Supplementary material for

A database for the comparison of measured datasets of human voice directivity

Proceedings of the Forum Acusticum 2023

https://doi.org/10.5281/zenodo.7834210

Christoph Pörschmann

Institute of Computer and Communication Technology, TH Köln - University of Applied Sciences

christoph.poerschmann@th-koeln.de

Technology
Arts Sciences
TH Köln

Directivity database

This supplementary material to [1] provides a database and Matlab scripts that allow comparing and visualizing the datasets from different studies. For interpolation and visualization, the scripts require the SUpDEq toolbox[1]. The considered datasets are listed in Table 1. The data has been collected from tables, plots, and datasets from the supplemental material of the respective studies. While some studies present directivity patterns averaged over a sentence or while singing, others report phoneme-dependent data.

Furthermore, the datasets vary in their sampling grids, with many measured in the horizontal plane and just a few measured spherically. Finally, most datasets included in this work present frequency-band averaged values, for example, in one-third octave bands, while a few newer studies provide the raw data in the form of transfer functions. As a result, the datasets are sparse both in their spatial and their frequency resolution.

While most datasets in our database are stored in form of an table in an Excel sheet (`Database_Directivity_Patterns.xlsx`), some of the newer studies provide the datasets of the directivity patterns as supplementary material. The following dataset can be included in the database: Leishman et al. [2] provide the datasets in digital form as supplementary to their paper[2] For this, the file `SpeechDirectivity.CSV` needs to be separately downloaded and put in the Matlab path. All datasets from our studies are stored in the SOFA format [3]. The datasets from [4, 5] contain in total 23 phonemes measured twice for 13 subjects[3]. In [6] we analyzed for the same subjects as above the influence of hand postures on voice directivity[4]. These datasets are stored according to the SOFA convention "GeneralTF". Finally, we presented the results of dummy head measurements with and without face masks [7], these datasets are stored as well as SOFA-Files according to the "GeneralFIR" convention in the supplementary material to this study[5]. For convenience, we have included the datasets of our studies in this supplementary material so that they do not have to be downloaded separately.

Additional dataset

As part of this supplementary material voice directivity datasets determined over a complete sentence which we measured in the same series and for the same subjects as in [4] are included. These datasets have not been published before and comprise one measurement for each of the 13 subjects (*subject#*) containing directivity patterns for the following phonetically balanced German sentence: "Die Schüssel mit Äpfeln haben wir auf den Küchentisch gedeckt." The naming of the measured and post-processed datasets with 32 sampling points is:
Directivity_sentence_{*subject#*}.sofa
Additionally, datasets that have been spatially upsampled to a dense Lebedev grid with 2702 sampling points using the SUpDEq method [8] are filed:
Directivity_sentence_{*subject#*}_upsampled.sofa
As in [9] the transfer functions for each direction of the directivity pattern are stored in the data field *Data* separately for the real and imaginary parts. The data field *ReceiverPosition* contains azimuth ϕ and elevation θ for each direction and in the third field, the subject's optimal head radius is given, which we determined according to Algazi et al. [10] based on measurements of each subject's head width, height, and length.

[1] Available: https://github.com/AudioGroupCologne/SUpDEq
[2] Available: https://scholarsarchive.byu.edu/directivity/1
[3] Available: https://doi.org/10.5281/zenodo.7452117
[4] Available: https://doi.org/10.5281/zenodo.5995215
[5] Available: https://doi.org/10.5281/zenodo.3952320

Visualizations

Our database allows direct comparison and visualization of datasets from the considered studies. For this, methods for interpolating the frequency scale and upsampling the sampling grid have been implemented in the Matlab scripts. The interpolation over frequency for the frequency-band-based datasets is carried out as a linear interpolation of the magnitudes. For the spatial upsampling, we implemented methods from the SUpDEq toolbox that can be applied to both the frequency-band and transfer-function-based datasets allowing for spatial upsampling to an arbitrary target grid. For the datasets for which only the magnitudes in respective frequency bands are given, we spatially upsampled the datasets using natural neighbor interpolation [27]. For the complex-valued datasets that provide transfer functions, we applied the SUpDEq (Spatial upsampling by Directional Equalization) method [8] which has been evaluated in [28] for voice directivities. The visualization comprises three different kinds of plots, polar plots in the horizontal and/or vertical plane, plots on a continuous frequency scale in the horizontal and/or vertical plane as well as balloon plots.

Matlab scripts

As mentioned above the Matlab scripts require the SUpDEq toolbox[1]. The Matlab script `runme.m` defines which dataset is to be analyzed and some properties of the plots, e.g. whether an averaged plot over the directivity patterns of all subjects is to be created or whether a single data set is to be considered (only available for some of the datasets).

The Matlab scripts `Voice_Directivity_Tables.m`, `Voice_Directivity_Leishman.m`, `Voice_Directivity_Poerschmann.m` read the directivity patterns from the database and store the datasets in the form of a struct. Databases from external sources must be downloaded and placed in the appropriate folders, or the Matlab path must be adjusted. If desired, the frequency scale can be interpolated. The struct contains information about the type of sampling grid, the type of measurement, the measurement distance, and the directivity data itself (magnitudes of transfer function values, frequency values, sampling grid).

The Matlab script `Plot_Voice_Directivity_FA_2023.m` generates three different types of plots. Polar plots show the directivity patterns in defined frequency bands for the horizontal and/or vertical planes. The function further creates plots in the horizontal and/or vertical planes on a continuous frequency axis, showing the structure of the directivity patterns over frequency and angle in a single plot. Finally, for an analysis of the spherical shape of the directivity, balloon plots for different frequencies are drawn. For each of the plots, the datasets are spatially interpolated using the SUpDEq toolbox.

[6]The digital dataset of [25] has not been used for our database due to technical reasons.

Authors	Year	No. of subj.	Dummy head model	Articulation	Data repres.	Sampl. grid	Dist.	Meas. proc.	Data acq.	Remarks	Data-base
Dunn and Farnsworth [11]	1939	1		fluent speech	1/2 Octave	spherical, 45° hor., 22.5° vert.	0.05 m - 1 m	seq.	table	one hemisphere, mouth centering, $\leq$ 500 Hz 1/1 octave	○
Flanagan [12]	1960		mannequin		1/1 Octave	20° hor.	0.3 m	seq.	plot	mouth centering	○
Moreno and Pfretzschner [13]	1978	10		fluent speech	1/3 Octave	hor. and vert. planes		seq.	plot		○
Marshall and Meyer [14]	1985	1		singing	1/1 Octave				plot	[a] [e] [o] only shown for 2 kHz, not considered	x
Huopaniemi et al.[15]	1999		B&K 4128		1/1 Octave		2 m	seq.	plot	2 measurement methods (reciprocal and direct)	x
Chu and Warnock[16]	2002	40	B&K	fluent speech	1/3 Octave	spherical, 15° hor., 9 elevations	1 m	seq.	table	one hemisphere, standard deviations of all data given in paper	○
Bozzoli et al. [17, 18]	2003, 2005	10	B&K 4230, Parma		1/1 Octave	5° hor.				two measurement series, 1 and 10 subjects 3 elevations measured in one series	○
Katz et al. [19]	2006	1		[a] [o] [i] [m] [n] [f] [s] [ch]	1/1 Octave	5° hor.		simult.	plot	one hemisphere, not all bands shown in plots	x
Monson et al. [20]	2012			fluent speech, [s] [sh] [f] [th], singing	1/1 Octave	15° hor.	0.6 m	simult.	table	one hemisphere	○
Kocon and Monson[21]	2018	15		fluent speech, [a [e] [i] [o] [u]	1/3 Octave	15° hor.	0.6 m	simult.	table	one hemisphere	○
Brandner et al. [22]	2018	2		singing	frequency response	11.25° hor., 11.25° vert.	1.23 m	simult.	plot	transferred from plot showing 2 frequency bins, 22.5° steps in hor. and vert. plane	x
Fischer et al. [23]	2019		Kemar 45BC		frequency response	4° hor.	0.6 m	seq.	plot	transferred from plot showing 4 frequency bins	x
Brandner et al. [24]	2020	2	B&K 4128		frequency response	11.25° hor., 11.25° vert.	1.23 m	simult.	plot	transferred from plot showing 2 frequency bins, 22.5° steps in hor. and vert. plane	x
Pörschmann et al.[7]	2020		HEAD HMS II.3		frequency response	spherical, 2702 Lebedev grid	2 m	seq.	dataset	5 face masks and reference without mask	○
Leishman et al.[2]	2021	6	Kemar 45BC	fluent speech	frequency response	spherical, 5° hor., 5° vert.	1.83 m	vert. simult., hor. seq.	dataset	multiple-capture transfer function technique	○
Pörschmann and Arend [4, 5]	2021, 2023	13		5 vowels, 6 plosives, 3 nasals, 7 fricatices, 2 alveolars	frequency response	spherical, 32 directions	1 m	simult.	dataset	spatial Upsampling using SUpDEq [8] to 2702 Lebedev grid	○
Pörschmann and Arend [6]	2022	13		articulation of an [a] with and without hand postures	frequency response	spherical, 32 directions	1 m	simult.	dataset	spatial Upsampling using SUpDEq [8] to 2702 Lebedev grid	○
Brandner et al. [25]	2022	10		[a] [e] [i] [o] [u]	frequency response	11.25° hor., 11.25° vert.	1.23 m	simult.	dataset	sustained vowels articulated over wide pitch range	x[6]
Pörschmann [26]	2023	13		fluent speech	continuous	spherical, 32 directions	1 m	simult.	dataset	spatial Upsampling using SUpDEq [8] to 2702 Lebedev grid	○

Table 1: Overview of all studies on voice directivity that are considered in the database. The last column indicates in which form the data sets could be made usable in the database. An [x] denotes single data sets suitable for comparative purposes, and a [○] denotes full data sets determined in multiple frequency bands and on a well-defined sampling pattern.

Visualizations

The following plots were obtained by applying the database and the Matab scripts. We show polar plots that depict the data in frequency bands in the horizontal and vertical planes. Furthermore, we depict visualizations of the datasets in the horizontal and vertical planes on a continuous frequency axis. Finally, as an example of spherical plots, we show some balloon plots.

Polar plots

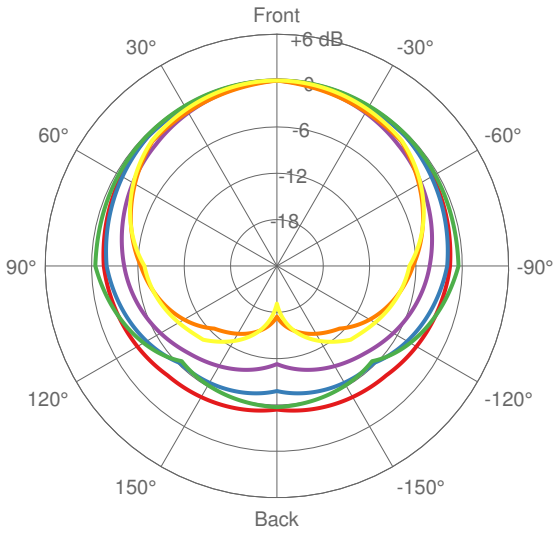
We show the directivity patterns in the horizontal and vertical planes. Figures 1 – 2 depict the directivity patterns in octave bands with center frequencies between 500 Hz and 8 kHz in the horizontal and vertical planes for the datasets from Table 1 that provide data for fluent speech. As the data measured by Moreno and Pfretzschner [13] is limited to 4 kHz, the 8 kHz octave band is not shown in these plots.

Frequency plots

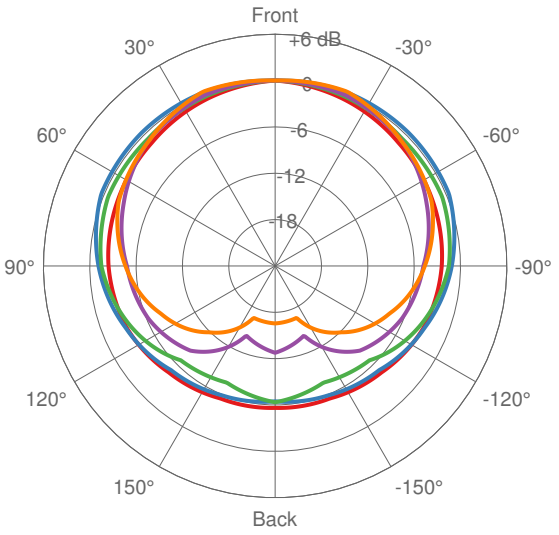
Figures 3 – 4 depict the directivity patterns in the horizontal and vertical plane for the datasets from Table 1 that provide data for fluent speech on a continuous frequency axis. The datasets apart from [1] were interpolated from the values given in the respective frequency bands to a continuous frequency scale.

Balloon plots

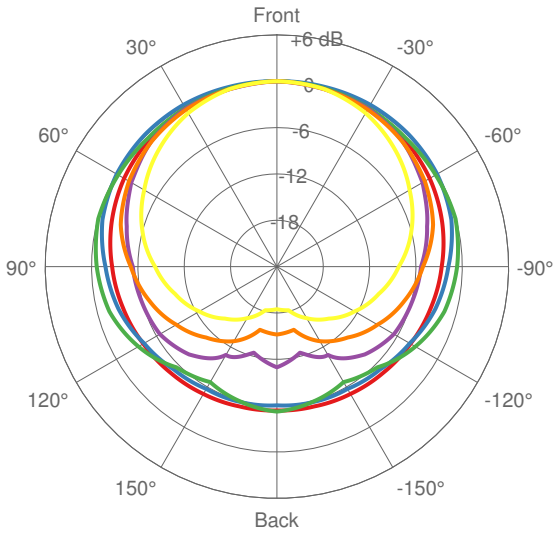
Figure 5 – 8 visualize the directivity for the datasets providing spherical measurements of fluent speech from Table 1. However, the expressiveness of the graphs in Figures is limited because the graphs cannot be rotated directly as in the original Matlab plots.



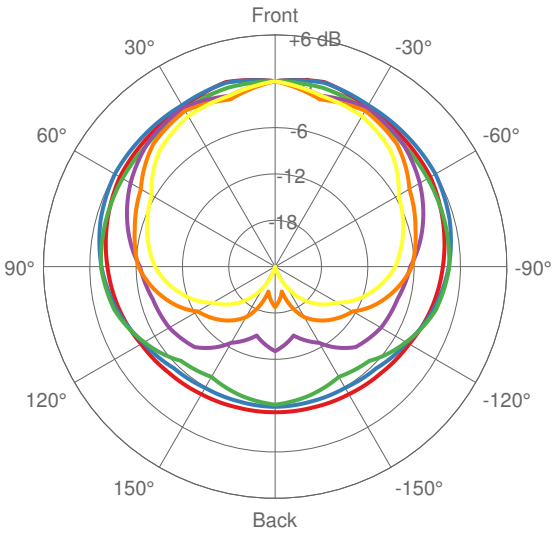
Dunn&Farnsworth [11]



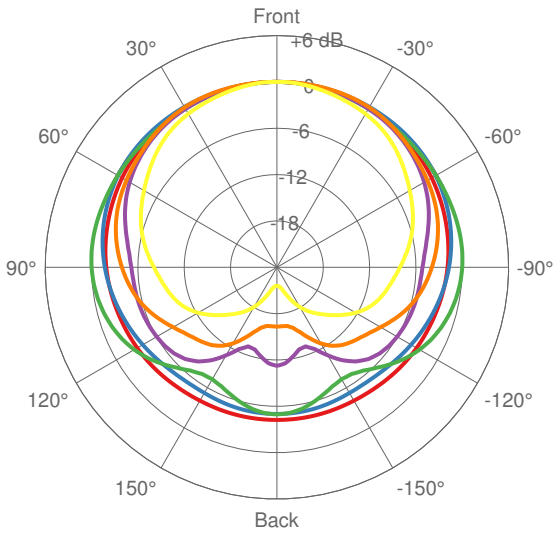
Moreno&Pfretzschner [13]



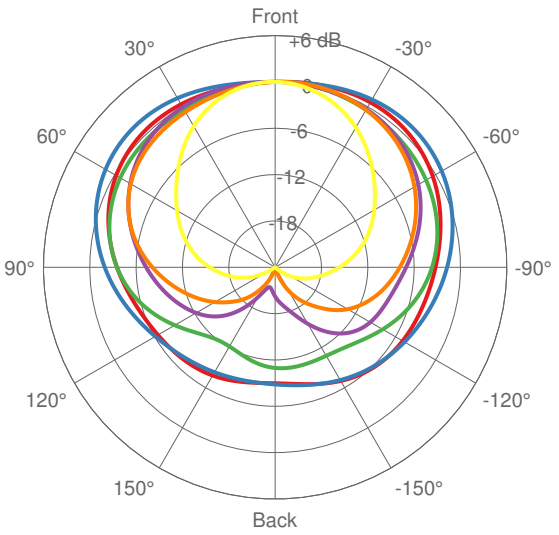
Chu&Warnock [16]



Kocon&Monson [21]



Leishman et al. [2]



Pörschmann [1]

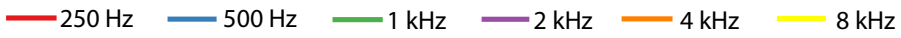


Figure 1: Polar plots in the horizontal plane for the measurements that provide directivity patterns of fluent speech. Shown are the mean values in octave bands with center frequencies between 250 Hz and 8 kHz ([13] only up to 4 kHz).

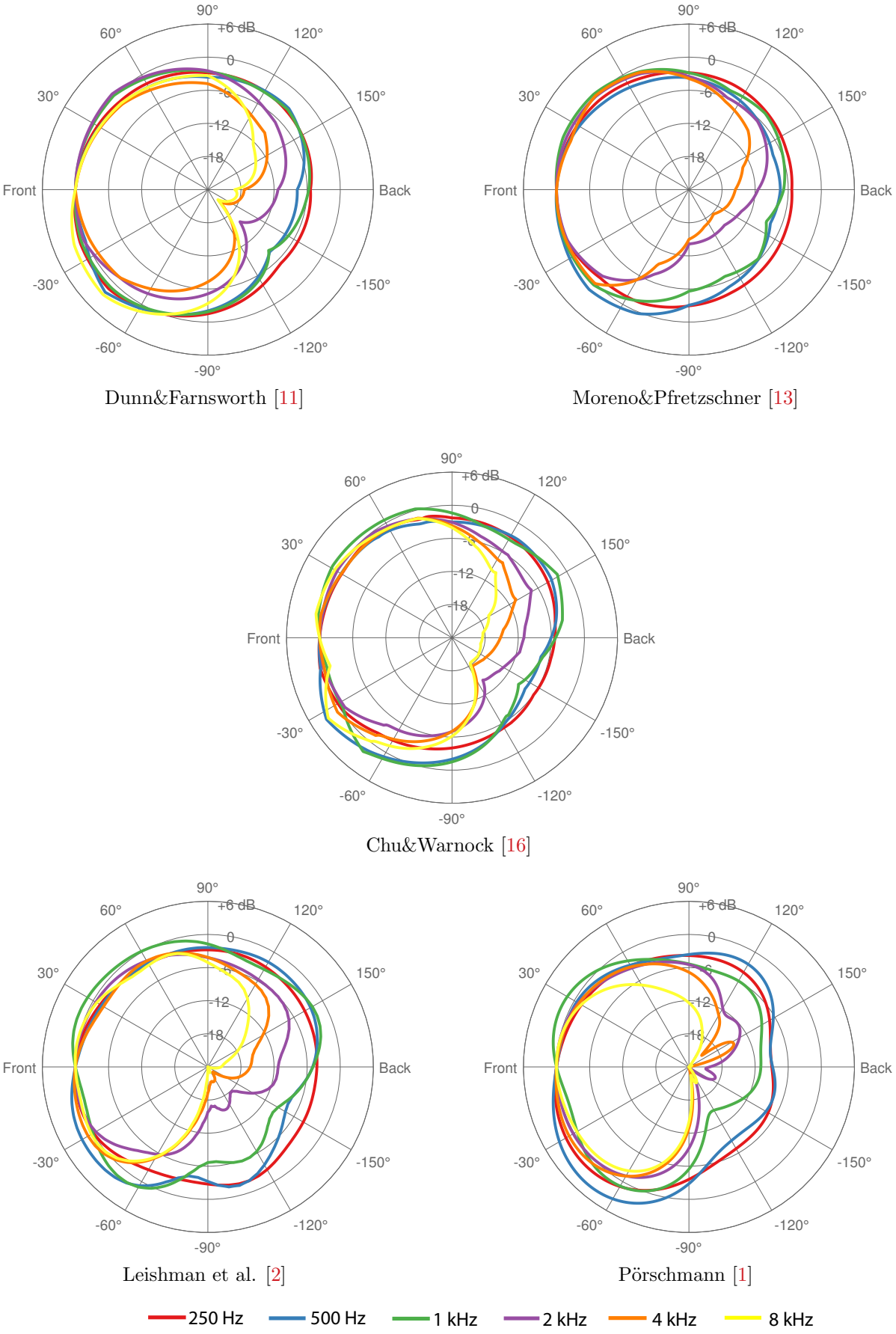


Figure 2: Polar plots in the vertical plane for the measurements that provide directivity patterns of fluent speech. Shown are the mean values in octave bands with center frequencies between 250 Hz and 8 kHz ([13] only up to 4 kHz).

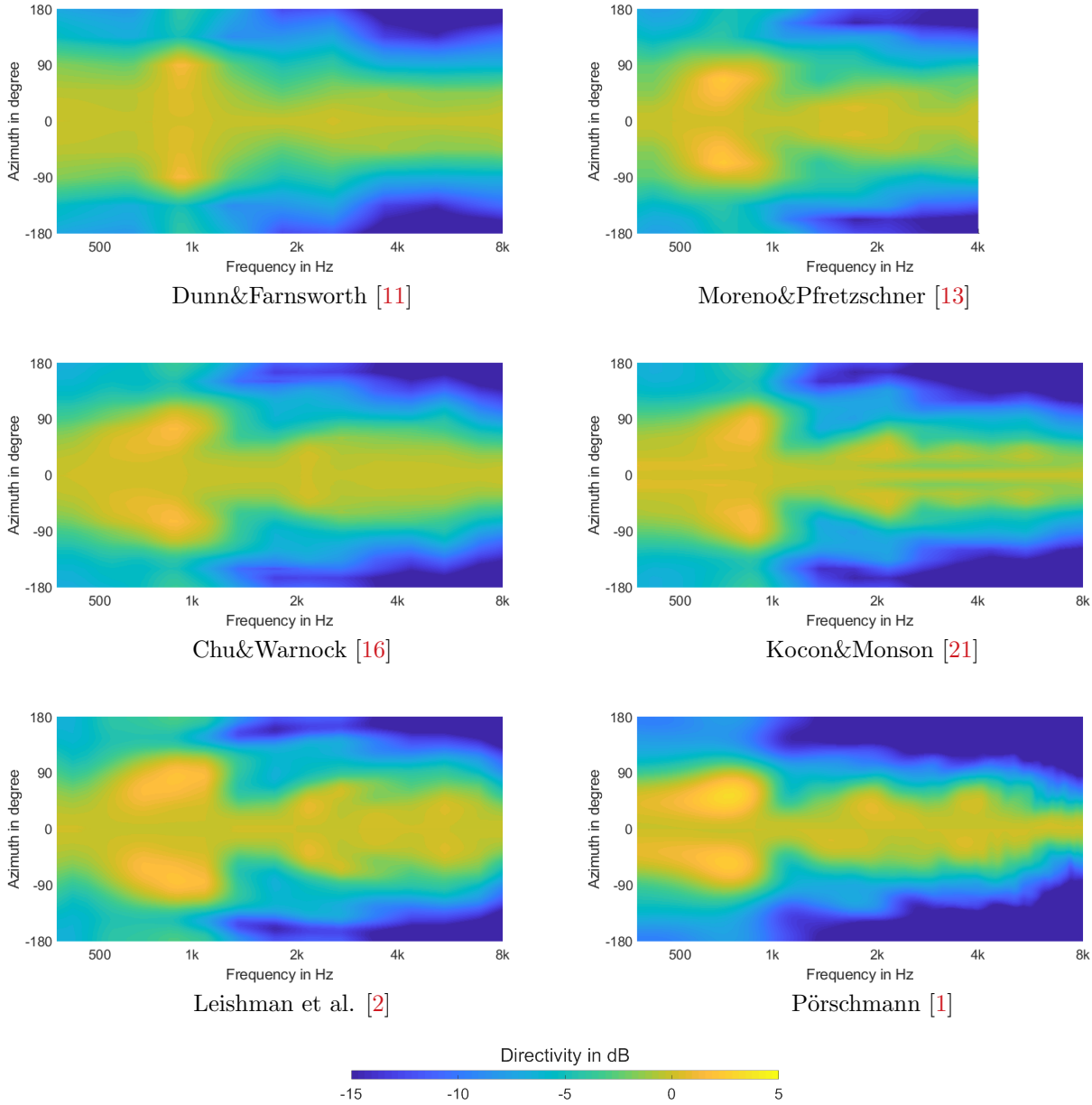


Figure 3: Directivity patterns over frequency in the horizontal plane for the measurements for which directivity patterns of fluent speech have been determined up to 8 kHz ([13] only up to 4 kHz).

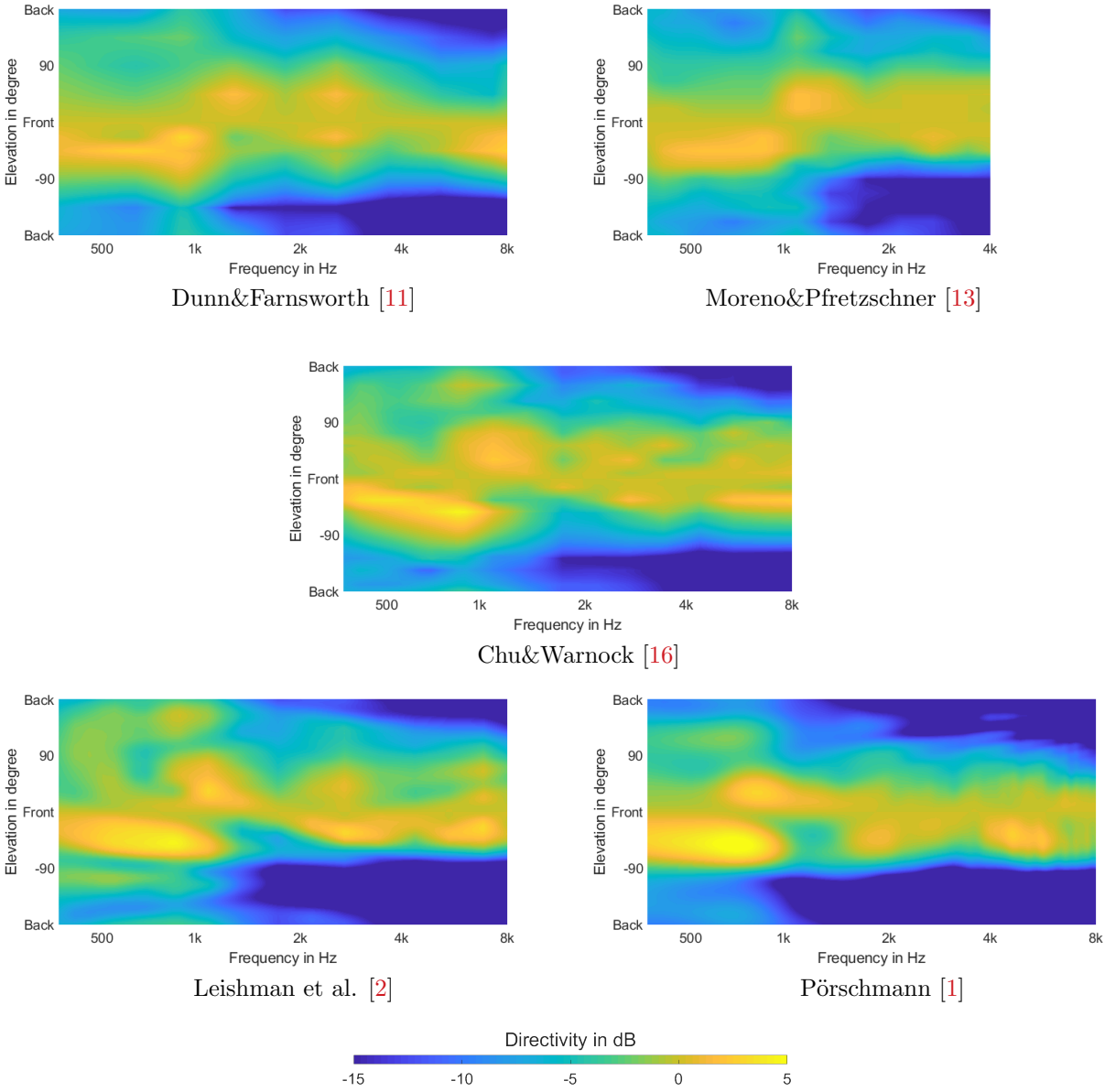


Figure 4: Directivity patterns over frequency in the vertical plane for the measurements for which directivity patterns of fluent speech have been determined up to 8 kHz ([13] only up to 4 kHz).

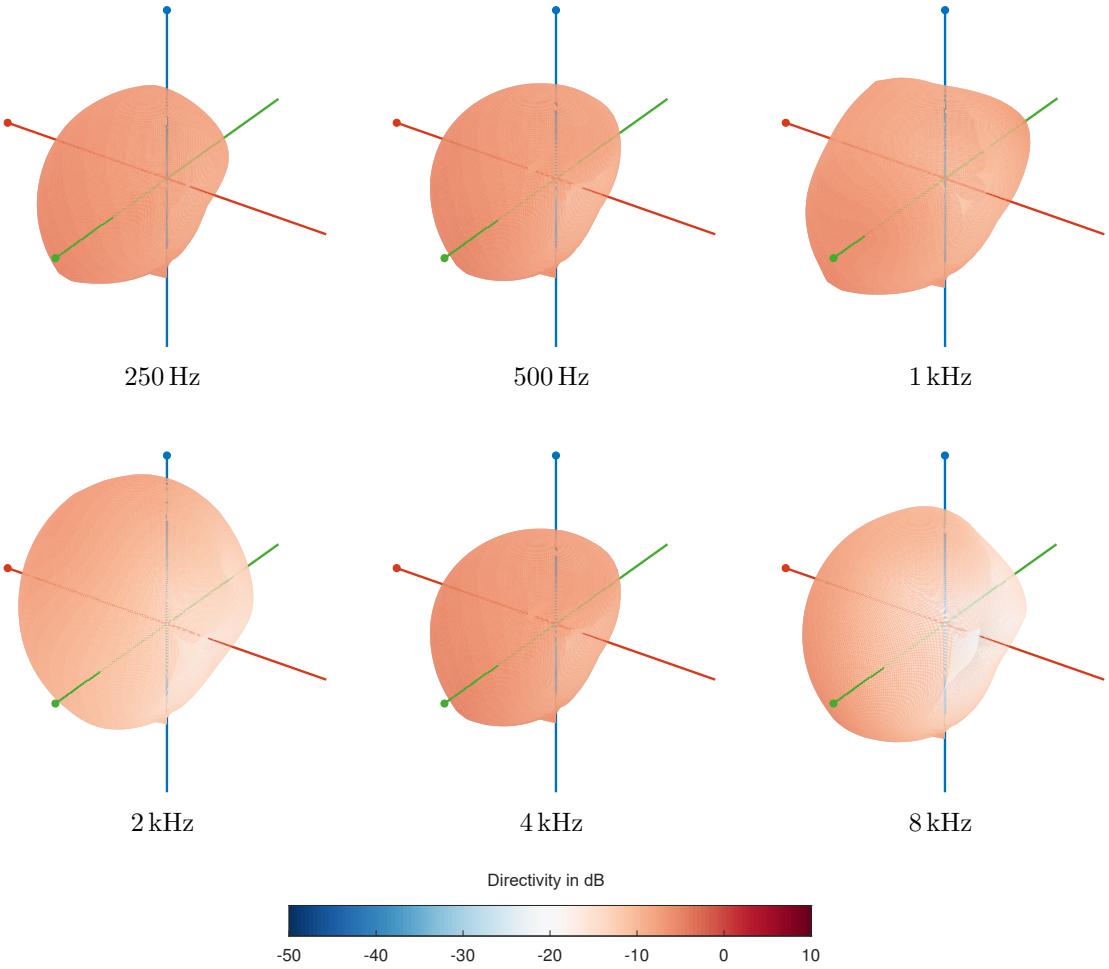


Figure 5: Balloon plots viewed from slightly backward and slightly upward (ϕ =215°, θ=30°) for the measurements from Dunn and Farnsworth [11]. Frequencies of 250 Hz, 500 Hz, 1 kHz, 2 kHz, 4 kHz, and 8 kHz are shown.

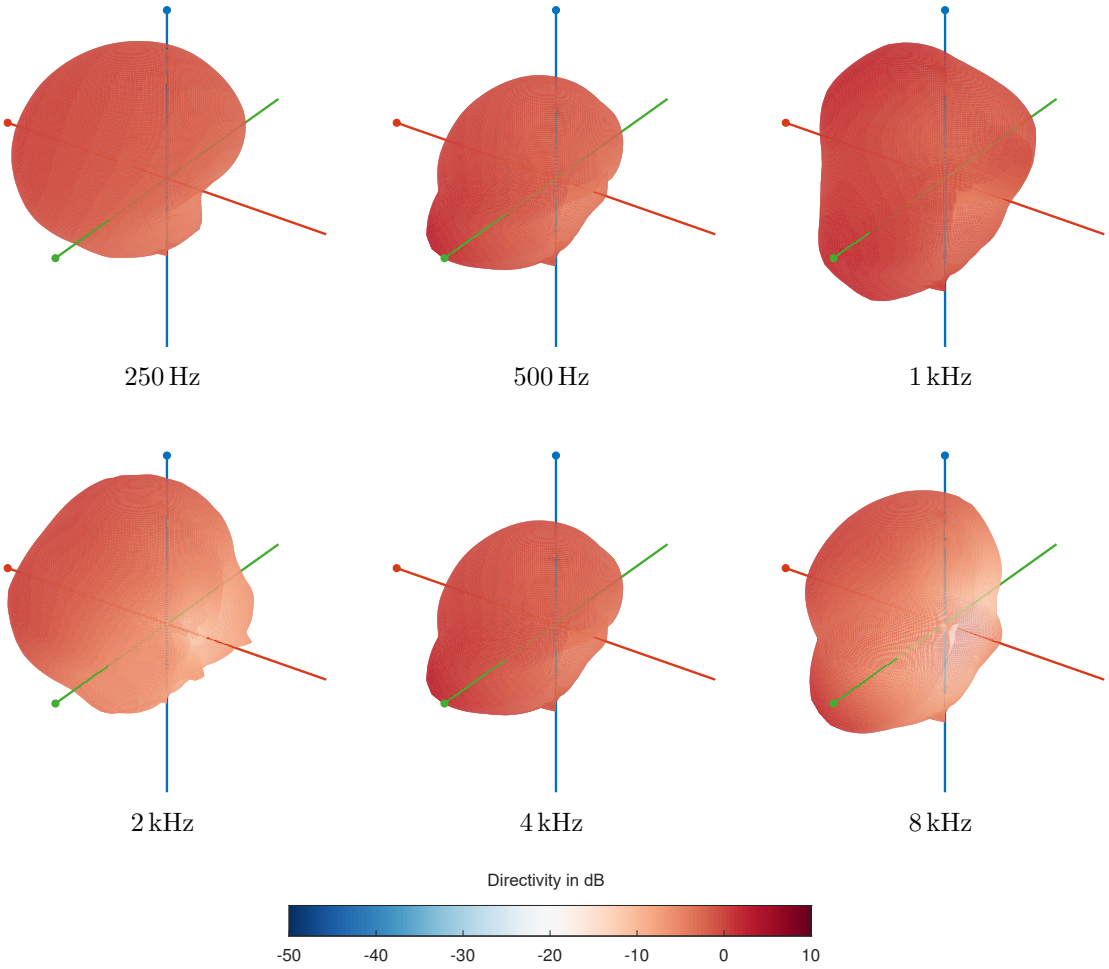


Figure 6: Balloon plots viewed from slightly backward and slightly upward (ϕ =215°, θ=30°) for the measurements from Chu and Warnock [16]. Frequencies of 250 Hz, 500 Hz, 1 kHz, 2 kHz, 4 kHz, and 8 kHz are shown.

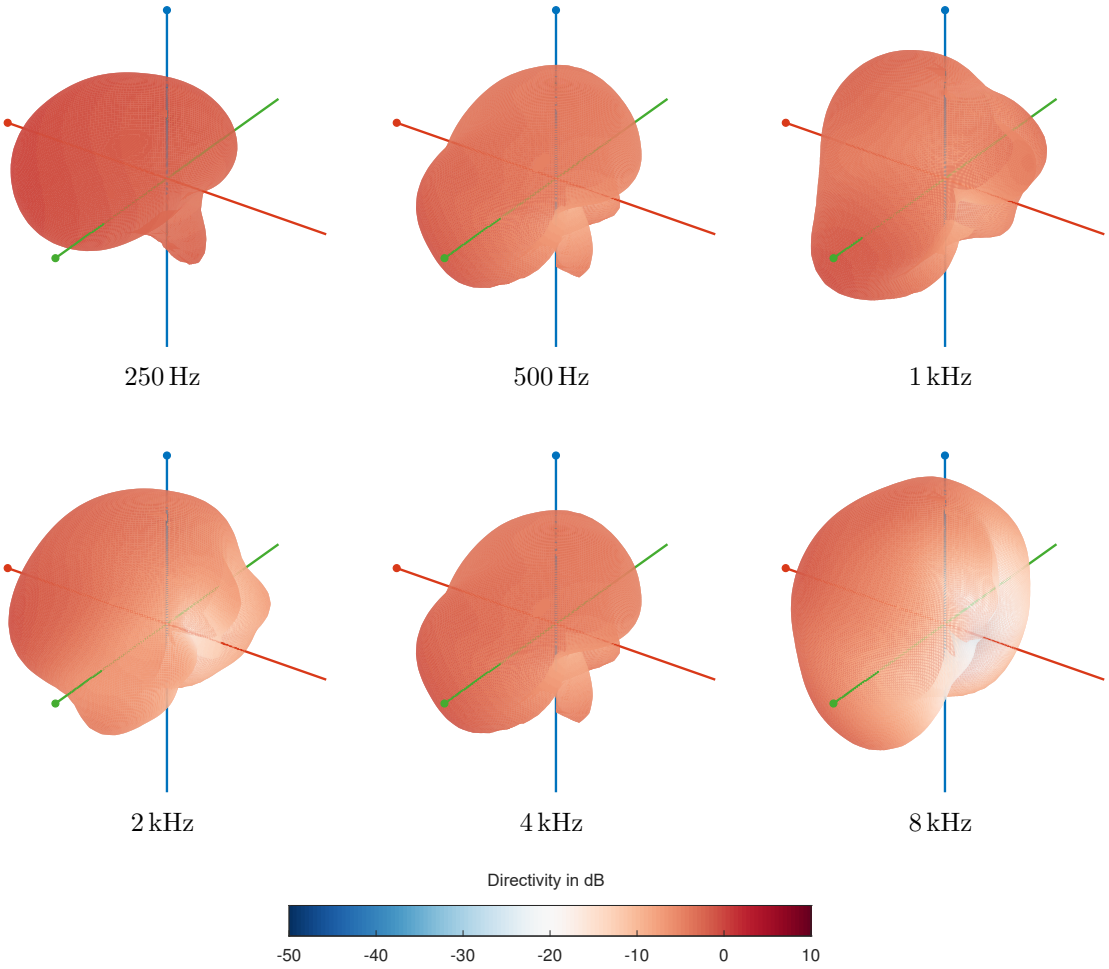


Figure 7: Balloon plots viewed from slightly backward and slightly upward (ϕ =215°, θ=30°) for the measurements from from Leishman et al. [2]. Frequencies of 250 Hz, 500 Hz, 1 kHz, 2 kHz, 4 kHz, and 8 kHz are shown.

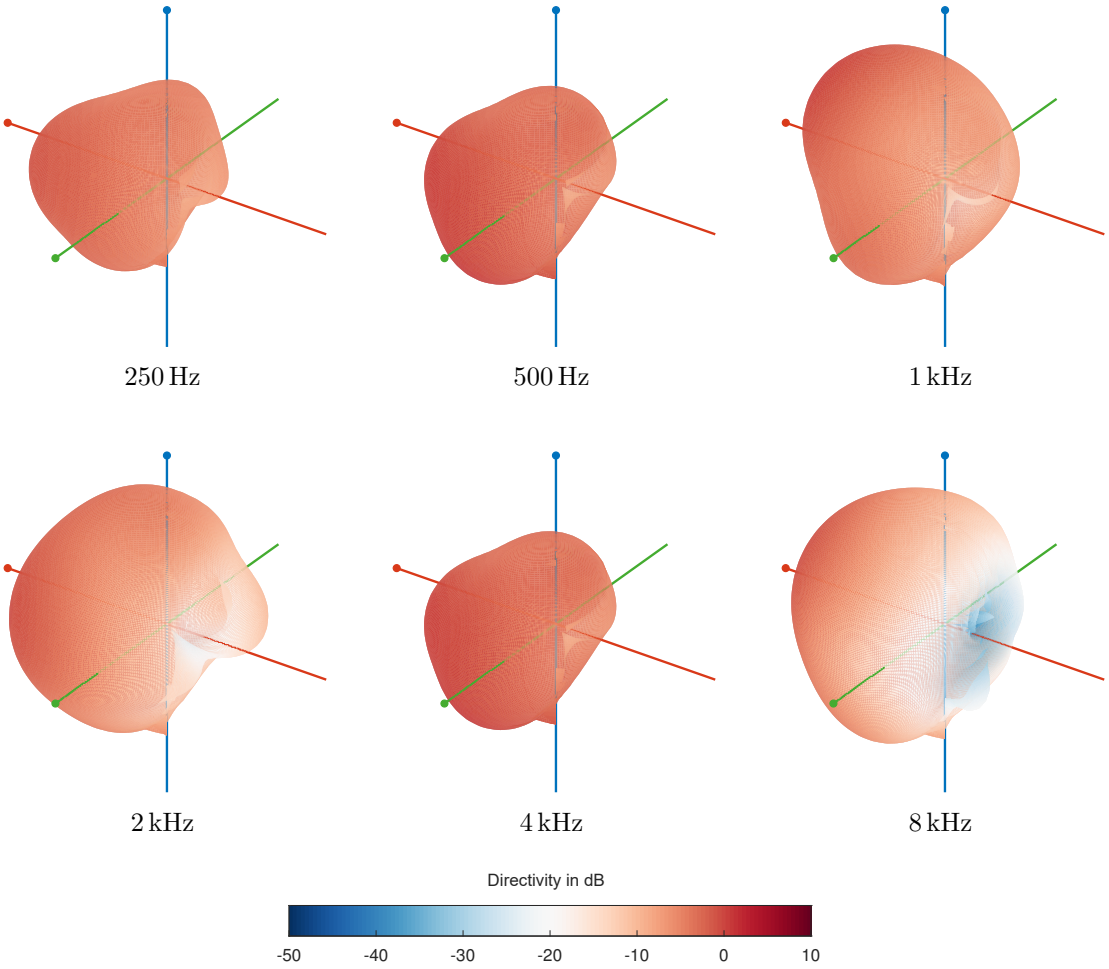


Figure 8: Balloon plots viewed from slightly backward and slightly upward (ϕ =215°, θ=30°) for the measurements from Pörschmann [1]. Frequencies of 250 Hz, 500 Hz, 1 kHz, 2 kHz, 4 kHz, and 8 kHz are shown.

References

[1] C. Pörschmann, "A database for the comparison of measured datasets of human voice directivity," in *Proc. Forum Acusticum*, 2023.

[2] 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," *The Journal of the Acoustical Society of America*, vol. 149, no. 3, pp. 1507–1523, 2021.

[3] P. Majdak, F. Zotter, F. Brinkmann, J. De Muynke, M. Mihocic, and M. Noisternig, "Spatially Oriented Format for Acoustics 2.1: Introduction and Recent Advances," *Journal of the Audio Engineering Society*, vol. 70, no. 7/8, pp. 565–584, Jul. 2022.

[4] C. Pörschmann and J. M. Arend, "Investigating phoneme-dependencies of spherical voice directivity patterns," *The Journal of the Acoustical Society of America*, vol. 149, no. 6, pp. 4553 – 4564, 2021.

[5] ——, "Investigating phoneme-dependencies of spherical voice directivity patterns II: Various groups of phonemes," *The Journal of the Acoustical Society of America*, vol. 153, no. 1, pp. 179–190, 2023, publisher: Acoustical Society of America. [Online]. Available: https://asa.scitation.org/doi/10.1121/10.0016821

[6] ——, "Effects of hand postures on voice directivity," *JASA Express Letters*, vol. 2, no. 3, p. 035203, 2022, publisher: Acoustical Society of America.

[7] C. Pörschmann, Tim Lübeck, and J. M. Arend, "Impact of face masks on voice radiation," *The Journal of the Acoustical Society of America*, vol. 148, no. 6, pp. 3663–3670, 2020.

[8] C. Pörschmann, J. M. Arend, and F. Brinkmann, "Directional Equalization of Sparse Head-Related Transfer Function Sets for Spatial Upsampling," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 6, pp. 1060 – 1071, 2019.

[9] C. Pörschmann and J. M. Arend, "Supplementary material for Investigating phoneme-dependencies of spherical voice directivity patterns II: Various groups of phonemes," Apr. 2023, doi: 10.5281/zenodo.7452117.

[10] V. R. Algazi, C. Avendano, and R. O. Duda, "Estimation of a Spherical-Head Model from Anthropometry," *Journal of the Audio Engineering Society*, vol. 49, no. 6, pp. 472 – 479, 2001.

[11] 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.

[12] J. L. Flanagan, "Analog Measurements of Sound Radiation from the Mouth," *The Journal of the Acoustical Society of America*, vol. 32, no. 12, pp. 1613–1620, 1960.

[13] A. Moreno and J. Pfretzschner, "Human Head Directivity in Speech Emission: A new approach," *Acoustics Letters*, vol. 1, pp. 78–84, 1978.

[14] A. H. Marshall and J. Meyer, "The directivity and auditory impressions of singers," *Acustica*, vol. 58, pp. 130–140, 1985.

[15] J. Huopaniemi, K. Kettunen, and J. Rahkonen, "Measurement and Modeling Techniques for Directional Sound Radiation from the Mouth," in *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 1999, pp. 183–186.

[16] W. T. Chu and A. C. C. Warnock, "Detailed Directivity of Sound Fields Around Human Talkers," National Research Council of Canada, Ottawa, Tech. Rep. NRC-IRC-15212, 2002.

[17] F. Bozzoli and A. Farina, "Directivity balloons of real and artificial mouth simulators for measurement of the Speech Transmission Index," in *Proceedings of the 115th AES Convention, New York, Preprint #5953*, 2003.

[18] F. Bozzoli, A. Farina, and M. Viktorovitch, "Balloons of Directivity of Real and Artificial Mouth Used in Determining Speech Transmission Index," in *Proceedings of the 118th AES Convention, Barcelona, Preprint #6492*, 2005.

[19] B. F. G. Katz, F. Prezat, and C. d'Alessandro, "Human voice phoneme directivity pattern measurements," *The Journal of the Acoustical Society of America*, vol. 120, no. 5, pp. 3359–3359, 2006, publication Title: 4th Joint Meeting Acoustical Society of America and Acoustical Society of Japan.

[20] 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.

[21] 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.

[22] M. Brandner, M. Frank, and D. Rudrich, "DirPat - Database and Viewer of 2D/3D Directivity Patterns of Sound Sources and Receivers," in *Proceedings of the 144th AES Convention, e-Brief 425*, 2018, pp. 1–5.

[23] G. Fischer, C. Schneiderwind, and A. Neidhardt, "Comparing the directivity of a mouth simulator and a simple physical model," in *Proceedings of the 45th DAGA*, 2019.

[24] M. Brandner, R. Blandin, M. Frank, and A. Sontacchi, "A pilot study on the influence of mouth configuration and torso on singing voice directivity," *The Journal of the Acoustical Society of America*, vol. 148, no. 3, pp. 1169 – 1180, 2020.

[25] M. Brandner, Matthias Frank, and A. Sontacchi, "Horizontal and Vertical Voice Directivity Characteristics of Sung Vowels in Classical Singing," *Acoustics*, vol. 4, pp. 849–866, 2022.

[26] C. Pörschmann, "Supplementary Material for A database for the comparison of measured datasets of human voice directivity," 2023, doi: 10.5281/zenodo.7834210. [Online]. Available: https://doi.org/10.5281/zenodo.7834210

[27] R. Sibson, "A brief description of natural neighbor interpolation," in *Interpreting Multivariate Data*, V. Barnett, Ed. John Wiley, 1981, pp. 21–36.

[28] C. Pörschmann and J. M. Arend, "A Method for Spatial Upsampling of Voice Directivity by Directional Equalization," *Journal of the Audio Engineering Society*, vol. 68, no. 9, pp. 649–663, 2020.