Speech directivity patterns generated from a high-fidelity speech corpus

Allison Trine[1], Margaret K. Miller[2], Emily Buss[3], G. Christopher Stecker[2], Brian B. Monson[1]

1 Department of Speech and Hearing Science, College of Applied Health Sciences, University of Illinois Urbana-Champaign

2 Boys Town National Research Hospital, Omaha, Nebraska

3 Department of Otolaryngology/HNS, University of North Carolina at Chapel Hill

Introduction

Human talkers are directional sound sources – a phenomenon that has consequences for speech perception in multi-talker environments. Directivity patterns for speech showing frequency- and angle-dependent radiation reveal that speech generally becomes more directional toward the front of the talker as frequency increases. Differences in physical attributes can lead to individual variability in directivity patterns across talkers. Here, we examine individual variability in speech directivity using frequency-dependent directivity indices and directivity maps.

Directivity Index (DI) is defined as the ratio of energy radiated toward 0º vs. the average energy radiated in all directions.

Aims

- To examine individual variability in speech directivity

Methods:

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

Demographics:

- Thirty subjects (15 female)
- Ages 21.3-60.5 (m = 35.8) years
- Native speakers of American English

Speech materials:

- Four (sixteen-sentence) lists of the Bamford-Kowal-Bench (BKB) sentences
- Recorded in a 4x4x4 (interior) ETS A100 anechoic chamber
  - 1m free field spanning 100-20,000 Hz
  - Acoustical noise floor <15 dB SPL from 160-20,000 Hz

Methods (cont)

- 17 B&K type 4189 1/2" free-field condenser microphones (48 kHz, 24 bits) were spaced at 11.25-degree intervals at 1m spacing around the seated talker, at the height of the talker's mouth, spanning from 0 to 180-degrees azimuth.

Post-processing of speech materials:

- Best exemplar was selected from three repeats of each BKB sentence
- Sentences were high-pass filtered (to suppress ambient noise) in MATLAB using a 4th-order Butterworth filter with a 70-Hz cut-off frequency

Directivity Analysis:

- Directivity maps were calculated for each subject to demonstrate individual variability
- The long-term average speech spectrum was utilized to calculate directivity indices in 1-ERB (equivalent rectangular bandwidth) bands – for a total of 41 ERB bands (bands 2 through 42)
- A linear mixed-effects (LME) model was used to evaluate influences of gender and ERB band on directivity indices.

Results

Gender differences in directivity indices were evaluated using an LME model.

- There was no main effect of gender (p = 0.43).

	Estimate	SE	t-value	p-value
(Intercept)	0.94	0.272	3.44	0.0006
Gender	-0.31	0.385	-0.79	0.4335
Gender: ERB 15	1.43	0.528	2.71	0.0069
Gender: ERB 16	1.29	0.528	2.45	0.0144
Gender: ERB 31	1.52	0.528	2.89	0.0040
Gender: ERB 32	1.73	0.528	3.27	0.0011

Table 1 Significant interactions present in the LME. ERB bands 15 and 16 correspond to center frequencies of 924 and 1055 Hz. ERB bands of 31 and 32 correspond to center frequencies of 6237 and 6973 Hz.

Results

- There was a main effect of ERB band with higher-frequency bands tending to have higher (i.e., more directional) directivity indices, however there was a nonmonotonic relationship between average directivity indices and frequency (Fig 1).
- There was an interaction of gender and ERB band (Table 1).

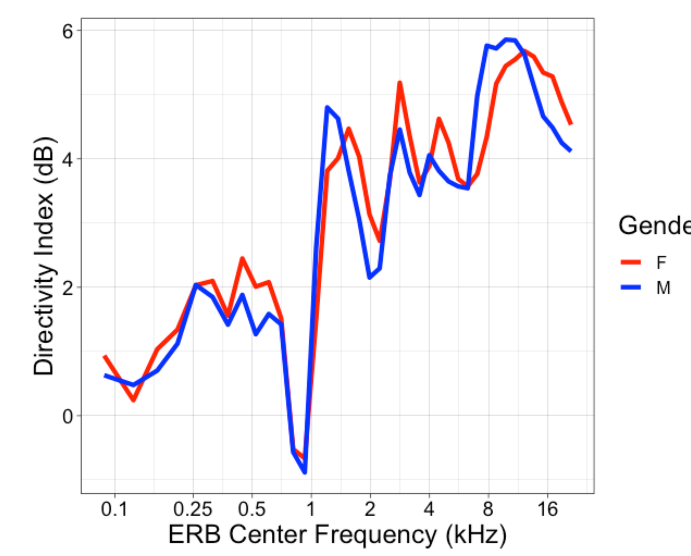


Fig 1 Average directivity index (dB) plotted against ERB center frequency (kHz) for males and females. Nonmonotonicity was present for both genders.

Differences in directivity patterns were present between subjects.

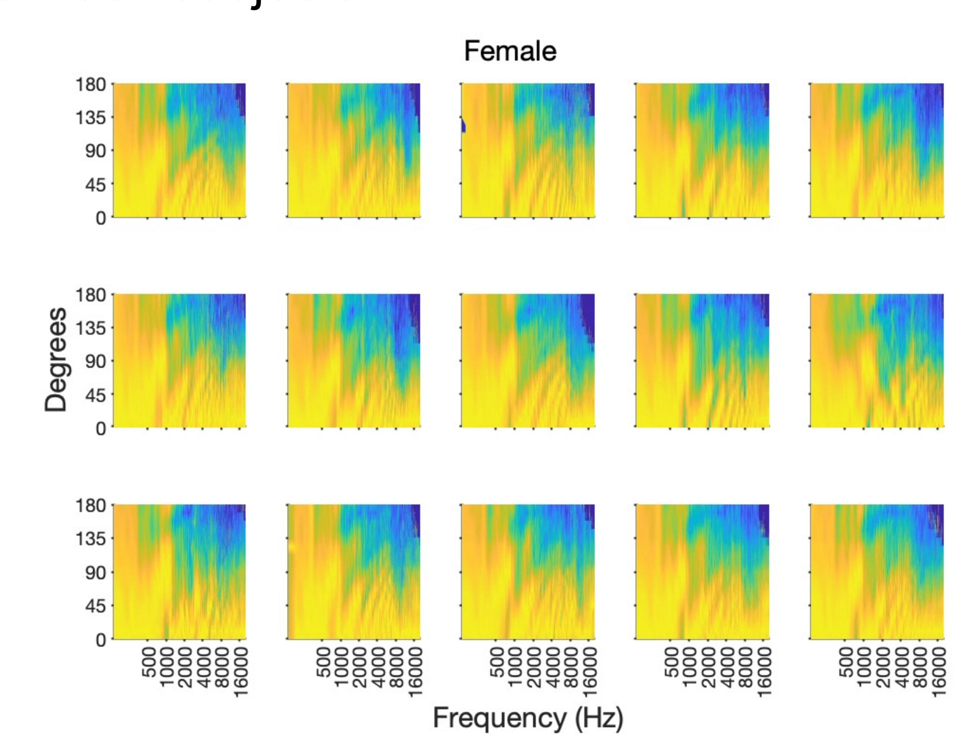
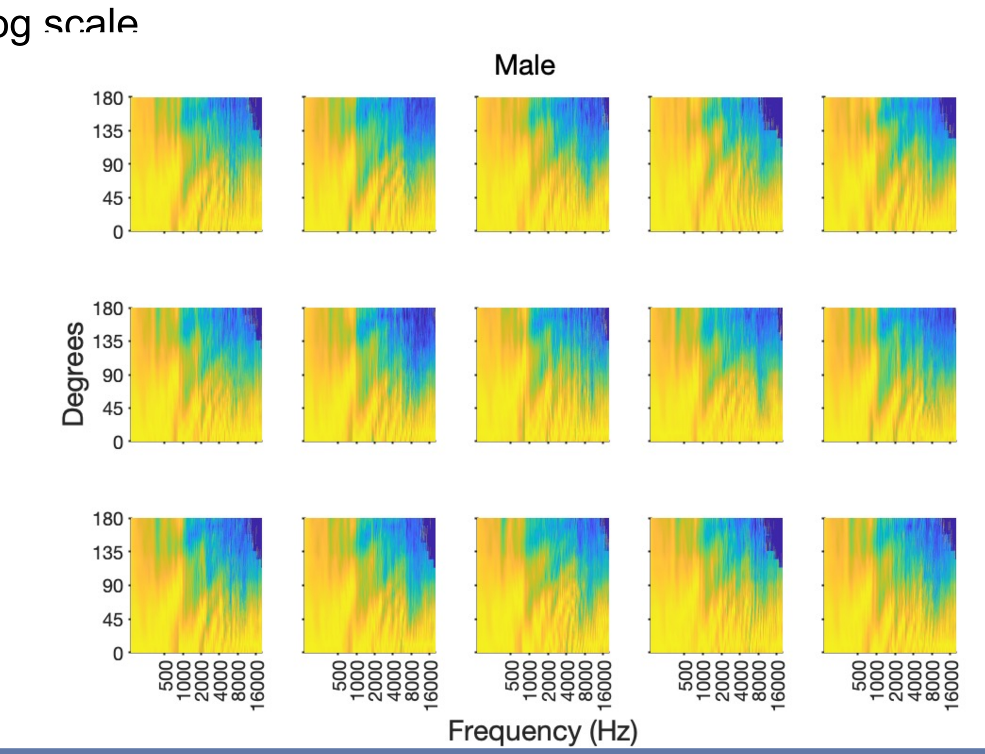


Fig. 2 & 3 Directivity maps demonstrate individual differences in speech radiation for ERB bands 2-42. Female speech (above) and male speech (below) are plotted on a log scale



Results (continued)

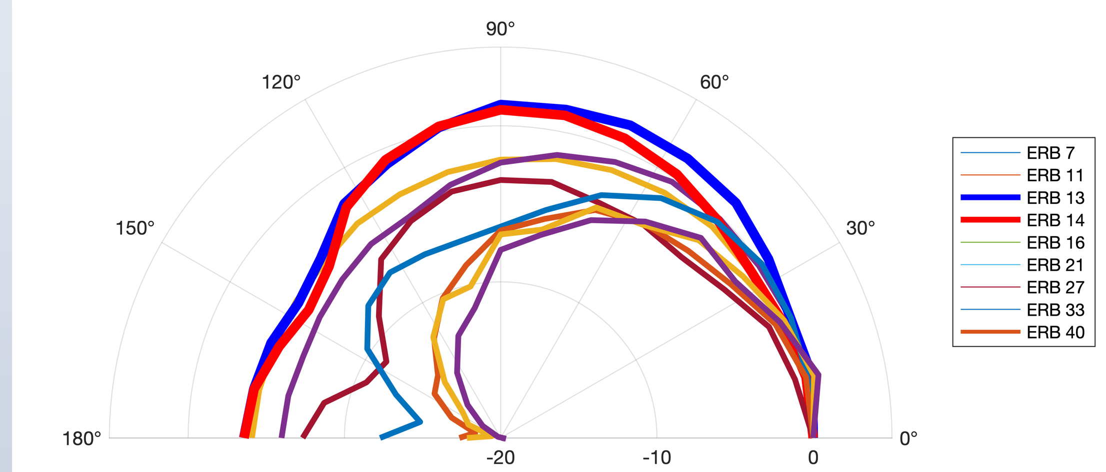


Fig. 4 Average octave band horizontal directivity patterns across all female talkers. Radius of curve indicates sound level relative to the level at 0-degrees. Bands 13 and 14 have highest directivity indices at 90-degrees.

Differences in directivity patterns were most prominent at ERBs 13 and 14 (corresponding to center frequencies of 700 and 806 Hz, respectively). Results were similar with both male and female speakers.

Conclusions

There was an interaction of gender and ERB band on directivity index.

Differences in directivity patterns at the ERBs of 13 and 14 were most prominent.

Male speakers may be more affected by noise floor in the higher frequencies.

References

Kocon, P., & Monson, B. B. (2018). Horizontal directivity patterns differ between vowels extracted from running speech. *The Journal of the Acoustical Society of America*, *144*(1), EL7–EL12. https://doi.org/10.1121/1.5044508

Monson, B. B., Hunter, E. J., & Story, B. H. (2012). Horizontal directivity of low- and high-frequency energy in speech and singing. *The Journal of the Acoustical Society of America*, *132*(1), 433–441. https://doi.org/10.1121/1.4725963

Acknowledgements

This study was supported by NIH grant number R01-DC019745.

ILLINOIS