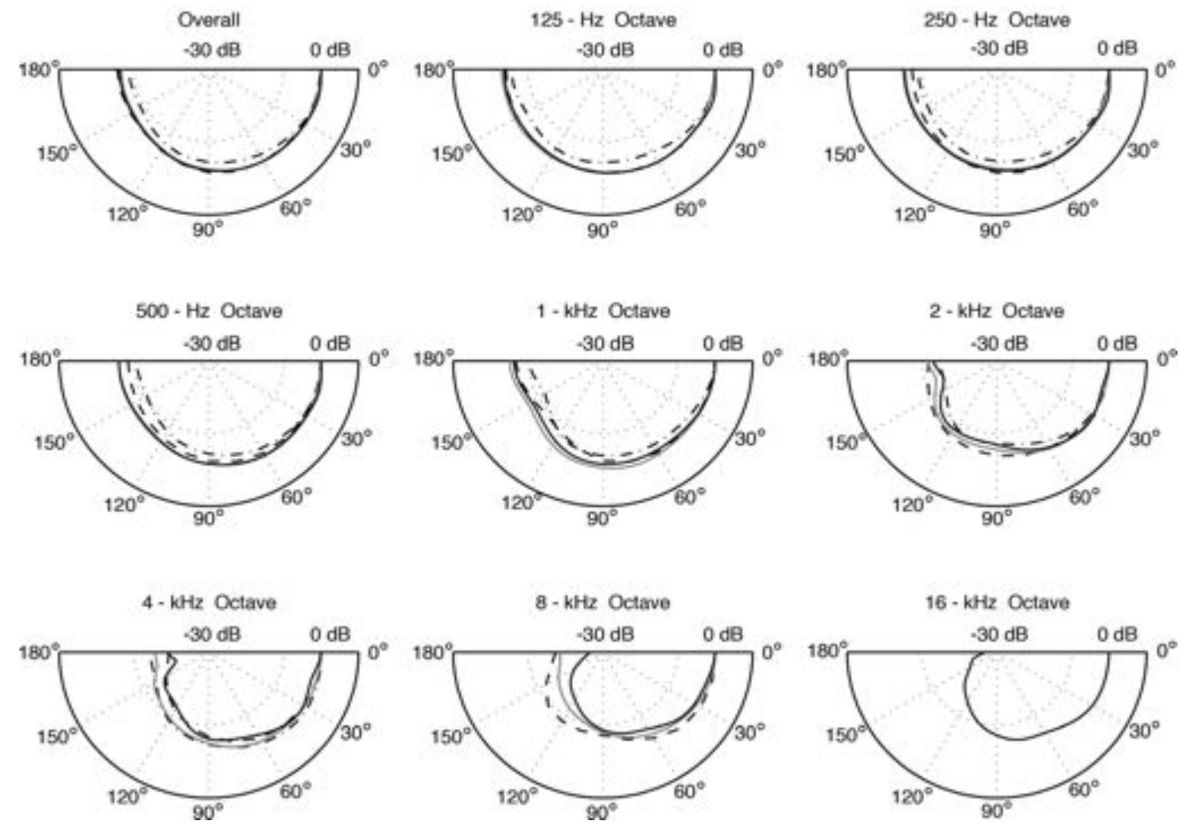

Figure 2:Directivity patterns showing overall and octave band directivity for normal speech. Radius indicates sound level at that location.

Method

Stimuli:

- Recorded phrase "amend the slower page"
- Recorded phrase by two male and two female talkers
- Stimuli for low-pass filtered condition generated by filtering a 32-pole Butterworth filter with cut-off frequency of 8kHz

Listener Subjects:

- 18 normal-hearing native English-speaking adults

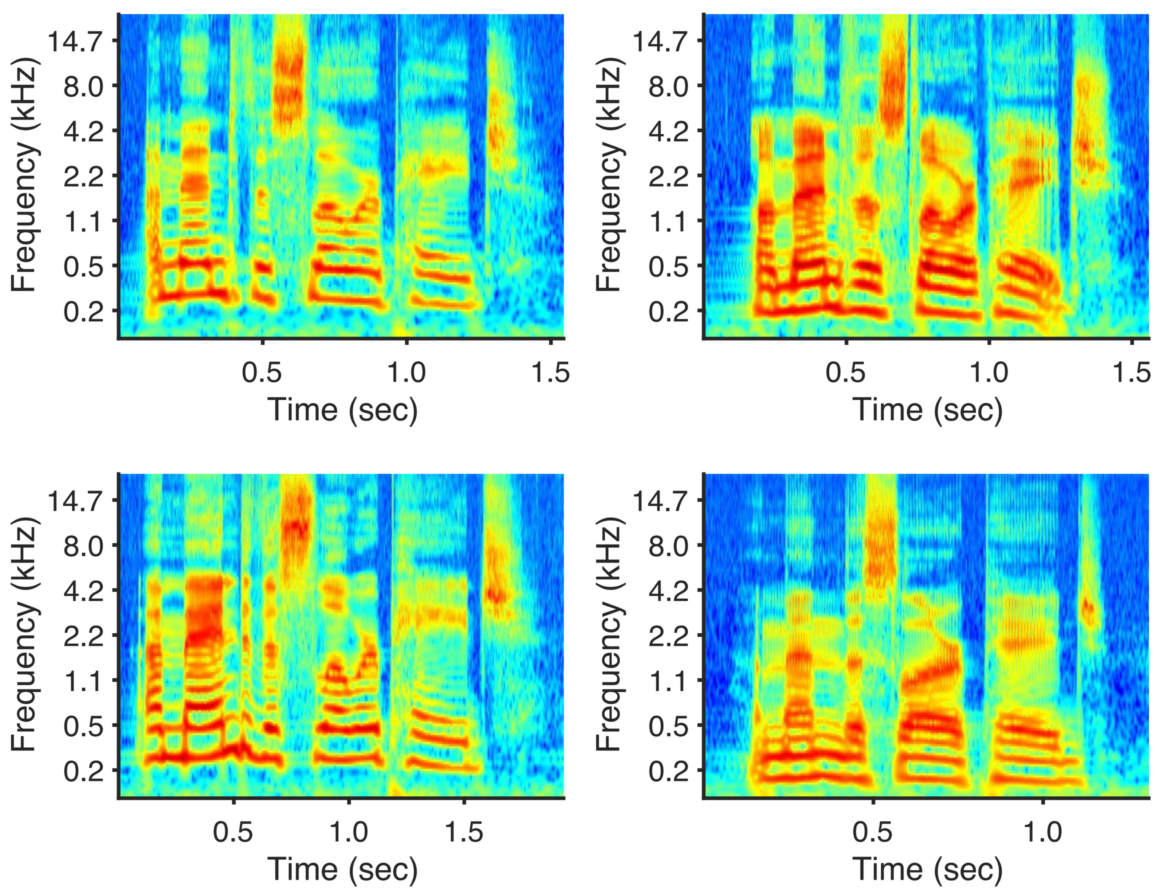

Figure 3: Cochleograms of four talkers uttering the phrase, "amend the slower page." (Left – female, right – male)

Method (continued)

Procedure:

- One-up, two-down, three alternative adaptive forced-choice task
- Stimuli presented in sound-treated booth over KRK Rokit 8G3 loud speaker, 70 dB SPL @ 1m
- Four tracks (one per talker) in separate runs
- Two conditions tested in separate blocks: full band *vs*. low-pass filtered @ 8kHz
- Run order within each block and block presentation both randomized
- **Reference** stimulus = 0° recording
- **Test** stimulus = Adaptively varied between 180° and 0° recordings based on listener performance (no feedback provided)
- Each run began with easily distinguishable test stimulus of 135° angle
- Step size changed from 45° to 15° after first two reversals; last six reversals averaged to obtain just noticeable difference (JND)

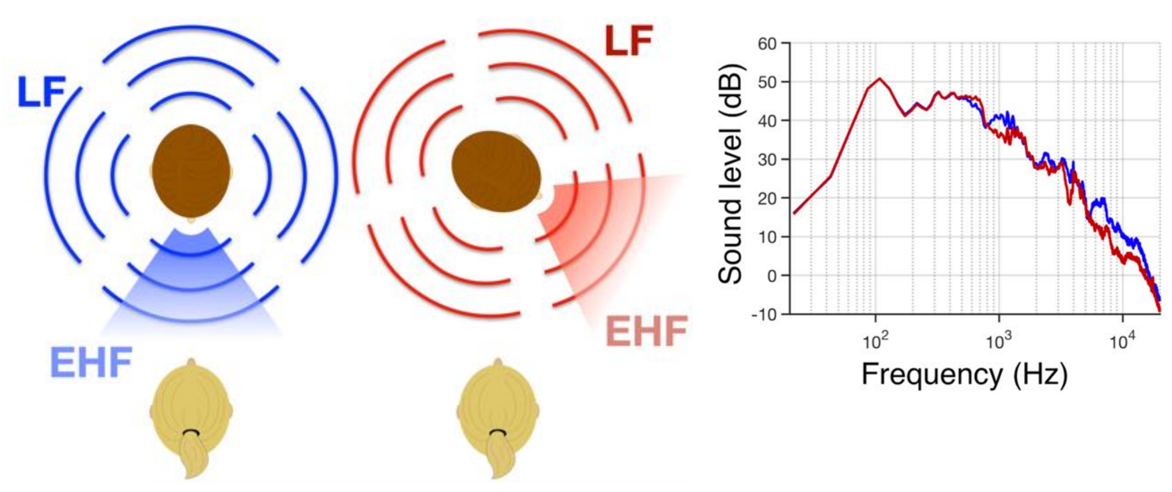

Figure 4: Speech spectra at the ear of the listener for talker head orientations of **0°** (**blue**) and 60° (red). Energy losses of 5-7 dB are observed at EHFs but not low frequencies (LF).

Results

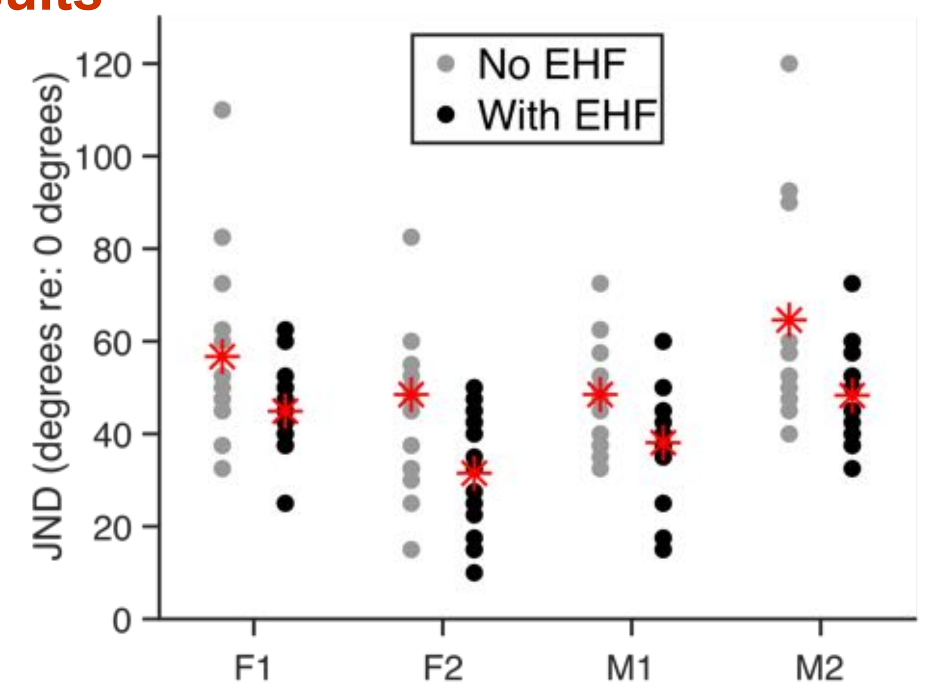

Figure 5: JND in talker head orientation (relative to 0°) for speech uttered by two female and two male talkers. Asterisks indicate mean JND.

Results (continued)

- Mean JND with EHF = 41°
- Mean JND without EHF = 55°
- Main effect of EHF condition ($p = 0.003$)
- Main effect of talker, with better performance for F2 and M1 ($p < 0.001$)
- No effect of talker gender ($p = 0.7$)

Conclusions

Listeners are sensitive to changes of approximately 40° in talker orientation using only auditory cues

This is less sensitive than what humans have displayed with only visual cues

These findings indicate that auditory cues are available for head orientation discrimination, which may be of greater utility when visual cues are unavailable

Access to EHFs improves auditory discrimination of talker head orientation

EHFs potentially serve as salient cue for determining "Are you looking at me?" (*i.e.*, is this vocalization directed at me?)

References

Dunn HK and Farnsworth DW (1939). Exploration of pressure field around the human head during speech. J. Acoust. Soc. Am. 10, 184–199.

Monson BB, Hunter EJ, and Story BH (2012) Horizontal directivity of low- and high-frequency energy in speech and singing. J. Acoust. Soc. Am., 132(1), 433-441.

Neuhoff JG (2003) Twist and Shout: Audible Facing Angles and Dynamic Rotation. Ecological Psychology, 15:4, 335-351.

Wilson HR, Wilkinson F, Lin LM and Castillo M (2000) Perception of head orientation. Vision Res. Vision Res. 40, 459-472.

Monson BB, Rock J, Schulz A, Hoffman E, and Buss E (submitted) Ecological cocktail party listening reveals the utility of extended high-frequency hearing.

ILLINOIS