Individual variability and similarity in horizontal directivity patterns for speech

Brian B. Monson[1,2,3*] **Rohit M. Ananthanarayana**[1]
[1] Department of Speech and Hearing Science, College of Applied Health Sciences
[2] Department of Biomedical and Translational Sciences, Carle Illinois College of Medicine
[3] Neuroscience Program
University of Illinois Urbana-Champaign, USA

ABSTRACT

Speech directivity patterns are highly complex and temporally dynamic during speech production, with notable variability across individuals. In this paper, we report some of the complexities observed in anechoic horizontal speech directivity data obtained from standing talkers, including variability and consistencies observed in speech directivity patterns across talkers. Although speech generally becomes more directional towards 0º azimuth as frequency increases, it is clear that the relationship between frequency and directionality is nonmonotonic. Our data indicate this relationship shows some consistency across talkers. For example, peak directionality towards 0º tends to occur around 7-9 kHz, with a decrease in directionality at higher frequencies. On the other hand, due to acoustic lobing, minimum directionality towards 0º and corresponding maximum directionality towards 90º tends to occur around 700-800 Hz. Whether these particular phenomena have perceptual consequences for listeners in complex auditory scenes is an intriguing but unresolved question. However, previously published perceptual data do indicate talker effects on listener detection of talker head orientation, suggesting individual variability in speech directivity patterns may be perceptually relevant.

Keywords: *speech directivity, speech acoustics, speech production, directivity index*

**Corresponding author*: monson@illinois.edu

1. INTRODUCTION

Directivity patterns for speech radiation are frequency and angle dependent [1-3]. In general, speech energy at lower frequencies radiates more omnidirectionally around the head of a talker, with increasing directionality toward the front of the talker as frequency increases. These patterns can be affected by differences in face and body anatomy and geometry, articulator placement or vocal tract geometry, size of the mouth opening, and site of sound source generation within the vocal tract [4]. For this reason, we previously observed that different phonemes differ in their directivity patterns [3, 5]. These physical factors, which vary across talkers, can lead to substantial individual variability in directivity patterns. One question is whether such individual differences in directivity are perceptually relevant. For example, it has been demonstrated that differences in head orientation between a target talker and background talkers can lead to improvements in speech recognition due to the low-pass filtering effect of turning the head [6-8]. Furthermore, we have demonstrated talker effects on listeners' ability to detect changes in head orientation [7], raising the possibility that individual differences in speech directivity may be perceptually relevant. Here, we aimed to examine more closely individual speech directivity patterns to help identify features that may or may not be consistent across talkers.

2. METHODS

Detailed methods of the data collection were published previously [3] but are repeated briefly here. Fifteen native speakers of American English (8 female) were recorded uttering semantically unpredictable phrases in an anechoic chamber located at Brigham Young University, Utah, USA. All talkers were trained vocalists. Talkers stood with the mouth located at the center of a 60-cm radius semicircular horizontal arc of 13 Class 1 microphones with flat frequency response. The angular separation between microphones was 15º, spanning 0º (directly in front of the talker) to 180º (directly behind the talker) around the right side of the talker (see Figure 1, upper left panel). The microphone boom was adjusted to the height of the standing talker's mouth. Recordings were made using a 44.1-kHz sampling rate at 24 bits.

To quantify directivity, a frequency-dependent horizontal directivity index (DI) was calculated for each talker as

$$DI(\omega) = 10\log_{10}\frac{|H_x(\omega)|^2}{\frac{1}{N}\sum_{n=0}^{N-1}|H_n(\omega)|^2} \qquad (1)$$

where H is the sound pressure, x is the reference angle, n is angle, and N is the total number of angles [9]. This DI represents the ratio of acoustical energy radiated toward the reference angle x to the average energy radiated in all directions. Frequency-dependent analyses were conducted using a 2048-point FFT to calculate the long-term average speech spectrum using all phrases recorded for each talker.

3. RESULTS AND DISCUSSION

Figure 1 shows frequency-dependent horizontal radiation patterns for all talkers using octave-wide analysis bands. Radiation sound level, indicated by radius of the curve, is plotted relative to sound level at 0º. Raw data are plotted in 15º increments with no smoothing applied. Whereas radiation patterns for 500 Hz and 1 kHz are quite similar across talkers, notable talker variability is observed in radiation patterns for higher frequencies. At extended high frequencies (EHFs; >8 kHz), talker differences of >10 dB can be observed for some angles.

Figures 2 and 3 show DI calculations across frequency for each talker, relative to radiation toward 0º and 90º, respectively. Although talker differences are apparent, some similarities can be identified. For 0º directionality, a prominent notch is observed between 700-800 Hz for nearly every talker (Figure 2). Peak directionality toward 0º is likewise similar, occurring between 7-9 kHz for most talkers. Some talkers exhibit substantially reduced DIs at EHFs beyond this peak frequency range. The notch between 700-800 Hz in 0º directionality corresponds to a peak in 90º directionality around this same frequency (Figure 3).

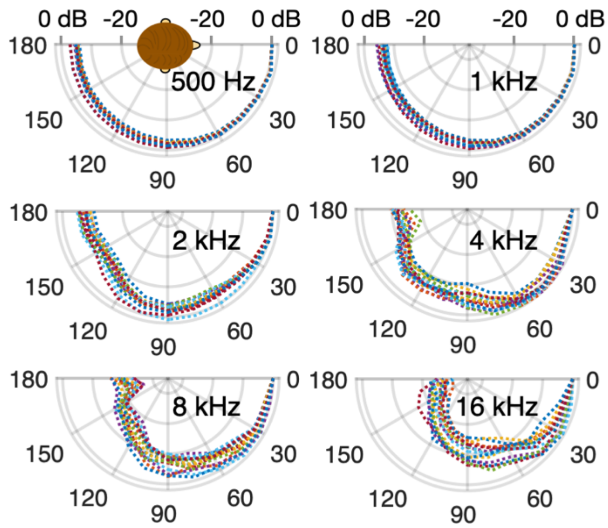

Figure 1. Horizontal radiation patterns for 15 talkers using octave-wide bands. Data are plotted relative to sound level at 0º. Concentric circles represent 10-dB steps.

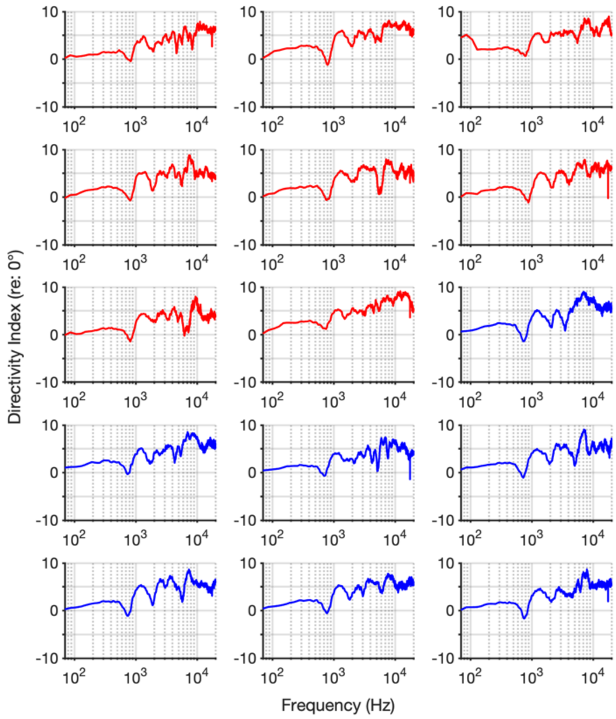

Figure 2. Directivity index calculations for directionality toward 0º. Red curves are female talkers.

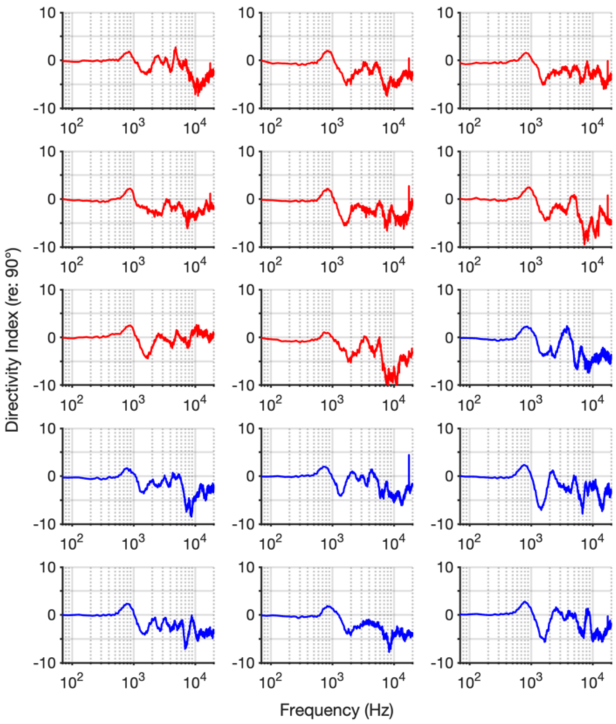

Figure 3. Directivity index calculations for directionality toward 90º. Red curves are female talkers.

To get some sense of whether the changes in DI might be perceptually relevant, Figures 4 and 5 show DIs calculated using cochlear filter bands (equivalent rectangular bandwidths; ERB) to ascertain what features may be present in a cochlear representation of directivity. The features identified earlier remain apparent.

The changes in the DI observed here are likely due to acoustic lobing that occurs for directional sound sources, in which the peak radiation changes direction as a function of frequency. Figure 6 is a directivity map showing this lobing pattern for one example female talker. Energy level is plotted in dB relative to the level at the angle of peak radiation. As may be deduced from the differences in the frequency loci of local peaks and notches in the DI curves shown in Figures 2-5, the lobing pattern differs markedly across talkers. It is important to note that, due to the microphone recording distance of 60 cm, it is possible that near-field effects may be influencing these patterns.

4. CONCLUSIONS

Speech directivity patterns exhibit marked variability across individual talkers. Although radiation of speech energy is broadly characterized as becoming more directional as frequency increases, there is a nonmonotonic relationship between directionality and frequency, and the pattern of this nonmonotonic relationship is talker dependent. However, some features of these patterns remain consistent across talkers, including reduced directionality toward 0º at ~700-800 Hz, and peak directionally toward 0º at ~7-9 kHz. Whether this variability and similarity lead to talker effects on the perception of directivity (e.g., talker head orientation discrimination) remains to be seen.

5. ACKNOWLEDGMENTS

We thank our participating talkers. This work was supported by NIH Grant R01 DC019745 (B. B. M.).

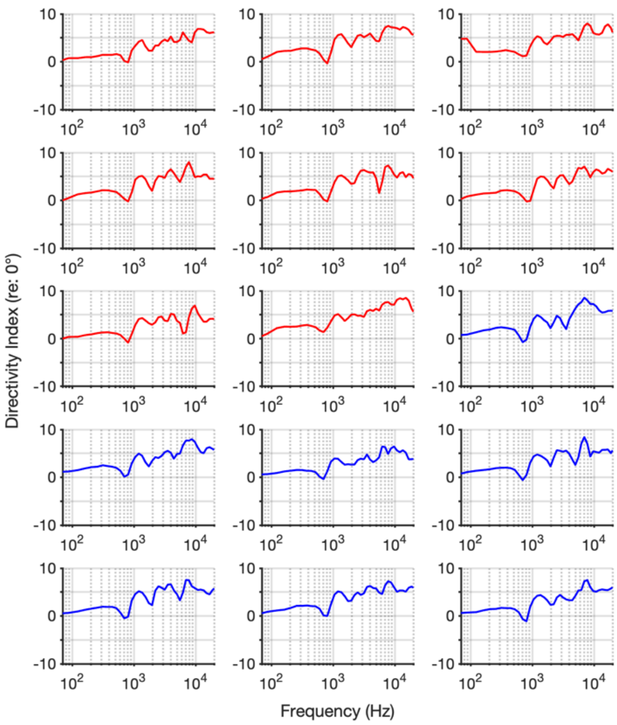

Figure 4. Directivity index calculations for directionality toward 0º using ERB bands. Red curves are female talkers.

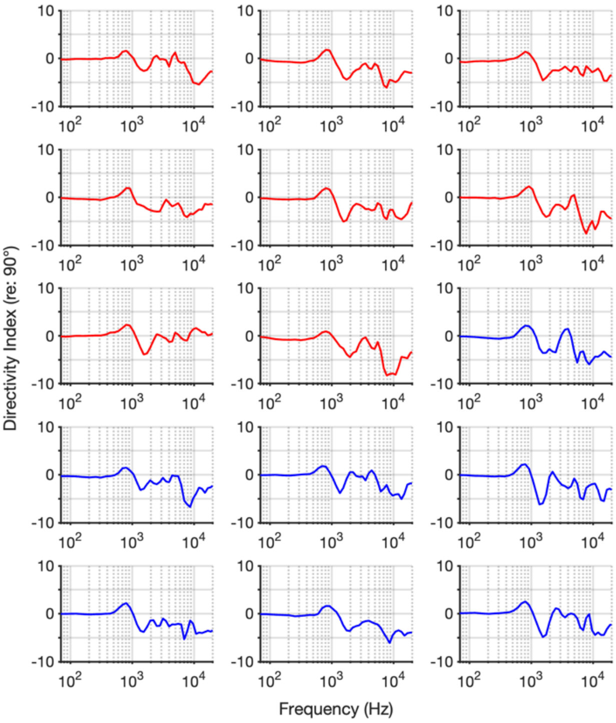

Figure 5. Directivity index calculations for directionality toward 90º using ERB bands. Red curves are female talkers.

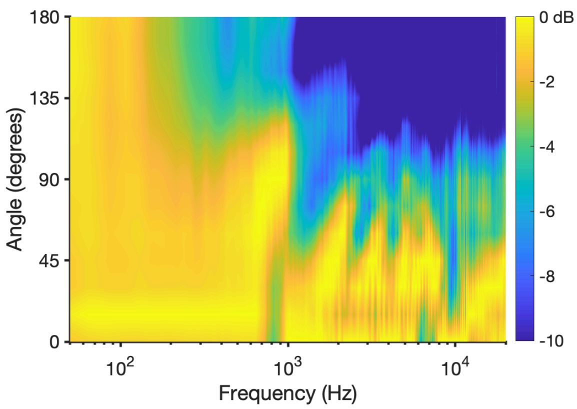

Figure 6. Directivity map of speech radiation for one example female talker, demonstrating acoustic lobing. The lobing results in minimal radiation at 0º and a peak at 90º at approximately 800 Hz, which was consistently observed across talkers.

6. 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] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.