DOI: 10.61782/fa.2025.0448

FORUM ACUSTICUM
EURONOISE 2025

IMPACT OF LEG DIFFRACTION AND SCATTERING ON VOICE DIRECTIVITY PATTERNS

Samuel D. Bellows[1,2,3*] **Brian F. G. Katz**[2] **Timothy W. Leishman**[3]

[1] Department of Electrical and Computer Engineering, University of Utah Asia Campus, South Korea
[2] Institue Jean le Rond d'Alembert, Sorbonne University, France
[3] Department of Physics and Astronomy, Brigham Young University, USA

ABSTRACT

Understanding sound radiation from the human voice has broad applications in room acoustical design, telecommunications, physical modeling of the singing voice, and virtual acoustics. Head simulators and head and torso simulators can provide simplified approximations to voice directivity, which motivate their use in room acoustical and other related measurements. Nonetheless, recent works have shown that scattering and diffraction due to the torso alter speech radiation patterns compared to those produced from an isolated head alone. Despite the improvements that including a torso provides, most commercial voice simulators neglect the effects of human legs. To better understand the impact of leg scattering and diffraction on voice directivity, this work presents measurements of a manikin with a head, torso, and legs. Comparing the results with those measured from human talkers shows that scattering and diffraction from human legs can impact voice radiation patterns, particularly above 1 kHz. The results also highlight the importance of high spatial sampling resolution when performing directivity measurements, as these scattering effects are easily spatially aliased in lower resolution sampling schemes.

Keywords: *voice, speech, directivity, sound radiation, room acoustics.*

**Corresponding author*: *samuel.bellows11@gmail.com.*

1. INTRODUCTION

The significance of voice directivity in applications such as telecommunications [1] or room acoustical design [2] has motivated numerous investigations on its characteristics [3–6]. Nonetheless, its technical assessment for the standardized spatial resolutions [7] that are commonly employed in architectural acoustics simulation packages [8,9] comes with many practical challenges. For example, because mouth shape influences sound radiation, voice directivity is inherently time-varying and phoneme dependent [1, 10–13]. As human subjects cannot exactly repeat the same speech utterances or sung musical passages during recording procedures, obtaining high-resolution directivity measurements for practical applications has required time-averages over running speech [6, 14], isolated, steady-state phonemes [13, 15], or interpolation techniques applied to lower-resolution results [12]. Each method has distinct advantages and limitations, making proper evaluation of voice radiation a difficult task. To circumvent these and other challenges, the use of head simulators (HSs) [16], head and torso simulators (HATSs) [1,11,17], and entire manikins [18,19] have become ubiquitous in voice directivity measurements. Nonetheless, the differences between these approximations and actual human voice radiation remain unclear.

The advantage of an HS and HATS is that they incorporate some basic human geometry compared to simplistic theoretical models, such as a radially vibrating cap on a sphere [11,14,18]. However, researchers and practitioners often indiscriminately apply these approximations without considering how geometrical deviations between the manikins and humans could influence their results. More recently, Brandner et al. [11] showed that torso diffrac-

tion impacts speech radiation in the transverse and median polar planes, particularly at higher frequencies. A study on the impact of head orientation on voice directivity likewise confirmed these trends [17]. Nonetheless, few studies have used legs with the torso and head; Flanagan [18] presented results from a standing manikin while Bellows and Leishman [19] contains some initial results for a seated KEMAR HATS with legs. While several studies include some analysis between select configurations (HS to HATS in [1, 11, 17] or HATS to human talkers in [6, 14]), no study has systematically analyzed differences between HS, HATS, full-body manikins, and human talkers.

This work addresses these needs by comparing measured high-resolution, spherical directivities of an HS, a HATS, a seated HATS with legs, and a human talker. Directivity balloons show differences between the different configurations at frequencies as low as 500 Hz. The results demonstrate that torso diffraction and scattering from the legs influence speech radiation. Directivity index and directivity factor function deviation levels quantify differences between the configurations.

2. METHODS

2.1 Measurement System

As suggested by Fig. 1, three different manikin configurations enabled an investigation on the impact of the torso and human legs on voice directivity. The manikin under consideration was the GRAS KEMAR 45BC HATS. The first configuration (Fig. 1(a)) was the isolated dummy head mounted in the center of the array. The second configuration (Fig. 1(b)) included the manikin torso. The final manikin configuration (Fig. 1(c)) included legs and a chair.

The directivity measurement system consisted of 36 precision 12.7 mm (0.5 in) microphones spaced in 5° increments in the polar angle along a semi-circular arc. Arc rotations around the manikin configurations produced a spherical scanning surface with dual-equiangular 5° resolution. The resultant 2,521 unique sampling positions are consistent with the Type-A AES directivity sampling standard used in many architectural acoustics simulation packages [7–9]. For the isolated head and combined head and torso configurations, the mouth was aligned to the center of the $R = 0.97$ m radius measurement array. The increased size of the seated configuration required moving the array microphones to a $R = 1.17$m radius. In this configuration, the mouth did not exactly align with the center of the array. The comparative human voice directivity measurements came from time-averaged speech results in [14]. Because subject averaging smooths many fine radiation details [14] and because body size strongly influences radiation characteristics, this work presents comparisons between a single female talker whose frequency-dependent trends aligned well with results produced by the KEMAR 45BC HATS.

2.2 Data Processing

The measurement input signal consisted of three repetitions of a 5 s exponentially weighted swept-sine. Narrowband frequency response functions (FRFs) between the electrical input signal and measured pressure at the array positions followed from cross and autospectral density estimates [14, 17]. Relative calibrations over the measurement channels corrected for variations that would appear as latitudinal bands in the directivity data.

Because the measurements were made at different radii, the mouth did not align at the same position, and the array radius did not fall in the far field for all frequencies, far-field propagation via spherical harmonic expansions is necessary [19–21]. The solution to the Helmholtz equation in spherical coordinates for $r > R$ is [19, 22]

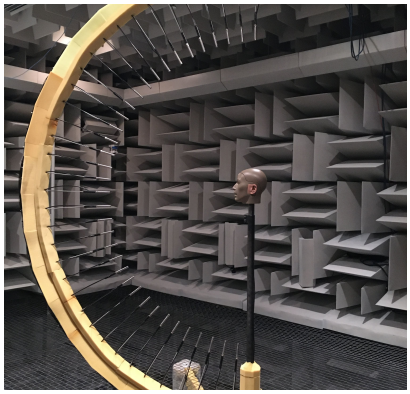
$$p(r,\theta,\phi,k) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} c_n^m(k) h_n^{(2)}(kr) Y_n^m(\theta,\phi), \quad (1)$$

where k is the wavenumber, c_n^m are the frequency-dependent expansion coefficients, $h_n^{(2)}$ are the spherical Hankel functions of the second kind (for $e^{i\omega t}$ time-dependance) and Y_n^m are the spherical harmonics. The expansion coefficients followed from numerical integration [23]. Applying the large-argument approximation of the spherical Hankel functions yields an unnormalized far-field directivity pattern [21]

$$D_{ff}(\theta,\phi,k) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} c_n^m(k) i^{n+1} Y_n^m(\theta,\phi). \quad (2)$$

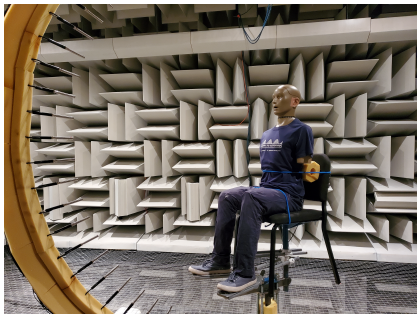
The magnitude of this pattern is invariant to source position or measurement radius [20], thus allowing fair comparisons across the three measurement configurations.

Avoiding spatial aliasing effects in the far-field results requires careful analysis of the convergence of the spherical harmonic expansions. The sampling density applied in this work allows expansions up to degree $N_s = 34$ [23].



(a)

(b)



(c)

Figure 1: Three different measurement configurations of the KEMAR HATS. (a) Isolated head simulator (HS). (b) Head and torso simulator (HATS). (c) Head, torso, legs, and chair (HTALS).

Convergence analysis of spherical harmonic expansions to the measured FRF values on the array surface indicated that this sampling resolution was sufficient for spherical-harmonic-based analysis up to approximately 4 kHz [24]. Above this frequency, narrowband interpolations or extrapolations of measured data are not reliable. Magnitude narrowband or broadband directivities can provide important insights above these frequencies, but the impact of measurement radius and source placement should be also taken into account.

The directivity factor function follows from the far-field directivity function as [22]

$$Q(\theta,\phi,f) = \frac{4\pi|D_{ff}(\theta,\phi,f)|^2}{\int_0^{2\pi}\int_0^{\pi}|D_{ff}(\theta,\phi,f)|^2\sin\theta d\theta d\phi} \tag{3}$$

and may be expressed on a logarithmic scale as the directivity index (DI):

$$\mathrm{DI}(\theta,\phi,f) = 10\log_{10}Q(\theta,\phi,f). \tag{4}$$

A directivity factor function deviation level [17, 21]

$$L_Q = 10\log_{10}(1+\sigma_Q) \tag{5}$$

with

$$\sigma_Q = \frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi}|Q_1(\theta,\phi,f) - Q_2(\theta,\phi)|\sin\theta d\theta d\phi \tag{6}$$

may be used to quantify differences between two directivity patterns.

3. RESULTS

Distinct differences between the directivities produced by each of the four conditions (HS, HATS, HTALS, and the human talker) appear, even at mid-to-lower frequencies. Figure 2 compares far-field projected narrowband (1 Hz resolution) results for five different frequencies: 250 Hz, 500 Hz, 800 Hz, 1.2 kHz, and 1.6 kHz. Directivity balloon color and radius indicate relative levels on a logarithmic scale. In all balloons, the mouth faces the 0° azimuthal marker so that the vantage point is from the right side of the body.

At 250 Hz, the lowest relative levels of the HS results remain within -3 dB, indicating that the source behaves roughly omnidirectionally. For the HATS, the lowest relative levels approach -5 dB, indicating that the increased dimensions of the body make the source slightly more directional. The HTALS and human talker results show the

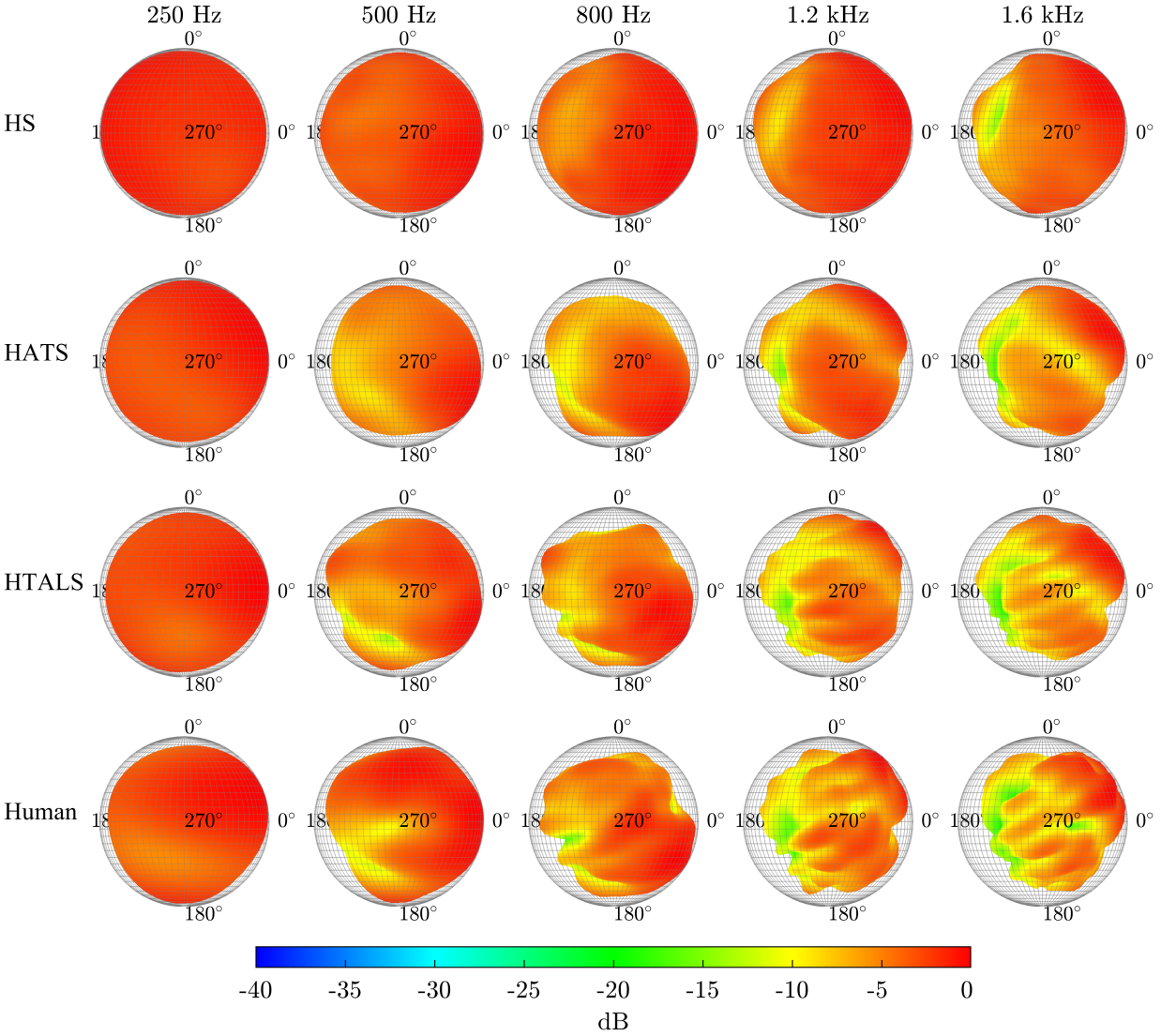


Figure 2: Comparison between the narrowband directivity of the HS, HATS, HTALS, and female talker at 250 Hz, 500 Hz, 800 Hz, 1.2 kHz, and 1.6 kHz.

greatest deviations from omnidirectionality, with relative levels approaching -7 dB and -9 dB, respectively.

Differences become even more pronounced at 500 Hz. While the HS still appears roughly omnidirectional, the HATS, HTALS, and human talker have developed a radiation lobe directed below the mouth, at an elevation angle of roughly -30°. While the HATS shows more diffraction effects relative to the HS, significantly reduced levels behind and below the HTALS and human talker appear in the region corresponding to the chair.

These trends continue further for the HATS, HTALS, and human talker at 800 Hz, including more radiation directed slightly below the mouth axis. Additionally, more detailed and intricate diffraction lobes appear around the body for the HTALS and human talker that are not apparent for the HATS. The HS results, while becoming more directional, appear most dissimilar from those of the HATS, HTALS, and human talker.

At 1.2 kHz, the primary direction of radiation for the HATS, HTALS, and human talker shifts to slightly above the mouth axis, at an elevation angle of roughly 45° (the HS remains roughly on-axis). Significantly reduced levels appear behind and below the source for these three conditions, highlighting the impact of diffraction and scattering off of the body compared to results from the HS.

The impact of the legs and chair becomes even more pronounced at this frequency. Comparing the results from the HATS and HTALS reveals evenly-spaced lobes or striations appearing in the balloon. Because the only difference between the HATS and HTALS is the chair and legs, these interference-like effects may be due to backscattering from the legs and chair. These same patterns appear in the balloon of the human talker, although their relative locations and orientation differ slightly. These variations are likely due to differing geometry and absorption properties of the human legs compared to the manikin approximation.

Similar scattering effects appear in the balloons of the human talker and HTALS but not the HATS at 1.6 kHz. At this frequency, the primary radiation direction has slightly lowered, although diffraction and absorption features behind the head appear similar to those at 1.2 kHz. For the HS, reduced levels appear behind the body, although they are not as significant as for the HATS, HTALS, and human talker.

Although the narrowband balloons illustrate some of the fine, frequency-dependent features of voice directivity, many common applications of directional data employ 1/1 and 1/3-octave-band data. Summation over frequency smooths many of the unique features appearing in narrowband results, although the general characteristics remain the same. Figure 3 plots the 1/3-octave band results of the four conditions for the 250 Hz, 500 Hz, 1 kHz, and 1.6 kHz bands.

At the 250 Hz, 500 Hz, and 800 Hz bands, differences between the narrowband and broadband results are minimal. However, the larger bandwidths of the 1 kHz and 1.6 kHz 1/3-octave bands do smooth the directivity balloons appreciably. For example, while the effects of scattering from the legs and chair distinctly appear in the narrowband HTALS and human talker balloons of Fig. 2, only small traces of these effects appear in the respective 1/3-octave-band balloons appearing in Fig. 3. Indeed, the results of the HTALS and human talker in these bands appear more similar to those of the HATS. Nonetheless, for the 1.6 kHz 1/3-octave-band results, the direction of the HATS primary lobe and null differs from that of the HTALS and human talker.

4. ANALYSIS

Visual inspection of the directivity balloons reveals several qualitative trends of each configuration. At lower frequencies, differences between all the configurations appear to be slight. By 500 Hz, the HS results show pronounced differences from the HATS, HTALS, and human talker results. The impact of the legs and chair become more visible near and beyond 1 kHz, although 1/1 and 1/3-octave bands smooth these effects.

To further quantify these trends, Fig. 4 plots the maximum 1/3-octave-band DI values over the sphere for each of the four configurations. In general, the DI of the isolated head is much lower than for the DIs of the other cases, which all exceed 1 dB by 200 Hz. Thus, using a torso over an isolated head provides more quantitatively realistic results. Interestingly, up to 1 kHz, the DIs for the HATS, HTALS, and female talker remained within 1 dB. Even above 1 kHz, the DIs have similar values. These results seem consistent with the directivity balloons, which only showed minor differences due to leg and chair scattering; the directional properties and relative amplitudes were similar even at higher frequencies.

To further assess deviations between the configurations, Fig. 5 compares L_Q values between the different sources. Because differences in body and mouth size and shape influence directivity patterns, Fig. 5(a) uses the HTALS for the reference directivity rather than the human talker. This choice ensures that the head size, mouth

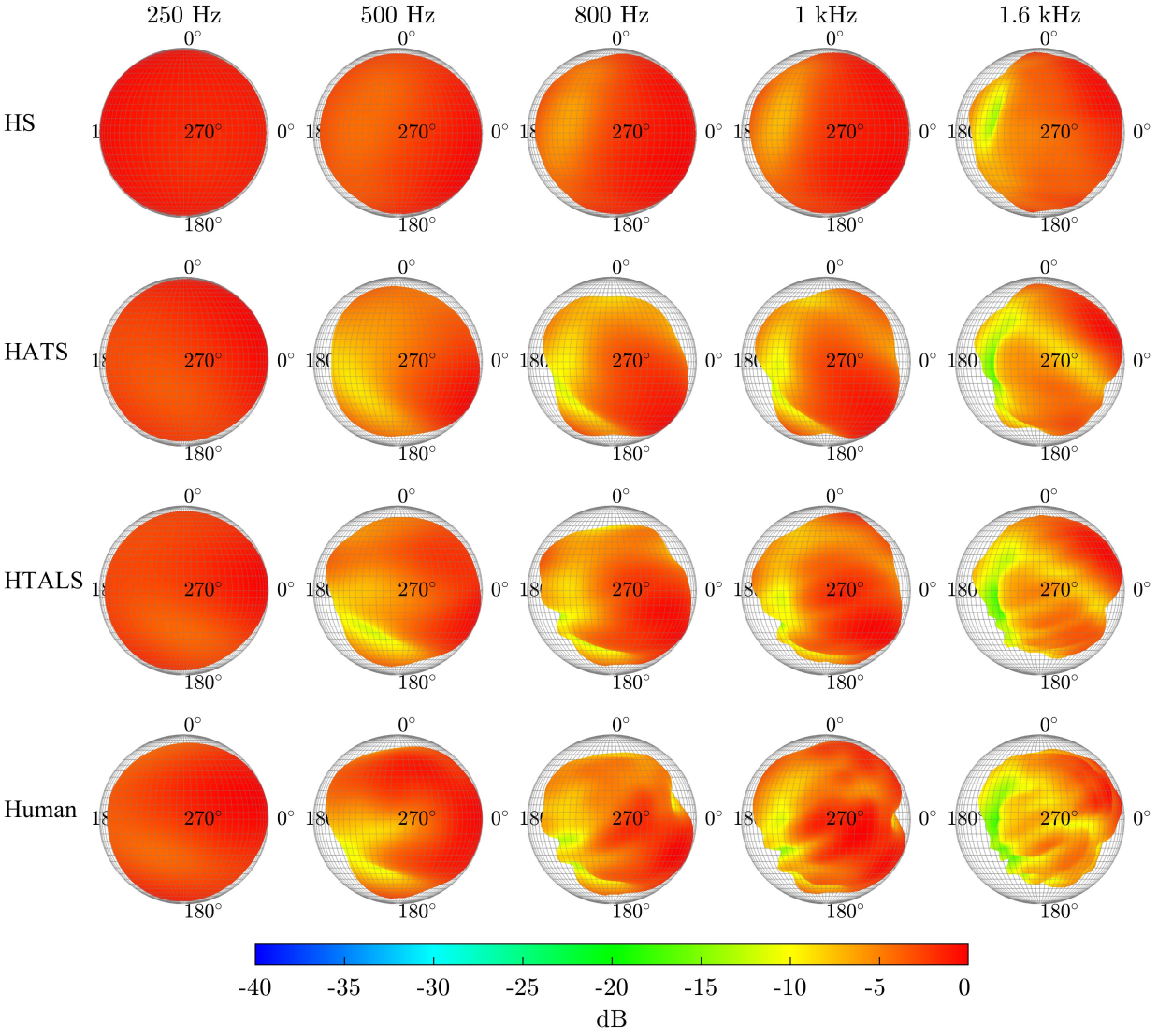


Figure 3: Comparison between the 1/3-octave-band directivity of the HS, HATS, HTALS, and female talker at the center frequencies 250 Hz, 500 Hz, 800 Hz, 1 kHz, and 1.6 kHz.

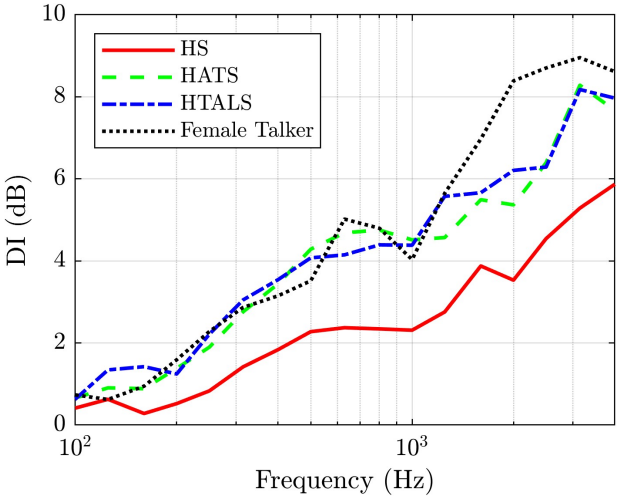


Figure 4: The 1/3-octabe band directivity indices for the four configurations.

shape, etc., remain consistent across configurations.

At the lowest frequencies, the deviations are less than 1 dB. Deviations exceed 1.0 dB by 315 Hz for the HS and 630 Hz for the HATS. Additionally, the deviations exceed 2.0 dB at 2.5 kHz for the HS, while the HATS remains below this value up to at least 4 kHz. Both curves show fairly consistent increases in deviation errors of approximately 1.2 dB per decade.

Figure 5(b) shows values of L_Q but with the seated female talker as the reference. The deviations exceed 1.0 dB by 250 Hz for the HS, 400 Hz for the HATS, and 1 kHz for the HTALS. Although the deviations for the HATS and HTALS are lower than those for the HS, above 1 kHz their deviation levels are similar.

5. CONCLUSIONS

This work compared the directivities of an HS, HATS, and HTALS with results from a seated human talker. Directivity balloons and derived directivity indices showed variations arising even by 500 Hz between the differing configurations. Overall, neglecting torso diffraction (HS configuration) led to the most significant deviations relative to the HTALS or human talker. Scattering from the legs and chair also altered radiation patterns and led to differences between the HATS and HTALS configurations at higher frequencies.

Future work should evaluate the perceptual relevance of each configuration and the impact on room acoustical measurements. Additional work could compare results from other commercial HS and HATS to further understand differences that arise when approximating human voice radiation using simulators. In addition, acoustic reciprocity suggests that torso, leg, and chair diffraction and scattering effects may reduce the reliability of head-related transfer function (HRTF) calculations derived from head geometry alone, motivating further research into this area.

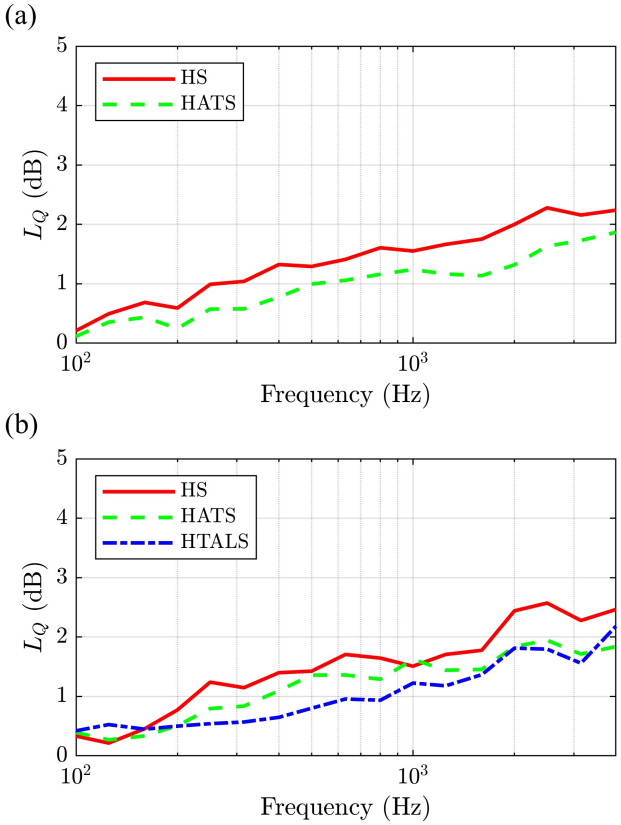


Figure 5: Directivity factor function deviation levels between the HS, HATS, and HTALS with the (a) HTALS and (b) female talker as reference.

6. ACKNOWLEDGMENTS

The William James and Charlene Fuhriman Strong Family Endowed Fellowship Fund for Musical Acoustics and the SONICOM project (European Union's Horizon 2020 research and innovation program grant agreement No. 101017743) supported this work.

7. REFERENCES

[1] T. Halkosaari, M. Vaalgamaa, and M. Karjalainen, "Directivity of artificial and human speech," *J. Audio. Eng. Soc.*, vol. 53, no. 7/8, pp. 620–631, 2005.

[2] D. Cabrera, P. J. Davis, and A. Connolly, "Long-term horizontal vocal directivity of opera singers: Effects of singing projection and acoustic environment," *J. Voice*, vol. 25, no. 6, pp. 291–303, 2011.

[3] H. Dunn and D. W. Farnsworth, "Exploration of pressure field around the human head during speech," *J. Acoust. Soc. Am.*, pp. 184–199, January 1939.

[4] F. S. McKendree, "Directivity indices of human talkers in english speech," *Paper presented at Inter-Noise 86 Conference: Progress in Noise Control, Cambridge, USA*, pp. 911–916, 1986.

[5] A. Moreno and J. Pretzschner, "Human head directivity in speech emission: A new approach," *Acoustics Letters*, vol. 1, pp. 78–84, 1978.

[6] W. T. Chu and A. C. Warnock, "Detailed directivity of sound fields around the human head during speech," *National Research Council Canada*, vol. IRC-RR-104, pp. 1–47, 2002.

[7] *AES56-2008 (r2019): AES standard on acoustics - Sound source modeling - Loudspeaker polar radiation measurments*. New York, New York: Audio Engineering Society, August 2019.

[8] CLF Group, "CLF: A common loudspeaker format," *Syn-Aud-Con Newsl.*, vol. 32, pp. 14–17, 2004.

[9] *GLL Loudspeaker File Format*. Ahnert Feistel Media Group, 2021.

[10] B. B. Monson, E. J. Hunter, and B. H. Story, "Horizontal directivity of low- and high-frequency energy in speech and singing," *J. Acoust. Soc. Am.*, vol. 132, pp. 433–441, 2012.

[11] M. Brandner, R. Blandin, M. Frank, and A. Sontacchi, "A pilot study on the influence of mouth configuration and torso on singing voice directivity," *J. Acoust. Soc. Am.*, vol. 148, no. 3, pp. 1169–1180, 2020.

[12] C. Pörschmann and J. M. Arend, "Investigating phoneme-dependencies of spherical voice directivity patterns," *J. Acoust. Soc. Am.*, vol. 149, no. 6, pp. 4553–4564, 2021.

[13] B. F. Katz and C. D'Alessandro, "Directivity measurements of the singing voice," in *Proceedings of the 19th International Congress on Acoustics*, (Madrid), 2007.

[14] T. W. Leishman, S. D. Bellows, C. M. Pincock, and J. K. Whiting, "High-resolution spherical directivity of live speech from a multiple-capture transfer function method," *J. Acoust. Soc. Am.*, vol. 149, no. 3, pp. 1507–1523, 2021.

[15] S. D. Bellows and B. F. G. Katz, "Combining multiple sparse measurements with reference data regularization to create spherical directivities applied to voice data," *Acta Acustica*, vol. 8, no. 14, pp. 1–12, 2024.

[16] C. Pörschmann and J. M. Arend, "A method for spatial upsampling of voice directivity by directional equalization," *J. Audio Eng. Soc.*, vol. 68, no. 9, pp. 649–663, 2020.

[17] S. Bellows and T. W. Leishman, "Effect of head orientation on speech directivity," in *Proc. Interspeech 2022*, pp. 246–250, 2022.

[18] J. L. Flanagan, "Analog measurements of sound radiation from the mouth," *J. Acoust. Soc. Am.*, vol. 32, no. 12, pp. 1613–1620, 1960.

[19] S. D. Bellows and T. W. Leishman, "On the low-frequency acoustic center," *J. Acoust. Soc. Am.*, vol. 153, no. 6, pp. 3404–3418, 2023.

[20] S. D. Bellows and T. W. Leishman, "Acoustic source centering of musical instrument directivities using acoustical holography," *Proc. Mtgs. Acoust.*, vol. 42, no. 1, p. 055002, 2020.

[21] S. D. Bellows, D. T. Harwood, K. L. Gee, and M. R. Shepherd, "Directional characteristics of two gamelan gongs," *J. Acoust. Soc. Am.*, vol. 154, no. 3, pp. 1921–1931, 2023.

[22] L. Beranek and T. Mellow, *Acoustics: Sound fields, transducers and vibration*. Academic Press, 2 ed., 2019.

[23] S. D. Bellows and T. W. Leishman, "Application of Chebyshev quadrature rules to equiangular spherical and hemispherical directivity measurements," *J. Audio Eng. Soc*, vol. 72, no. 1/2, pp. 44–58, 2024.

[24] S. D. Bellows and T. W. Leishman, "A spherical-harmonic-based framework for spatial sampling considerations of musical instrument and voice directivity measurements," in *Proceedings of Forum Acusticum*, (Turin, Italy), pp. 4747–4754, 2023.