

Horizontal directivity of low- and high-frequency energy in speech and singing

Brian B. Monson^{a)} and Eric J. Hunter

National Center for Voice and Speech, University of Utah, 136 S. Main Street, Suite 320, Salt Lake City, Utah 84101

Brad H. Story

Department of Speech, Language, and Hearing Sciences, University of Arizona, P.O. Box 210071, Tucson, Arizona 85721

(Received 20 December 2011; revised 1 May 2012; accepted 1 May 2012)

Speech and singing directivity in the horizontal plane was examined using simultaneous multi-channel full-bandwidth recordings to investigate directivity of high-frequency energy, in particular. This method allowed not only for accurate analysis of running speech using the long-term average spectrum, but also for examination of directivity of separate transient phonemes. Several vocal production factors that could affect directivity were examined. Directivity differences were not found between modes of production (speech vs singing) and only slight differences were found between genders and production levels (soft vs normal vs loud), more pronounced in the higher frequencies. Large directivity differences were found between specific voiceless fricatives, with /s,ʃ/ more directional than /f,θ/ in the 4, 8, 16 kHz octave bands.

© 2012 Acoustical Society of America. [http://dx.doi.org/10.1121/1.4725963]

PACS number(s): 43.70.Bk, 43.75.Rs, 43.20.Rz, 43.71.An [JW]

Pages: 433–441

I. INTRODUCTION

The directivity patterns of human speech and voice have been measured for use in a variety of applications. In addition to experimental verification of acoustical modeling techniques (Flanagan, 1960; Moreno and Pfitzschner, 1978), directivity measures are necessary to verify the accuracy of artificial human head and speech simulators (Flanagan, 1960; Chu and Warnock, 2002; Halkosaari *et al.*, 2005; Jers, 2007), which are in turn used experimentally for applications such as calculation of the speech transmission index in architectural acoustical design (McKendree, 1986; Bozzoli *et al.*, 2005). For this purpose, directivity has been measured specifically to aid in open office layout design (McKendree, 1986; Chu and Warnock, 2002) and small enclosure (vehicle) design (Bozzoli *et al.*, 2005). Others have measured directivity for singing in an attempt to enhance singing performance conditions for both soloists (Marshall and Meyer, 1985; Cabrera *et al.*, 2011) and group/choir singers (Marshall and Meyer 1985; Jers, 2007). Additional potential uses of speech/singing directivity patterns include optimization of microphone placement for recording (Dunn and Farnsworth, 1939; Cabrera *et al.*, 2011) and telecommunication purposes (Halkosaari *et al.*, 2005), and for acoustical modeling of human talkers in virtual environments (e.g., video games).

Like other areas of speech research, speech directivity data have largely been reported for frequencies only up to the 4 kHz octave band. Recent findings give evidence that high-frequency energy [(HFE) defined here as the energy in the 8 and 16 kHz octave bands] is of perceptual significance

(Monson *et al.*, 2011) and affects percepts of intelligibility (Lippmann, 1996; Stelmachowicz *et al.*, 2001; Pittman, 2008; Fullgrabe *et al.*, 2010; Moore *et al.*, 2010; Monson, 2011). Directivity at high frequencies in speech is therefore of potential interest for the applications listed earlier. More specifically, here it was desired to know which recording angles gave reasonably accurate representation of HFE in the signal so that high-fidelity speech and singing recordings could be made for use in perceptual experiments. Naturally, it is assumed that HFE is highly directional as increased directionality is generally seen with increased frequency, but this has not yet been shown for speech, and the extent of speech/voice HFE directionality is unknown. Furthermore, possible differences in HFE directivity caused by differing types of vocal productions have not yet been examined.

Perhaps the greatest difficulty met in trying to obtain accurate directivity measurements for humans is the need for simultaneous data collection from multiple locations around a talker. Of course, the ideal apparatus for obtaining global directivity patterns would consist of a full sphere of microphones surrounding a human subject with high spatial resolution in an anechoic environment. Assuming left–right symmetry for human talkers (as is typical in directivity studies) decreases this to a hemisphere of microphones, but such an arrangement still seems impractical. Methods used in directivity data collection must combat this problem. One of the first published studies collecting directivity data for human speech (Dunn and Farnsworth, 1939) used only two microphones—a reference microphone and a single “exploring” microphone moved to each location around a seated talker—to record data in both the vertical and horizontal (azimuthal) planes. This necessitated repetition of the recorded material by the talker for every recording location, and the authors noted the extensive time taken to collect data from their single male subject. While the authors did

^{a)} Author to whom correspondence should be addressed. Present address: Duke-NUS Graduate Medical School, Singapore. Electronic mail: brian.monson@duke-nus.edu.sg

publish some HFE directivity data (up to 12 kHz), it is difficult to know the accuracy of this data due to the equipment available at the time, the fact that it was collected from a single seated subject, the poor spatial resolution used (45°), and the multiple repetitions necessary for their data collection method.

Nowadays the more common and efficient approach has been to use an arc of microphones placed in the vertical plane to simultaneously record data (in the vertical plane) and to rotate either the arc or the subject to collect data in the horizontal plane (see, e.g., Chu and Warnock, 2002; Bozzoli et al., 2005; Jers, 2007). Again, the limitation of this method is that the subject must repeat the recorded material for each new recording position, introducing error in the horizontal plane data due to variations made by the subject during repetition (e.g., in mouth shape, posture, loudness, etc.) and over time. Of studies utilizing this method, Chu and Warnock (2002) recently reported some HFE directivity (third-octave data up to the 8-kHz third-octave band) for human talkers ($n = 40$), though these subjects were sitting down.

An alternative to the rotation/repetition method is to collect simultaneous data in a single plane of interest. Simultaneous horizontal plane data was collected by McKendree (1986) and more recently by Cabrera et al. (2011) using microphones to create a full semicircle from directly in front of to directly behind the subject. While no data were collected in the vertical plane, the reported horizontal plane directivity patterns are likely more accurate than those in other studies as no repetitions were necessary. These two studies were not without other limitations, however. Specifically, while McKendree (1986) recorded a large number of subjects ($n = 56$, 17 male), only seven microphones were used, resulting in a high spatial resolution of 30° , and the subjects were again seated during the data collection process. (That subjects were seated in these studies is not a limitation *per se*—and in fact may be desirable depending on the application of interest—but this arrangement will likely introduce reflections from the legs and chair that could affect directivity patterns, particularly at high frequencies.) Cabrera et al. (2011), on the other hand, recorded only eight subjects (two male), and only singing data were collected (not speech) as the authors were interested in directivity of singers in performance halls.

Authors from these previous studies have made several claims and conjectures regarding speech and voice directivity. The following four specific claims are examined here: (1) McKendree (1986) reported some large gender differences in directivity for certain octave bands, although Chu and Warnock (2002) found no such differences. (2) Chu and Warnock (2002) suggested the speech production loudness level (low vs normal vs high) would affect the directivity pattern, particularly for low-level speech, but used only one male subject to show this. (3) Marshall and Meyer (1985) showed directivity differences for individual phonemes (vowels), suggesting that humans can volitionally manipulate their directivity patterns. (4) Cabrera et al. (2011) predicted that singing would be more directional than speech, especially at high frequencies, on the assumption that a larger aperture (mouth) size is used for singing. Although these claims are interesting and potentially useful for applications utilizing speech directivity data, their valid-

ity (and pertinence to the higher frequencies) is questioned here due to the limitations of the previous studies.

In this study we attempted to provide accurate horizontal plane directivity patterns by collecting simultaneous full audio bandwidth data in the horizontal plane from standing talkers and singers. At the same time, the accuracy of previously reported directivity data was examined. Specifically, the effects on directivity caused by gender, production level (soft vs normal vs loud), production mode (speech vs singing), and different phonemes were studied. Voiceless fricatives were selected for analysis in this study because they are characterized as having high levels of HFE (Jongman et al., 2000; Maniwa et al., 2009). The main drawback of the method used here was that full spherical directivity patterns could not be obtained because data were recorded only in a single plane. Thus, no commentary on directivity patterns in the vertical plane can or will be made.

II. METHOD

A. Recordings

For the description that follows, the reader is referred to Fig. 1, which depicts the recording apparatus. A semicircular microphone boom 2.35 m in diameter was crafted using a 3/4 in. metal pipe. Thirteen 1/2 inch Type 1 (Larson Davis 2551, Provo, UT) precision microphones were used for the recordings. To minimize the effects of high-frequency scattering caused by reflection off of the microphone boom, the microphones were attached to the end of 0.57 m rods protruding perpendicularly from the microphone boom. Thirteen rods were attached at 15° increments from 0° to 180° , and were adjusted in length of protrusion inward from the semicircular boom such that the semicircle radius (after microphone placement) was 60 cm. The apparatus was then attached horizontally to two stands crafted from 1 in. threaded rods that were attached vertically to two stationary columns located just beneath the cable floor of the anechoic chamber. The height of the boom was adjustable. The microphones were regularly calibrated throughout the recording process (1 kHz, 114 dB).

Subjects were positioned facing directly on-axis from the microphone at 0° (channel 1) with the mouth measured to be at the center point of the 60 cm radius semicircle. The height of the microphone boom was set to the level of the

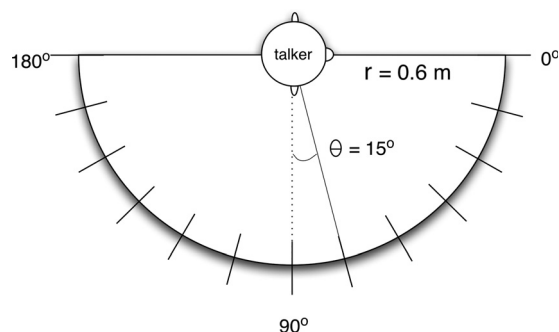


FIG. 1. Depiction of the semicircular recording apparatus and setup (top-down view with talker facing right). Microphone distance from the subject's mouth was 60 cm, with microphones placed at 15° increments from directly in front of the subject (0°) to directly behind the subject (180°).

mouth of the subject. No previous directivity study has physically constrained subject movement. The subjects here were likewise not constrained as this would have introduced difficulty in vocalizing naturally (particularly for singing), but were instructed not to move. Although this likely introduced some error, subjects were observed to maintain head location during recording. When mouth distance was measured between recordings, deviation was minimal (1–2 cm). The microphone signals were patched into a separate control room located adjacent to the chamber. Acoustical data from the microphones were collected using a National Instruments PXI 36-channel data acquisition system (Austin, TX), recording at 24 bits with a 44.1 kHz sampling rate. Data were recorded as binary files using a customized LABVIEW interface, and were then imported into MATLAB (The MathWorks, Inc., Natick, MA) for analysis and conversion to wave files.

Post-processing was necessary to correct for two of the microphones that exhibited frequency response characteristics deviating from the other channels. The response from channel 2 (15°) decreased at a higher slope above 2.5 kHz and continued to drop relative to the rest of the channels as frequency increases. A correction factor for channel 2 was designed to increase the level recorded in channel 2 by the difference between the ambient room noise measurements in channels 1 and 2, approximated to the nearest 0.5 dB, for frequencies above 2.5 kHz. The response from channel 12 (165°) had a higher noise floor that affected data collection above ~14 kHz, except in louder conditions where energy levels above 14 kHz were greater than the microphone noise floor. In these conditions (some loud speech, loud singing, and fricatives) the level at 165° was observed to fall between the level at 150° and 180°, as might be expected. Using this observation as a guide, any corrupted data above 14 kHz obtained from channel 12 were interpolated between levels obtained at channels 11 and 13. The chamber and microphone noise floors were sufficient to capture data (including HFE data) at all other microphones (see the Appendix for octave band measurements of the noise floor).

B. Subjects

Recordings were made from 15 singer subjects (8 female) who were native speakers of American English with no reported history of a speech or voice disorder. All subjects had at least 2 years of post-high school private singing voice training. Age ranged from 20 to 71 years (mean = 28.5).

C. Recording procedure

The subject recorded the first 20 six-syllable low-predictability phonetically representative phrases with alternating syllabic strength given in Appendix A of Spitzer *et al.* (2007). The phrases were spoken at three different levels: Normal, soft, and loud. The phrases were also sung at three different levels: Normal, pianissimo, and fortissimo. The subject sang in his/her preferred singing style. Singing styles included classical/opera ($n = 10$), classical/choral ($n = 2$), jazz/pop ($n = 2$), and musical theater ($n = 1$). The variety of styles used allowed for generalization of the results. (For greater detail, see Monson, 2011.)

D. Acoustical analysis

A long-term average spectrum (LTAS) of each recording from channel 1 for each subject in each condition was created using a 2048 point fast Fourier transform, resulting in a frame length of $2048/44.1 \text{ kHz} = 46.44 \text{ ms}$, with a Hamming window and 50% overlap. Frames containing only silence were excluded from analysis. To ensure precise directivity data, the identical time frames used in channel 1 were analyzed in creating a LTAS from each of remaining 12 channels (for each subject in each condition). A “mean” LTAS for each condition and channel was then calculated by normalizing the LTAS for each subject in a given condition and channel to an overall level of 0 dB. The linear versions of each subject’s normalized LTAS for that condition and that channel were averaged across subjects. The resulting LTAS was then normalized to the mean overall level calculated for that condition at that channel. Similarly, a mean LTAS was calculated for each gender in each condition at each channel. To examine effects of specific phonemes, the voiceless fricatives /s,ʃ,f,θ/ were extracted from the normal speech condition of the phrases recorded by each subject and a mean LTAS across subjects was calculated for each fricative and each channel. All levels were adjusted to the equivalent level at 1 m from the mouth using the inverse square law.

III. RESULTS AND DISCUSSION

Data are presented in two fashions: Directivity plots showing radial patterns in the horizontal axis (assuming left/right symmetry), and linear plots showing level change as a function of angle. Unless otherwise noted, all data in the radial directivity plots (including data from previous studies) were plotted as level relative to a channel 1 (which was set arbitrarily to a level of 30 dB) and were smoothed by interpolation by a factor of 3. All data in the linear plots were plotted as level relative to channel 1. Data are presented for overall level, and each octave band level from 125 Hz to 16 kHz. Where available, data from previous studies are included.

Figures 2–4 show the directivity patterns obtained for normal speech, plotted with data from comparable frequency bands reported in three previous studies. The results indicate normal speech is not as directional as the McKendree (1986) data (McK) for octave bands up to 2 kHz—a result also seen in the Chu and Warnock (2002) data (CW). In general, the patterns have a similar contour to the CW patterns, though the data here indicate that radiation above 1 kHz is more directional than they reported, particularly in the 4 and 8 kHz octave bands. This may be due to scattering of the high frequencies caused by the legs and chair of the seated subjects. The Dunn and Farnsworth (1939) data (DF) fit remarkably well for many of the bands shown, despite the use of only a single seated subject and a single microphone. In fact, the only band where the data appear to deviate significantly from the results here was in the 8 kHz octave. HFE in the 8 and 16 kHz octaves is clearly very directional toward the front of the talker. The overall and octave levels at each angle for normal speech (and all other conditions presented) are included in the Appendix. HFE levels are down by ~3 dB at 45°.

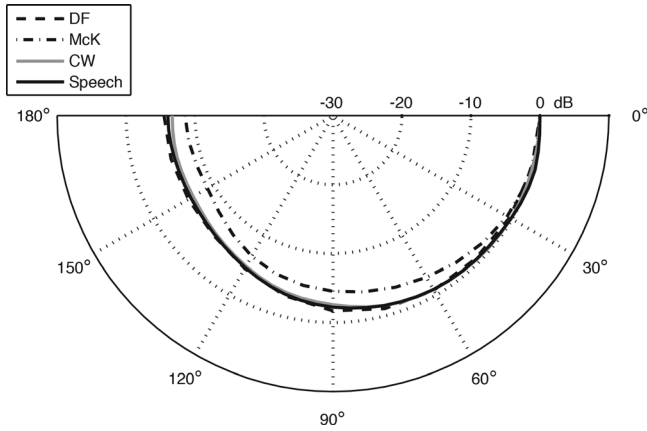


FIG. 2. Overall level directivity of normal speech comparing data from Dunn and Farnsworth (DF), McKendree (McK), Chu and Warnock (CW), and this study (Speech). The talker is facing 0° (to the right).

A. Effects of gender

Figures 5 and 6 show the directivity patterns obtained for normal speech, separated by gender. The McK results are included in Fig. 6 for comparison. Very few appreciable differences in directivity were seen between genders for octave bands below the 8 kHz octave, at which point male speech appears to become slightly more directional than female speech. While not shown here, this latter phenomenon was not observed in CW, but their data did confirm the absence of gender differences in the lower octaves. It is possible that the increase in directionality at higher frequencies for males is due to larger mouth sizes for men vs women, however, the gender differences in the 8 and 16 kHz octaves were all less than 3 dB. Mean directivity indices (DI)

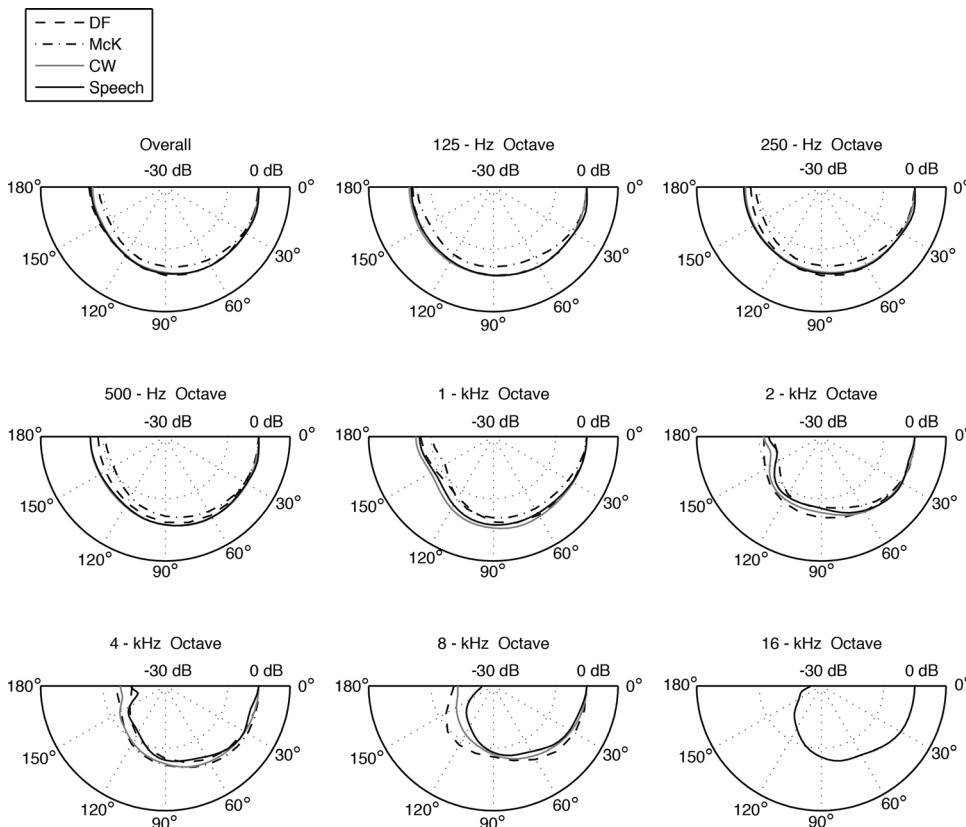


FIG. 3. Radial data showing overall and octave band directivity of normal speech comparing data (where available) from Dunn and Farnsworth (DF), McKendree (McK), Chu and Warnock (CW), and this study (Speech).

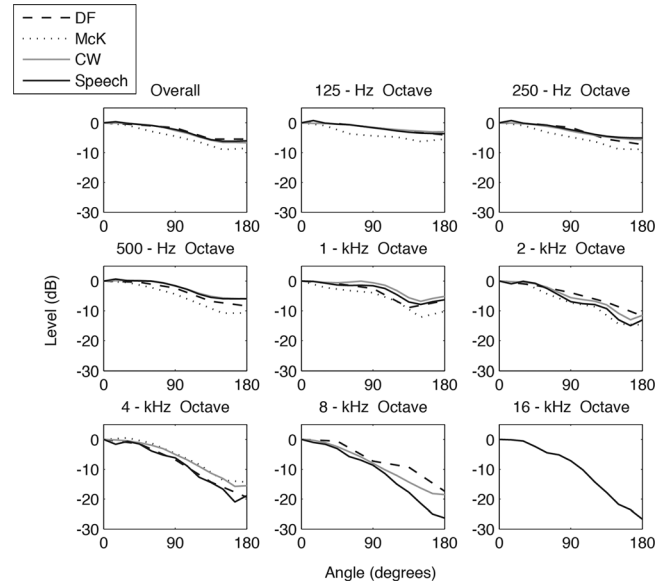


FIG. 4. Linear data showing overall and octave band directivity of normal speech comparing data (where available) from Dunn and Farnsworth (DF), McKendree (McK), Chu and Warnock (CW), and this study (Speech).

calculated for overall level for each gender showed no significant differences between males (1.9 dB) and females (2.3 dB) using an independent-samples *t*-test.

B. Effects of production level

Figures 7 and 8 show the normalized directivity patterns for soft, normal, and loud speech. In Fig. 8 an interesting (albeit very subtle) trend was seen starting at the 1 kHz

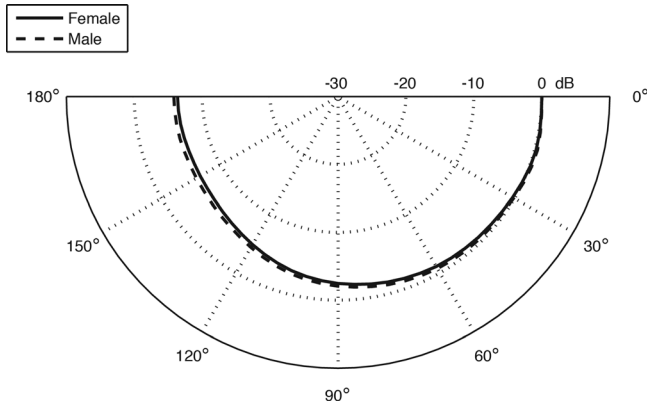


FIG. 5. Overall level directivity of normal speech comparing female and male speech.

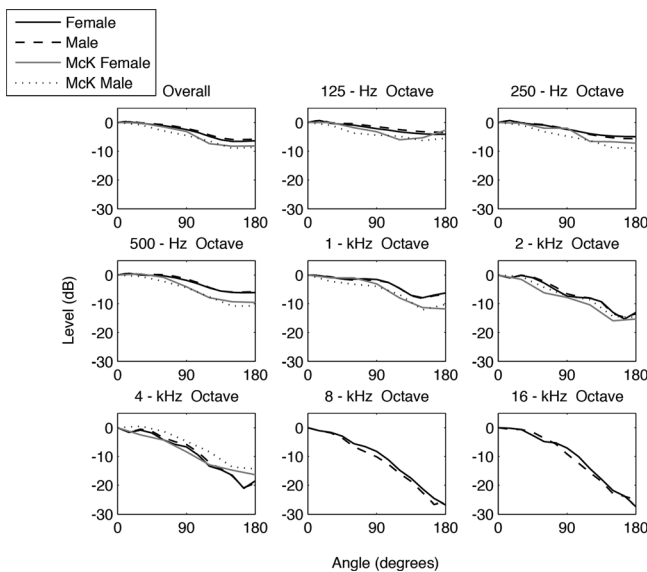


FIG. 6. Linear data showing overall and octave band directivity of female and male speech comparing data (where available) from McKendree (McK).

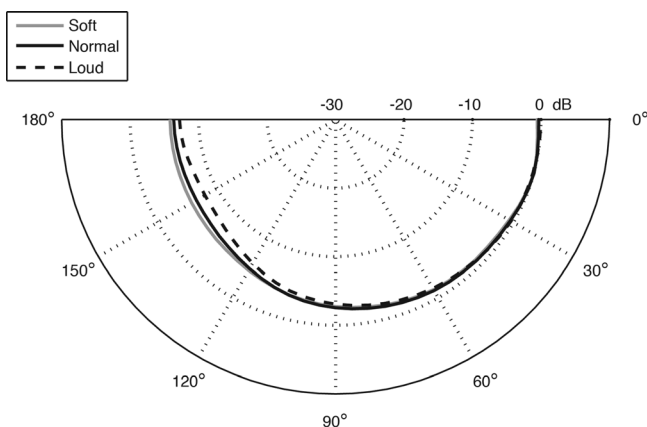


FIG. 7. Overall level directivity comparing soft, normal, and loud speech.

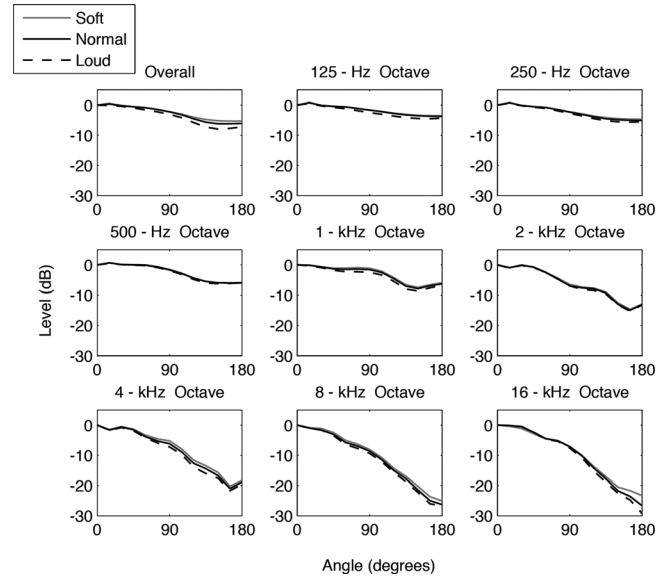


FIG. 8. Linear data comparing overall and octave band directivity of soft, normal, and loud speech.

octave and carrying through to higher bands. That is, as production level increased, speech appeared to become slightly more directional. Mean overall DIs calculated for soft, normal, and loud speech were 1.9, 2.1, and 2.7 dB, respectively, and the mean differences were found to be highly significant with a repeated-measures analysis of variance (ANOVA) ($F(2,28) = 23.629, p < 0.001$). Pairwise comparisons showed significant differences for each increase in production level. The effect was evident when examining overall directivity, but was most pronounced in the 16 kHz octave at 180°. Differences between loud and soft speech approached 3 dB at the higher angles for overall level, but increased above 3 dB at only 165° and 180° in the 16 kHz octave. The directivity differences reported for the single-subject CW data were not observed between levels here. In fact, the opposite effect was seen in CW. It is likely that their single-subject data are not generalizable.

C. Effects of production mode

Figures 9 and 10 show the directivity patterns obtained for both normal speech and normal singing. There was

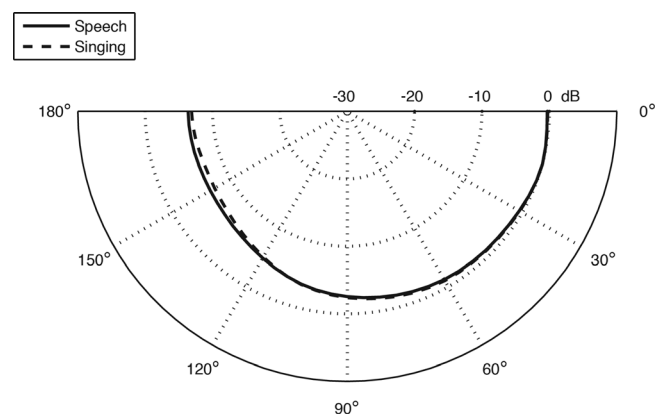


FIG. 9. Overall level directivity of normal speech and singing.

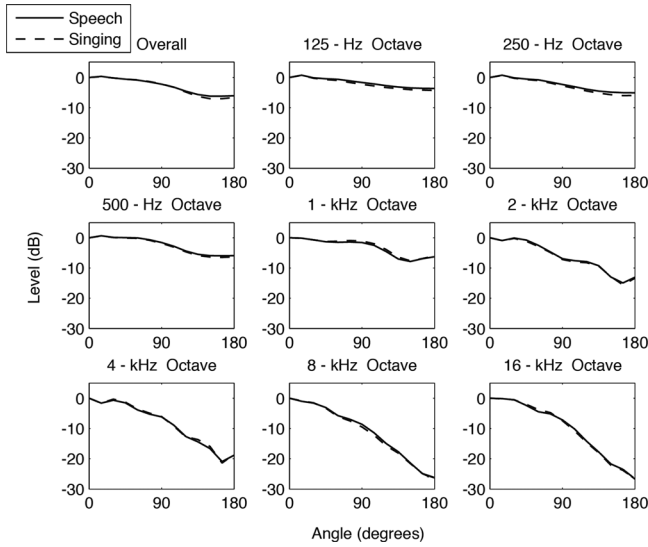


FIG. 10. Linear data comparing overall and octave band directivity of normal speech and singing.

virtually no difference in directivity between normal speech and singing as singing had the same mean overall DI as speech (2.1 dB). Any slight differences revealed no discernible consistencies across frequency bands; all differences were within 1 dB. The conjecture that singing would be more directional than speech does not appear to hold true, suggesting that the difference in the size of the mouth opening between singing and speech is not enough to affect directivity. However, mouth shape is likely dependent on the style of singing used, and the varied styles included in this analysis may have limited this effect.

D. Effects of phoneme

Figures 11 and 12 show the directivity patterns obtained for the four voiceless fricatives /s,ʃ,f,θ/. For higher frequencies starting at the 2 kHz octave band, fairly systematic and consistent differences were seen in directivity, with /s,ʃ/ being most directional and /θ/ being least directional. The largest differences were found between /θ/ and /s,ʃ/ in 8 kHz octave, reaching nearly 9 dB at angles beyond 90°.

Overall levels showed even larger differences in directivity patterns. This is likely attributable to the differing

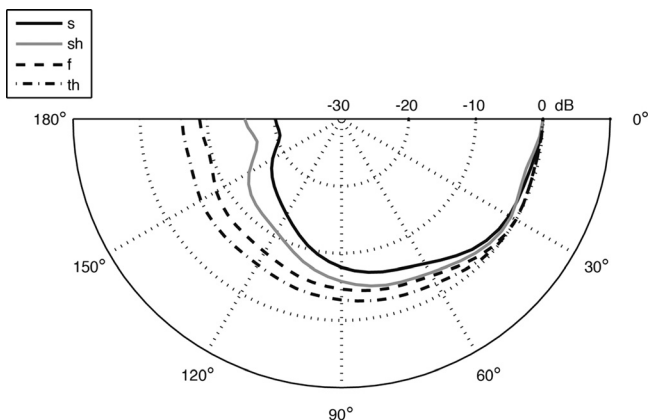


FIG. 11. Overall level directivity of the voiceless fricatives /s,ʃ,f,θ/.

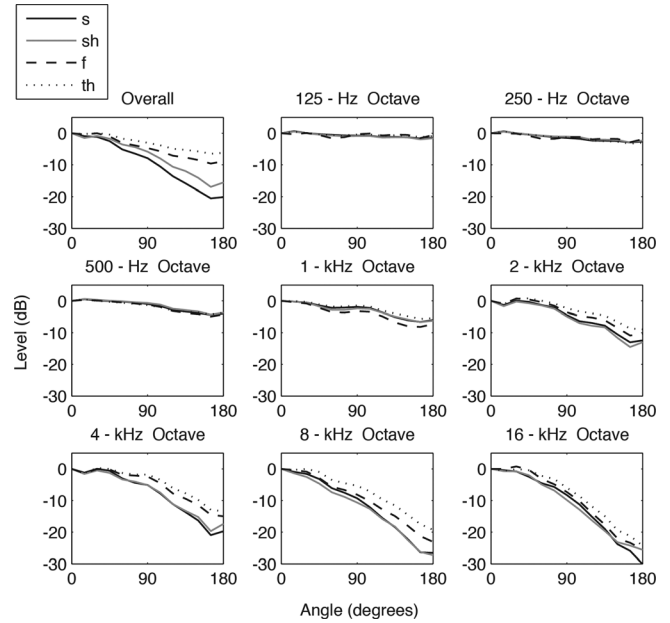


FIG. 12. Linear data comparing overall and octave band directivity of /s,ʃ,f,θ/.

frequency content of each fricative (Monson, 2011). That is, since the majority of energy in /s/ is in the 8 kHz octave, the overall level directivity should be influenced greatly by the directivity in that band. Likewise, since the majority of energy in /ʃ/ is in the 4 kHz octave, the overall directivity will be influenced by the directivity in that band. The fricatives /f/ and /θ/, on the other hand, are more broadband in nature, and will be more influenced by low-frequency directivity. Mean overall DIs calculated for /s,ʃ,f,θ/ were 5.3, 4.6, 3.2, and 2.5 dB, respectively, with the differences found to be highly significant using a repeated-measures ANOVA ($F(2,28) = 71.301, p < 0.001$). Pairwise comparisons showed significant differences between each fricative pair except /f/ and /θ/.

IV. CONCLUSIONS

Although previous studies have reported on speech/voice directivity data, limitations of the methods used in these studies have apparently resulted in some inaccurate data and conclusions. The results here reveal that appreciable directivity differences between genders were only evident in HFE in the 8 and 16 kHz octave bands, with male speech being slightly more directional than female speech. Increasing production level from soft to loud appears to increase the directionality of speech somewhat (again, more evident in the higher frequencies), perhaps due to changes in mouth shape that accompany an increase in vocal effort. However, differences between speech and singing were negligible, indicating that changes in mouth shape between normal speech and singing are not significant enough to affect directivity. These results suggest that, for the most part, modeling of human speech directivity does not need to account for differences caused by production mode, or by gender, unless the high frequencies are of particular concern. The differences caused by changes in production level might also be small enough to disregard overall, but should be included if HFE is of consequence.

On the other hand, specific mouth shapes and sound source generation mechanisms can greatly affect directivity, as seen by the large directivity differences in the voiceless fricative data shown here. These differences were likewise most pronounced in the higher frequency bands. This result reveals that certain phonemes are significantly more directional than others, showing the potential for humans to have some volitional control over their directivity patterns. Furthermore, this suggests that the phonetic content of the recording material used when obtaining directivity patterns could have an effect on the resulting patterns, which might account for some of the discrepancies found between directivity studies.

As HFE is highly directional, it is recommended to use a recording angle of 30° or less when making high-fidelity speech and singing recordings for experimental use. The results here indicate that recording at higher angles decreases HFE level by at least 3 dB—a level difference that is detectable by human listeners (Monson *et al.*, 2011). While not surprising, the directional nature of HFE is of scientific interest because, as studies on HFE continue to reveal its perceptual significance, HFE directivity may offer explanation for certain perceptual phenomena. For example, Best *et al.* (2005) have shown that successful vertical localization of a speech source (including front/back localization) is dependent upon HFE. As HFE is highly directional, a listener's ability to locate a talker would depend upon the orientation of the talker's head relative to the ear of the listener—an observation made by Lord Rayleigh in 1908 (Rayleigh, 1908).

The directionality of HFE suggests that HFE could become the optimal localization and selective attention cue in a listening situation consisting of a target talker directly facing the listener with multiple background talkers not facing a listener (the “cocktail party” problem). Furthermore, since the competing noise from background talkers in this situation would be largely in the low frequencies (due to their omni-directional behavior), and since HFE contains perceptual information useful for intelligibility, attending to HFE would likely be an effective strategy for successful communication in this situation. This notion is corroborated by recent results from Badri *et al.* (2011) who found that increased audiometric thresholds at nonstandard high frequencies (above 8 kHz) was a trait common to listeners who complained of and exhibited impaired speech-in-noise performance despite clinically normal standard audiometric thresholds (up to 8 kHz). To aid such individuals or other hearing-impaired listeners, it may be helpful to create a large HFE signal-to-noise ratio by ensuring the target talker is directly facing the listener. However, this might also necessitate improvements in hearing aid technology to effectively amplify HFE.

ACKNOWLEDGMENTS

This work was carried out at Brigham Young University in Provo, UT, and was funded by NIH Grant Nos. F31DC010533 and R01DC8612. The authors thank Dr. Andrew Lotto for helpful suggestions with the recording material, and Dr. Kent Gee and the BYU Acoustics Research Group for use of facilities and equipment.

APPENDIX

See Table I for overall and octave band levels.

TABLE I. Overall and octave band levels for all conditions at all angles. Values are given in dB SPL at 1 m.

Angle	Overall								
	level	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz	16 kHz
Normal speech									
0	62.0	50.2	56.0	57.4	55.0	51.2	46.3	46.4	38.0
15	62.4	51.0	56.7	58.0	54.8	50.2	44.6	45.4	37.9
30	61.8	50.1	55.8	57.5	54.3	51.1	45.6	44.7	37.5
45	61.5	49.8	55.5	57.4	53.6	50.4	44.7	43.3	35.7
60	61.1	49.6	55.2	57.3	53.5	48.7	42.4	40.6	33.6
75	60.5	49.1	54.4	56.7	53.6	46.5	41.0	39.4	32.9
90	59.8	48.6	53.7	55.8	53.4	44.2	40.0	37.7	30.8
105	58.8	48.1	52.9	54.6	52.5	43.6	37.3	34.9	27.9
120	57.5	47.5	52.1	53.0	50.4	43.3	33.6	31.4	23.8
135	56.4	47.1	51.5	51.9	48.0	41.9	31.8	28.6	20.2
150	55.8	46.7	51.1	51.4	47.1	38.2	29.6	24.9	16.3
165	55.8	46.6	51.0	51.4	48.1	36.2	25.3	21.4	14.5
180	56.0	46.6	50.9	51.5	48.7	38.1	27.4	20.1	11.3
Female speech									
0	61.3	39.0	55.6	56.7	54.8	51.1	45.0	46.7	39.2
15	61.6	39.7	56.3	57.3	54.6	50.1	43.4	45.6	39.0
30	61.0	38.8	55.4	56.8	54.0	51.0	44.2	45.0	38.6
45	60.6	38.4	55.0	56.6	53.4	50.2	43.2	43.7	36.5
60	60.2	38.1	54.7	56.5	53.4	48.3	40.8	41.1	34.4
75	59.6	37.5	54.0	55.8	53.4	46.0	39.2	40.1	33.9
90	58.9	36.9	53.2	54.9	53.2	43.8	38.3	38.4	32.1
105	57.9	36.4	52.5	53.7	52.2	43.3	35.6	35.5	29.1
120	56.5	35.8	51.7	52.2	50.1	43.1	31.6	31.8	24.9
135	55.3	35.4	51.2	51.1	47.7	41.8	30.4	29.1	21.2
150	54.8	35.1	50.9	50.6	46.9	38.2	28.4	25.4	17.3
165	54.8	34.9	50.8	50.6	47.8	35.8	23.9	22.0	15.4
180	55.0	34.9	50.7	50.6	48.5	38.1	26.5	19.8	11.8
Male speech									
0	62.7	54.0	56.2	58.0	55.0	51.0	47.5	45.6	34.7
15	63.1	54.9	57.0	58.7	54.9	50.0	45.9	44.5	34.1
30	62.6	54.0	56.1	58.2	54.4	50.9	46.9	43.8	34.0
45	62.3	53.7	55.8	58.1	53.6	50.3	46.2	42.0	33.1
60	62.0	53.6	55.5	58.1	53.3	48.8	43.9	39.0	31.0
75	61.4	53.0	54.7	57.5	53.3	46.8	42.6	37.2	29.0
90	60.7	52.6	53.9	56.6	53.2	44.4	41.6	35.5	25.5
105	59.7	52.1	53.1	55.4	52.4	43.5	39.0	32.9	22.3
120	58.4	51.5	52.1	53.6	50.4	43.0	35.3	29.7	19.0
135	57.3	51.1	51.4	52.4	47.8	41.5	33.2	26.9	16.0
150	56.8	50.8	50.9	52.1	46.9	37.5	30.7	23.0	11.7
165	56.8	50.7	50.6	52.1	47.9	36.3	26.6	19.0	10.7
180	56.9	50.6	50.6	52.2	48.4	37.5	28.2	20.0	9.5
Soft speech									
0	54.8	46.9	50.7	48.6	43.3	38.8	38.9	41.8	32.2
15	55.4	47.7	51.5	49.3	43.2	37.8	37.4	41.0	31.8
30	54.7	46.8	50.5	48.7	42.7	38.7	38.4	40.7	31.0
45	54.3	46.5	50.2	48.5	42.2	38.0	37.6	39.4	29.3
60	54.0	46.3	49.9	48.4	42.2	36.4	35.6	36.8	27.6
75	53.4	45.8	49.2	47.8	42.3	34.4	34.4	35.5	27.0
90	52.6	45.2	48.5	46.8	42.1	32.3	33.7	33.7	25.1
105	51.8	44.7	47.8	45.7	41.2	31.4	31.0	31.1	22.2
120	50.8	44.0	47.0	44.2	39.1	31.1	27.3	27.5	18.4
135	50.0	43.6	46.5	43.1	36.6	29.8	25.5	24.9	15.1

TABLE I. (Continued.)

Angle	Overall									
	level	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz	16 kHz	
150	49.7	43.3	46.2	42.6	35.8	26.2	23.3	21.5	11.7	
165	49.5	43.1	46.0	42.5	36.7	24.1	18.7	18.2	10.5	
180	49.6	43.1	46.0	42.5	37.3	25.9	20.6	16.7	8.9	
Loud speech										
0	73.8	55.4	64.0	68.5	69.1	66.3	61.5	54.7	47.6	
15	73.6	55.9	64.5	68.9	68.7	65.2	59.7	53.6	47.3	
30	73.4	55.1	63.7	68.5	68.1	66.1	60.6	53.0	47.1	
45	72.8	54.6	63.3	68.4	67.3	65.4	59.6	51.2	45.3	
60	72.2	54.2	62.9	68.2	66.8	63.8	57.3	48.2	42.8	
75	71.5	53.5	62.1	67.5	66.8	61.7	55.5	46.9	42.2	
90	70.7	52.9	61.2	66.6	66.6	59.3	54.2	45.2	39.9	
105	69.6	52.3	60.4	65.3	65.7	58.2	51.7	42.5	36.8	
120	68.0	51.7	59.5	63.6	63.6	57.8	47.5	39.0	32.3	
135	66.5	51.2	58.8	62.5	61.1	56.6	45.5	35.9	28.6	
150	65.8	51.0	58.5	62.2	60.4	53.1	43.9	32.4	25.0	
165	66.1	50.9	58.3	62.3	61.5	51.1	39.7	28.6	21.4	
180	66.6	51.1	58.5	62.6	62.3	53.1	41.9	28.0	18.4	
Singing										
0	73.9	55.8	64.2	70.8	67.7	63.3	59.9	50.2	42.3	
15	74.1	56.5	64.8	71.4	67.6	62.4	58.3	49.3	42.1	
30	73.7	55.5	63.9	70.9	67.0	63.0	59.6	48.6	41.8	
45	73.3	55.0	63.5	70.8	66.5	62.1	58.6	46.9	40.3	
60	73.0	54.8	63.1	70.7	66.6	60.4	56.3	44.1	38.5	
75	72.5	54.1	62.3	70.0	66.8	58.4	54.9	42.7	37.6	
90	71.7	53.5	61.4	69.0	66.6	56.1	53.8	40.6	34.8	
105	70.6	52.9	60.6	67.8	65.7	55.3	50.7	37.9	31.5	
120	69.0	52.4	59.7	66.2	63.8	55.0	47.1	34.4	27.7	
135	67.6	52.1	59.0	65.0	61.3	53.9	46.2	32.0	24.2	
150	66.8	51.7	58.5	64.5	60.1	50.4	43.9	28.5	20.2	
165	66.8	51.6	58.2	64.4	60.8	47.7	38.4	25.2	18.6	
180	67.2	51.4	58.2	64.5	61.7	49.9	41.3	23.4	15.7	
/s/										
0	58.5	32.7	31.0	28.4	28.4	36.1	51.1	57.0	48.2	
15	57.6	33.3	31.6	28.9	28.1	34.8	49.9	56.1	47.8	
30	57.4	32.7	30.9	28.5	27.8	36.1	51.1	55.4	47.4	
45	56.0	32.3	30.5	28.3	26.9	35.7	50.5	53.8	45.7	
60	53.4	32.2	30.1	28.2	26.2	34.8	48.0	50.9	43.7	
75	52.1	31.9	29.7	27.9	26.2	34.0	46.7	49.5	42.5	
90	50.6	31.9	29.5	27.4	26.5	31.6	46.0	47.7	39.9	
105	48.0	31.8	29.2	26.8	26.1	29.7	43.3	44.9	36.9	
120	44.9	31.4	28.7	25.6	24.7	29.2	39.9	41.4	33.0	
135	42.7	31.5	28.6	25.0	23.3	28.3	37.4	38.5	29.2	
150	40.4	31.4	28.5	24.7	22.3	25.6	34.8	34.8	24.4	
165	38.0	31.0	28.1	24.0	21.8	23.1	30.1	30.6	20.9	
180	38.4	31.2	28.4	24.5	22.4	23.7	31.3	30.5	18.1	
/sh/										
0	57.2	34.0	30.6	28.0	29.1	49.2	54.9	50.8	37.7	
15	55.7	34.5	31.1	28.7	28.8	47.5	53.3	49.4	37.0	
30	56.3	33.9	30.6	28.4	28.5	48.9	54.4	48.2	36.9	
45	55.4	33.5	30.2	28.1	27.3	48.4	53.6	46.2	35.7	
60	53.6	33.2	29.9	27.9	26.2	47.7	51.5	43.4	32.8	
75	52.8	33.0	29.6	27.6	26.3	46.7	50.8	41.9	30.9	
90	51.4	33.1	29.6	27.4	26.7	44.2	49.7	40.1	27.8	
105	49.2	33.1	29.4	26.8	26.6	42.2	47.3	38.2	24.8	
120	46.6	32.6	28.6	25.6	25.3	41.3	43.7	34.9	21.3	
135	45.2	32.7	28.6	25.1	24.1	40.9	41.6	32.5	18.0	
150	43.1	32.6	28.4	24.7	23.2	37.5	39.6	28.3	14.5	
165	40.2	32.0	27.7	23.7	22.4	34.7	35.2	24.5	13.5	
180	41.6	32.4	28.3	24.3	22.8	36.2	37.5	23.5	12.1	

TABLE I. (Continued.)

Angle	Overall									
	level	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz	16 kHz	
/f/										
0	44.0	28.8	27.1	28.5	32.2	33.9	33.2	39.2	38.7	
15	43.6	28.5	27.1	28.8	31.8	33.0	31.7	38.9	38.7	
30	44.0	28.9	27.0	28.5	31.3	34.6	33.1	38.3	39.4	
45	42.9	28.2	26.3	28.3	30.0	34.3	32.9	36.3	38.4	
60	40.9	27.2	25.4	28.0	28.6	33.5	31.3	33.6	35.5	
75	40.2	27.6	25.4	27.6	28.5	32.4	31.0	32.5	33.8	
90	39.3	28.5	25.9	27.3	28.9	30.3	30.7	31.1	31.3	
105	38.3	28.7	26.0	26.6	28.8	29.1	28.6	29.0	28.4	
120	36.9	28.1	25.2	25.3	27.3	28.7	25.4	26.4	24.6	
135	36.3	28.4	25.4	24.7	25.5	27.9	23.7	24.3	21.0	
150	35.6	28.3	25.3	24.3	24.3	25.5	21.7	21.2	16.6	
165	34.4	27.3	24.3	23.6	24.0	23.1	18.6	17.9	15.2	
180	35.2	28.2	25.3	24.3	24.9	23.6	18.1	16.1	13.2	
/th/										
0	43.8	33.2	32.1	29.9	28.8	31.0	31.9	36.7	39.5	
15	43.6	33.8	32.7	30.3	28.7	30.0	30.7	36.4	38.8	
30	43.9	33.2	32.1	29.9	28.5	31.7	32.2	36.7	39.5	
45	43.4	32.9	31.6	29.5	27.7	31.8	31.9	35.7	39.0	
60	42.1	32.6	31.2	29.2	27.0	31.0	30.6	33.2	37.2	
75	41.5	32.5	30.9	28.9	27.1	30.1	30.2	32.2	36.2	
90	40.8	32.8	30.8	28.5	27.4	28.7	30.1	31.3	33.5	
105	40.0	32.8	30.4	27.9	27.1	27.6	28.5	29.6	30.8	
120	38.9	32.5	29.7	26.8	25.8	27.1	25.8	27.0	27.5	
135	38.6	32.6	29.6	26.5	24.8	26.4	24.2	25.1	24.5	
150	38.1	32.5	29.4	26.3	24.0	24.4	22.1	22.2	20.0	
165	37.3	32.0	28.8	25.6	23.2	22.3	19.0	19.2	17.8	
180	37.7	32.4	29.1	26.1	23.5	22.1	18.5	17.4	15.1	
Ambient noise ^a										
0	27.8	15.1	5.6	5.4	7.0	8.7	10.4	11.0	10.3	
15	28.5	16.2	5.7	5.9	7.8	8.8	8.6	8.2	6.8	
30	27.9	15.0	5.8	4.8	6.6	8.4	10.0	10.6	9.3	
45	27.8	14.8	6.4	5.3	6.8	8.6	10.3	11.1	10.1	
60	28.1	15.0	6.7	5.8	7.8	9.5	10.7	11.1	10.1	
75	28.2	15.9	5.7	4.9	6.4	8.2	9.5	9.9	8.6	
90	28.5	16.4	5.1	5.1	6.6	8.4	10.2	10.9	10.0	
105	28.5	15.9	5.9	5.6	7.2	9.0	10.6	11.2	10.5	
120	28.6	15.9	6.5	5.7	7.1	8.8	10.1	10.5	9.7	
135	28.9	16.6	6.6	5.2	6.2	7.8	9.5	10.6	10.5	
150	29.0	16.8	6.1	5.7	7.3	9.0	10.7	11.4	10.6	
165	29.3	17.1	6.2	4.4	6.9	9.4	11.7	13.7	15.4	
180	29.1	17.2	3.5	2.7	4.4	6.8	9.1	10.6	10.6	

^aAmbient noise levels are reported as exact measured values. As all other values reported here were extrapolated to 1 m from the mouth using the inverse square law, accurate comparison to ambient noise levels requires extrapolation back to 60 cm (an increase of ~4.44 dB).

Badri, R., Siegel, J. H., and Wright, B. A. (2011). "Auditory filter shapes and high-frequency hearing in adults who have impaired speech in noise performance despite clinically normal audiograms," *J. Acoust. Soc. Am.* **129**, 852–863.

Best, V., Carlile, S., Jin, C., and van Schaik, A. (2005). "The role of high frequencies in speech localization," *J. Acoust. Soc. Am.* **118**, 353–363.

Bozzoli, F., Viktorovitch, M., and Farina, A. (2005). "Balloons of directivity of real and artificial mouth used in determining speech transmission index," *Convention Papers, 118th Audio Engineering Society Convention*, Barcelona, Paper 6492.

- Cabrera, D., Davis, P. J., and Connolly, A. (2011). "Long-term horizontal vocal directivity of opera singers: Effects of singing projection and acoustic environment," *J. Voice* **25**, e291–e303.
- Chu, W. T., and Warnock, A. C. C. (2002). "Detailed directivity of sound fields around human talkers," *Technical Report, Institute for Research in Construction* (National Research Council of Canada, Ottawa ON, Canada), pp. 1–47.
- Dunn, H. K., and Farnsworth, D. W. (1939). "Exploration of pressure field around the human head during speech," *J. Acoust. Soc. Am.* **10**, 184–199.
- Flanagan, J. L. (1960). "Analog measurements of sound radiation from the mouth," *J. Acoust. Soc. Am.* **32**, 1613–1620.
- Fullgrabe, C., Baer, T., Stone, M. A., and Moore, B. C. (2010). "Preliminary evaluation of a method for fitting hearing aids with extended bandwidth," *Int. J. Aud.* **49**, 741–753.
- Halkosaari, T., Vaalgamaa, M., and Karjalainen, M. (2005). "Directivity of artificial and human speech," *J. Audio Eng. Soc.* **53**, 620–631.
- Jongman, A., Wayland, R., and Wong, S. (2000). "Acoustic characteristics of English fricatives," *J. Acoust. Soc. Am.* **108**, 1252–1263.
- Jers, H. (2007). "Directivity measurements of adjacent singers in a choir," *19th International Congress on Acoustics*, Madrid, Spain.
- Lippmann, R. P. (1996). "Accurate consonant perception without mid-frequency speech energy," *IEEE Trans. Speech Aud. Proc.* **4**, 66–69.
- Maniwa, K., Jongman, A., and Wade, T. (2009). "Acoustic characteristics of clearly spoken English fricatives," *J. Acoust. Soc. Am.* **125**, 3962–3973.
- Marshall, A. H., and Meyer, J. (1985). "The directivity and auditory impressions of singers," *Acustica* **58**, 130–140.
- McKendree, F. S. (1986). "Directivity indices of human talkers in English speech," *Proceedings of Inter-Noise 86*, Cambridge, pp. 911–916.
- Monson, B. B. (2011). "High-frequency energy in singing and speech," Doctoral dissertation, University of Arizona, Tucson, AZ.
- Monson, B. B., Lotto, A. J., and Ternström, S. (2011). "Detection of high-frequency energy changes in sustained vowels produced by singers," *J. Acoust. Soc. Am.* **129**, 2263–2268.
- Moore, B. C., Füllgrabe, C., and Stone, M. A. (2010). "Effect of spatial separation, extended bandwidth, and compression speed on intelligibility in a competing-speech task," *J. Acoust. Soc. Am.* **128**, 360–371.
- Moreno, A., and Pfretzschner, J. (1978). "Human head directivity and speech emission: A new approach," *Acoust. Lett.* **1**, 78–84.
- Pittman, A. L. (2008). "Short-term word-learning rate in children with normal hearing and children with hearing loss in limited and extended high-frequency bandwidths," *J. Speech Lang. Hear. Res.* **51**, 785–797.
- Rayleigh, L. (1908). "Acoustical notes.-VIII," *Philos. Mag.* **16**, 235–246.
- Spitzer, S. M., Liss, J. M., and Mattys, S. L. (2007). "Acoustic cues to lexical segmentation: A study of resynthesized speech," *J. Acoust. Soc. Am.* **122**, 3678–3687.
- Stelmachowicz, P. G., Pittman, A. L., Hoover, B. M., and Lewis, D. E. (2001). "Effect of stimulus bandwidth on the perception of /s/ in normal- and hearing-impaired children and adults," *J. Acoust. Soc. Am.* **110**, 2183–2190.