DOI: 10.61782/fa.2025.0887

THE ROLE OF SPEECH DIRECTIVITY IN SPEECH PERCEPTION

Brian B. Monson[1*]
[1] Department of Speech and Hearing Science, University of Illinois Urbana-Champaign, USA
[2] Department of Biomedical and Translational Sciences, Carle Illinois College of Medicine, USA

ABSTRACT

Human speech directivity plays a role in speech perception, including for speech recognition in complex acoustic environments. When multiple talkers are in a room, the target talker's speech will be masked by both direct and reflected sound from background talkers. Because of speech directivity, the interfering signal at the ear of the listener will be dependent on background talkers' head orientations and locations relative to the listener and to reflective surfaces, as well as absorption characteristics of the reflective surfaces. We have examined horizontal speech directivity using anechoic, multi-channel, high-fidelity recordings of male and female talkers. Although speech generally becomes more directional as frequency increases, the relationship between directionality and frequency is nonmonotonic. There is some evidence for a sex effect, but at a limited set of frequencies. These factors may have implications for speech perception.

Keywords: *speech directivity, speech acoustics, speech perception*

1. INTRODUCTION

Speech directivity patterns convey acoustic cues to determine a talker's head orientation. To a first approximation, rotating the talker's head acts as a shallow low-pass filter, with the cutoff frequency decreasing as head angle increases (see **Figure 1**) [1-3]. Speech directivity patterns are also temporally dynamic, with different phonemes differing in their radiation patterns [3-4]. This is because speech directivity is affected by subtle changes in articulator placement or vocal tract geometry, size of the mouth opening, face and body geometry, and site of sound source generation within the vocal tract [5].

2. HEAD ORIENTATION DISCRIMINATION

Human listeners' ability to utilize acoustic directivity cues for talker head orientation discrimination has been demonstrated previously using a live talker or speech presented over a rotating loudspeaker. Taking a different approach, we measured the minimum audible change in talker head orientation using speech signals recorded simultaneously from different microphone locations surrounding a talker [6-7]. A large effect of talker on head orientation discrimination was consistent with the variation in individual speech directivity patterns: stimuli from two talkers who exhibited more directional radiation patterns yielded better discrimination performance than stimuli from two less directional talkers [7]. Additionally, we found that highly directional extended high frequencies (EHFs; >8 kHz), typically thought to be inconsequential for speech perception, support a listener's ability to discriminate talker head orientation [6-7].

3. SPEECH-IN-SPEECH RECOGNITION

Whereas the traditional speech-in-speech experiment simulates a listening scenario where both the target talker and maskers are facing a listener (owing to speech materials being recorded with a microphone directly in front of each talker), it would be unusual to have multiple talkers all facing a listener and talking at the same time in a real-world "cocktail party." We and others have demonstrated talker head-orientation-related (THOR) benefits when masker talkers face away from a listener, relative to when masker talkers face the listener [8-9]. Furthermore, due to the directionality of EHFs in speech, a listener at the real-world

*Corresponding author: monson@illinois.edu

cocktail party receives EHF spectral energy primarily from a talker facing the listener, whereas masking energy from background talkers is lacking EHF energy (**Figure 1**). This directionality makes EHF energy a salient cue for detection and segregation of the target talker from background talkers, in addition to increasing accessibility of phonetic information at EHFs. Similar to how visual cues of a social partner's head orientation and gaze might direct attention to that partner or other objects of interest, highly directional EHFs could herald the potential importance of speech signal, thereby drawing the listener's attention to that signal (i.e., high-amplitude EHF energy will only be received from a talker that is directly facing a listener [see **Figure 1**], which likely indicates that the listener is the intended recipient of this utterance). We have demonstrated the consequences of EHF directionality, showing that low-pass filtering speech at 8 kHz reduced speech-in-speech recognition when the target talker was facing the listener, but co-located maskers faced away from the listener [6, 10-12].

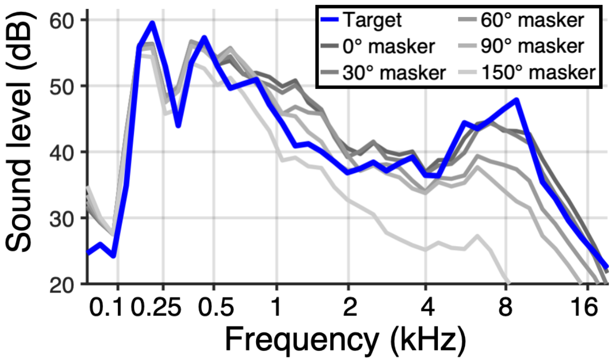

Figure 1. Long-term average speech spectrum for a target talker facing 0º (blue) and a two-talker masking facing different angles (grayscale).

4. SPEECH DIRECTIVITY

Given the role of speech directivity in speech perception, we undertook to analyze a publicly available multi-directional anechoic speech corpus of 15 female and 15 male talkers [13] to examine factors that might influence speech directivity. Utilizing methods we used previously [14], we calculated a frequency-dependent speech directivity index [15], and replicated our previous finding of the nonmonotonicity of the directivity index function [14]. We also found that talker sex had a small effect on the directivity index at a limited set of frequencies: male talkers exhibited slightly greater directionality toward 0º at approximately 1 kHz and 6-7 kHz.

5. CONCLUSIONS

Speech directivity patterns provide THOR cues for head orientation discrimination and speech recognition. In future work, we intend to examine the perceptual relevance of THOR cues for speech recognition in real-world, reverberant, complex listening environments.

6. ACKNOWLEDGMENTS

This work was supported by NIH Grant DC019745.

7. REFERENCES

[1] H. K. Dunn and D. W. Farnsworth, "Exploration of pressure field around the human head during speech," *The Journal of the Acoustical Society of America*, vol. 10, pp. 184–199, 1939.

[2] W. T. Chu and A. C. C. Warnock, "Detailed directivity of sound fields around human talkers," Tech. Report, *Institute for Research in Construction* (National Research Council of Canada, Ottawa, ON, Canada), pp. 1–47, 2002.

[3] B. B. Monson, E. J. Hunter, and B. H. Story, "Horizontal directivity of low- and high-frequency energy in speech and singing," *The Journal of the Acoustical Society of America*, vol. 132, no. 1, pp. 433–441, 2012.

[4] P. Kocon and B. B. Monson, "Horizontal directivity patterns differ between vowels extracted from running speech," *The Journal of the Acoustical Society of America*, vol. 144, no. 1, pp. EL7–EL12, 2018.

[5] R. Blandin, A. Van Hirtum, X. Pelorson, and R. Laboissière, "Influence of Higher Order Acoustical Propagation Modes on Variable Section Waveguide Directivity: Application to Vowel [α]," *Acta Acustica United With Acustica*, vol. 102, no. 5, pp. 918–929, 2016.

[6] B. B. Monson, J. Rock, A. Schulz, E. Hoffman, and E. Buss, "Ecological cocktail party listening reveals the utility of extended high-frequency hearing," *Hearing Research*, vol. 381, no. 107773, 2019.

[7] B. Moriarty B. B. Monson, E. J. Hunter, and B. H. Story, "Horizontal directivity of low- and high-

frequency energy in speech and singing," *The Journal of the Acoustical Society of America*, vol. 132, no. 1, pp. 433–441, 2012

[8] O. Strelcyk, S. Pentony, S. Kalluri, and B. Edwards, "Effects of interferer facing orientation on speech perception by normal-hearing and hearing-impaired listeners," *The Journal of the Acoustical Society of America*, vol. 135, no. 3, pp. 1419–1432, 2014.

[9] M. D. Braza, N. E. Corbin, E. Buss, and B. B. Monson, "Effect of Masker Head Orientation, Listener Age, and Extended High-Frequency Sensitivity on Speech Recognition in Spatially Separated Speech," *Ear & Hearing*, vol. 43, no. 1, pp. 90–100, 2022.

[10] A. Trine and B. B. Monson, "Extended high frequencies provide both spectral and temporal information to improve speech-in-speech recognition" *Trends in Hearing*, vol. 24, no. 2331216520980299, 2020.

[11] M. Flaherty, K. Libert, and B. B. Monson, "Extended high-frequency hearing and head orientation cues benefit children during speech-in-speech recognition," *Hearing Research*, vol. 406, no. 108230, 2021.

[12] R. M. Ananthanarayana, E. Buss, and B. B. Monson, "Band importance for speech-in-speech recognition in the presence of extended high-frequency cues," *The Journal of the Acoustical Society of America*, vol. 156, no. 2, pp. 1202-1213, 2024.

[13] M. K. Miller, V. Delaram, A. Trine, R. M. Ananthanarayana, E. Buss, B. B. Monson, and G. C. Stecker, "An anechoic, high-fidelity, multidirectional speech corpus," *Journal of Speech, Language, and Hearing Research*, vol. 68, no. 1, pp. 411-418, 2025.

[14] B. B. Monson and R. M. Ananthanarayana, "Individual variability and similarity in horizontal directivity patterns for speech," *Proceedings of Forum Acusticum*, September 11-15, Torino, Italy, pp. 4743-4746, 2023.

[15] J. G. Tylka and E. Y. Choueiri, "On the calculation of full and partial directivity indices," *3D Audio and Applied Acoustics Laboratory Technical Report* (Princeton University Press, Princeton, NJ), 2014.