PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Influence of the vocal tract on voice directivity

Rémi Blandin[(1)], Manuel Brandner[(2)]

[(1)]Institute of Acoustics and Speech Communication, TU Dresden, 01062 Dresden, Germany, remi.blandin@tu-dresden.de
[(2)]University of Music and Performing Arts, Institute of Electronic Music and Acoustics, Graz, brandner@iem.at

Abstract
Voice directivity induces variations of the sound amplitude and frequency content with the direction. Voice directivity is important for the efficient transmission of speech and singing. Therefore, it is taken into account by concert hall designers in order to enhance the quality of artistic performances. While in voice directivity studies the shape of the head and the torso have long been considered to study this phenomenon, the influence of the vocal tract shape/configuration has received very little attention. However, it has been recently shown that the vocal tract configuration influences voice directivity through the internal acoustic field and the frequency content. Within this paper, the contribution of the vocal tract to the voice directivity is characterized through physical modelling and direct measurements on a professional classical singer. In particular, the role of higher order propagation modes is discussed as well as changes of voice directivity due to different vocal tract configurations.
Keywords: Voice, directivity, singing, higher order modes

1 INTRODUCTION

The directivity of speech and singing induces variations of the amplitude and the frequency content of the radiated sound with the direction. Various measurements performed with microphone arrays on speakers and singers show that this phenomenon increases toward high frequencies [1, 2, 3, 5].

It was also shown that the directivity patterns of the various vowels and consonants are different [2, 4, 5, 6]. Also, analyzing data within smaller frequency intervals, such as third of octave bands and even linear frequency discretization [6] highlighted more complex changes of the directivity patterns with the frequency. Moreover, it was shown that the internal acoustic field can potentially influence the directivity of the radiated sound above 4 kHz to 5 kHz, inducing significant changes of directivity patterns within frequency intervals of the order of 100 Hz [7].

Katz and d'Alessandro highlighted differences in directivity patterns measured with different vocal techniques [4]. A global increase of directivity observed with the projected singing technique was attributed to an increase of acoustic energy in the 2.5 kHz third of octave band. However, to our knowledge, no study has focused on the potential influence of the mouth configuration on the directivity. Little speech studies compared measurements with theoretical radiation models, and only with simple models describing the mouth as a vibrating piston.

Thus, the objective of this work is to investigate the influence of the size of the mouth opening on voice directivity by comparing measurements and simulation. The directivity of a professional singer singing the vowel /a/ with two different mouth configurations was recorded with a microphone array. These measurements are compared with two theoretical models:

- a simple vibrating plane piston model set inside an infinite baffle [8],
- and the multimodal theory which allows one to take into account the potential influence of the vocal tract as well as the radiation patterns of the higher order modes [16].

The diffraction by the head and the torso is currently not taken into account in order to focus only on the potential effects of the mouth shape, whereas stronger influences are expected to occur only up to 2 kHz to 2.5 kHz for an average person (head and torso dimensions).

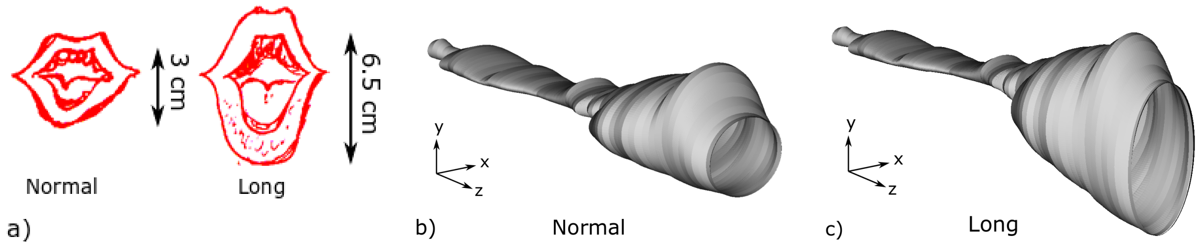

Figure 1. a) Normal and long mouth configurations used by a singer with vertical lower lip to upper teeth distance. b) and c), vocal tract geometries from magnetic resonance image [14] modified to simulate the b) normal and c) long mouth configurations, respectively. The lower part of the oral cavity is rotated to match the lower lip to upper teeth distance measured on the singer.

2 METHOD

2.1 Directivity measurement

The sound radiated by a professional female singer was recorded with a double circle microphone array constituted of 62 microphones regularly distributed on two circles in the horizontal and vertical planes [9]. The singer was asked to sing the vowel /a/ doing a glissando with a normal mouth opening and a long mouth opening. The long mouth opening corresponds to a lower position of the jaw which increases the vertical mouth opening compared to the normal configuration (see Fig 1a). Lowering the jaw is a common practice in classical singing [10]. The mouth opening was characterized with the vertical lower lip to upper teeth distance. This represents the effective mouth opening since the upper lip is above the upper teeth and the lower lip covers the lower teeth. The normal mouth opening corresponds to a vertical distance of about 3 cm and the long one to about 6 cm (see Fig. 1a). The transfer functions between each point of the microphone array and a reference microphone placed as close as possible (< 3 cm) to the singer were computed using a glissando method [11, 12].

2.2 Simulations

An initial geometry obtained by simplifying a three-dimensional geometry extracted from magnetic resonance image [13] was transformed to approximate the different vocal tract configurations investigated. This initial geometry corresponds to the vowel /a/ pronounced by a male speaker. It is constituted of a succession of straight tubes with elliptical cross-sections [14].

Since the singer involved in this study was a female, the length and the volume of the oral and pharyngeal cavity of the original geometry were adapted to female dimensions. The ratios of the average dimensions of 120 male and female subjects [15] were used to adapt the length and cross-sectional area of the sections to female average dimensions.

The initial geometry is axis-symmetric: the centers of each section share a common axis. However, it has been shown that this configuration is not well suited to reproduce the acoustic properties of the vocal tract above 4-5 kHz [16, 7]. Thus, 25 % and 75 % of eccentricity have been introduced in the horizontal (x, z) and the vertical (y, z) plane, respectively. The centers of the cross-sections have been shifted toward the positive values of x and y by 25 % and 75 % of the half width and the half height of the ellipses, respectively.

In order to adapt the mouth opening to the vertical lower lip to upper teeth distance measured on the test subject, a downward rotation of the lower points of the cross-sections of the oral cavity has been operated. The center of rotation was placed at the limit between the pharyngeal and the oral cavities. The height of the ellipses was increased to match the position of the rotated points (see Figs. 1b and 1c).

The transfer functions between an input volume velocity imposed at the glottis and the acoustic pressure radiated by the mouth was computed with the multimodal method described by Blandin et al. [16]. In order to simulate the vibrating piston model, the same method was used with the plane mode only.

2.3 Directivity maps

The transfer functions computed from the measurements and the simulations are presented as directivity maps in Fig. 2. These maps present the radiated amplitude in color-scale as a function of the angular position and the frequency. Frequency resolutions of 43.1 Hz and 10 Hz are used for the analysis of the measurements and simulations, respectively. The amplitudes are normalized over all angular positions at each frequency by the maximum of amplitude over the different angular positions. Since the simulations use and infinite baffle boundary condition, it can provide data only between -90°and 90°(0°corresponds to the direction normal to the frontal plane). Therefore, the visualization of the measurements will also be limited to the same angular region.

3 RESULTS

Within this section the results of the measurements and the simulations of the two proposed methods are presented.

All the directivity maps presented in Fig. 2 show a beam which becomes narrower as the frequency increases for the measurement and both simulation types. The directivity maps of the measurements and both simulations in Fig.2 show an increasing main beam width as the frequency increases. This main beam is globally narrower for the long configuration (Figs. 2d to 2f).

The measurements (Figs. 2a and 2d) and the simulation performed with higher order modes (Figs. 2b and 2e) show more complex patterns than the simulation obtained with the vibrating piston model (Figs. 2c and 2f). Abrupt changes of the direction (abrupt upward or downward shifts of the beam) and the shape of the directivity patterns as well as localized minima occur in small intervals of the order of 100 Hz. Similar trends can be seen both in the measurement and the multimodal simulations. In particular, in the normal configuration, a localized minimum around 9.5 kHz and -30°, exists both in the measurements and the multimodal simulation (Figs. 1a and 1b). In the long configuration, similar trends can be seen around 3.5 kHz or between 5-6 kHz in the 45°to 90°angular region. However, these complex variations appear only at relatively high frequency in the multimodal simulations: from 5 kHz for the normal configuration (see Fig. 2b), and from 3-4 kHz for the long configuration (see Fig. 2e). The measurements show more complex radiation patterns than the simulations between 0 and 3-4 kHz.

Between 0 and 2 kHz the measurements performed in both mouth configurations are very similar. They have the same radiation patterns, including minima and maxima evolving downward with increasing frequency. These patterns are not seen on the simulations which show little amplitude variations in this frequency range. Small deviations between the measurements performed with the normal and the long configurations occur between 2-3 kHz within the 0°to -90°angular region. More substantial differences are visible from 3kHz on.

In the measurements, from 2 kHz on, globally more energy is radiated downward (between 0°and -90°) in the long configuration compared to the normal one. This is especially seen in the frequency region around 4 kHz. However, with increasing frequency the main beam direction fluctuates between being radiated upward and downward. This trend is seen only for the experimental data and the simulations performed with higher order modes in some frequency bands. Thus, a downward orientation of the beam is present between 3-4 kHz, 5-6 kHz, 7-8 kHz and 9-10 kHz for the multimodal simulations of the long configuration (see Fig. 2e). However, the beam also tend to be orientated upward between 4-5 kHz, 6-7 kHz and 8-9 kHz.

4 DISCUSSION

Within this section the results obtained are discussed. Not only the proposed models are compared with each other and the measurements, but also the differences that occur if the jaw is lowered and the distance between lower lip and upper incisors increases by more than double the height.

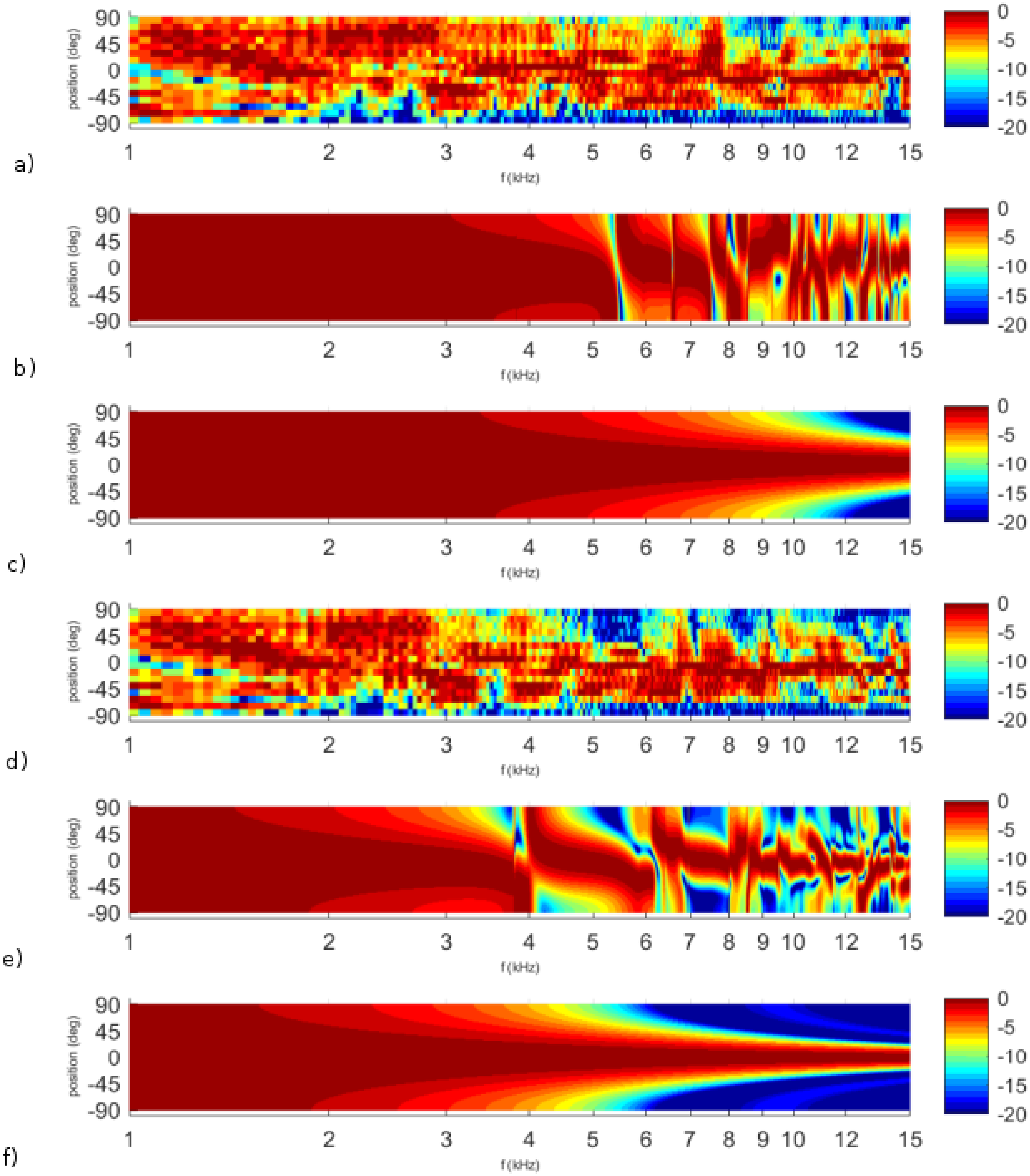

Figure 2. Normalized amplitude of the acoustic pressure radiated by a female professional singer and obtained with theoretical models as a function of the frequency and the angular position in the vertical plane (0°is located in front of the mouth). a) measurement with a normal mouth opening, b) simulation of a normal mouth opening based on the multimodal theory, c) simulation of a normal mouth opening using the vibrating piston model, d) measurement with a long mouth opening, e) simulation of a long mouth opening based on the multimodal theory and f) simulation of a long mouth opening using the vibrating piston model. No diffraction by the head and torso is taken into account in the simulations.

The plane piston model explains the general trend of a beam narrowing as the frequency increases. It also describes the influence of the size of the mouth aperture on the main beam width: a large mouth aperture tends to generate more focused radiation patterns. This could affect the distance impression: the source seems to be nearer because less reverberation is heard. However, the plane piston model fails to explain the complexity of the patterns and the global downward direction of the beam above 2 kHz.

The multimodal method allows one to extend the plane piston model by adding the propagation of the higher order propagation modes inside the vocal tract and shows the influence on the radiation patterns. Taking into account higher order modes allows one to reproduce complex pattern variations, but only above the cut-off frequency of the first higher order modes (above 5 kHz and between 3-4 kHz for the normal and long configurations respectively). These complex patterns correspond to the radiation patterns of the higher order modes [17, 7]. The qualitative similarities between the measured patterns and the simulated ones (abrupt transitions and localized minima) show that the complexity observed experimentally can be, at least partially, attributed to higher order modes above 3-4 kHz. For the long configuration, regions of high and low levels attributable to higher order mode propagation occur around similar frequencies and angles. The patterns are not exactly similar, which was expected as the exact vocal tract geometry is not reproduced in the simulations.

An increase of mouth opening results in a shift of the effect of higher order modes to lower frequencies as well as an increase of their bandwidth. In fact, this is due to the larger vertical dimensions which reduce the cut-off frequencies of the higher order modes compared to the normal configuration.

The complexity of the measured radiation patterns below the cut-off frequency of the higher order modes can be explained by the diffraction by the head and the torso. Thus, the overall complexity of the measured radiation patterns probably results of a combination of the effect of the diffraction by the head and the torso and the radiation of higher order modes.

Since the experimental patterns of both mouth configurations are very similar up to 2 kHz, the effect of diffraction by the head and the torso is probably predominant in this interval. Above 2 kHz the influence of the mouth shape probably becomes more significant and starts to influence the radiation patterns:

- first through the width of the radiation beam, related to the size of the mouth opening. In fact, the deviation between both mouth configurations between 2-3 kHz is not predicted by the multimodal theory because the wavelength is still large compared to the mouth opening. Therefore, it is likely that the diffraction by the head of different width of radiation beam introduce unequal effects.
- Above the cut-off frequency of the first higher order modes, the mouth configuration substantially changes the radiation. Since higher order modes are influenced by the shape of the mouth opening, the internal shape of the mouth can potentially influence the radiation.

The fact that the a downward orientation of the beam can be reproduced by the multimodal simulation indicates that the shape of the mouth can potentially influence the main direction of radiation. However, the trend is not as clearly marked in the simulation as in the measurement. As the vocal tract of the singer has not been exactly replicated and merely approximated and no modeling of the lips nor the head and the torso are included, it was not expected to have an exact matching in the frequency interval of occurrence of the higher order modes. Still, the multimodal simulations predict that a larger mouth opening changes radiation drastically above 3.5 kHz.

A significant amount of energy is radiated in the 2 kHz to 4 kHz interval which also corresponds to the maximum of hearing sensitivity of the hear, and frequencies in the 8 kHz and 16 kHz octave bands can be relevant for the perception of singing quality [18]. On the other hand, such changes of beam width and direction of a sound source have been shown to be perceptually relevant [19, 20, 21]. Thus, the influence of the mouth configuration on directivity can potentially lead to perceptible effects. On the other hand, the higher order modes can increase the width of the beam and strongly modify its direction in some specific frequency intervals. This could be perceptible and play a role in the naturalness of speech and singing.

5 CONCLUSION

The measurement of the sound radiated by a professional singer shows that the use of two different mouth configuration leads to noticeably different radiation patterns. Thus, from 2 kHz on, a more focused radiation beam shifted downward is generated using a long mouth opening (which increases the vertical dimensions of the mouth).

The comparison of these measurements with simulations performed with higher order modes and a vibrating piston model shows that the general characteristics of the radiated sound can be estimated from the vibrating piston properties. Thus, the decrease of the width of the radiation beam at high frequency and with a greater mouth opening corresponds to the prediction of this model. However, it fails to predict the complex local variations of the directivity patterns as well as the global angular downward shift of the radiation beam in the case of the long mouth configuration.

On the contrary, complex local variations of the directivity patterns and downward shifts of the radiation beam are observed both experimentally and in the multimodal simulations. However, the global downward tendency is weaker. Accounting for the diffraction by the head and the torso and using a more realistic geometry could help to simulate in a more realistic way the radiation of speech and singing.

ACKNOWLEDGEMENTS

Part of this research was funded by the German Research Foundation (DFG), grant no. BI 1639/7-1.

REFERENCES

[1] Flanagan, JL. Analog measurements of sound radiation from the mouth, The Journal of the Acoustical Society of America, Vol 32, 1960, pp 1613–1620.

[2] Marshall, AH.; Meyer, J. The directivity and auditory impressions of singers, Acta Acustica united with Acustica, Vol 58 (3), 1985, pp 130-140.

[3] Chu, WT.; Warnock, ACC. Detailed directivity of sound fields around human talkers, Technical Report, Institute for Research in Construction (National Research Council of Canada, Ottawa ON, Canada), 2002, pp. 1–47.

[4] Katz, B.; d'Alessandro, C. Directivity measurements of the singing voice, International Congress on Acoustics (ICA 2007), 2007.,

[5] Monson, BB.; Hunter, EJ.; Story, BH. Horizontal directivity of low-and high-frequency energy in speech and singing, The Journal of the Acoustical Society of America, Vol 132 (1), 2012, pp 433-441.

[6] Kocon, P.; Monson, BB. Horizontal directivity patterns differ between vowels extracted from running speech, The Journal of the Acoustical Society of America, Vol 144 (1), 2018, pp EL7-EL12.

[7] Blandin, R,; Van Hirtum, A.; Pelorson, X.; Laboissière, R. The effect on vowel directivity patterns of higher order propagation modes, Journal of Sound and Vibration, Vol 432, 2018, pp 621-632.

[8] Kinsler, LE.; Frey, AR.; Coppens, AB.; Sanders, JV. Fundamentals of Acoustics, 4th edition, Wiley, 1999, p 189.

[9] Brandner, M.; Frank, M.; Rudrich, D. DirPat—Database and Viewer of 2D/3D Directivity Patterns of Sound Sources and Receivers. Audio Engineering Society Convention 144, 2018.

[10] Nair, A.; Nair, G.; Reishofer, G. The low mandible maneuver and its resonential implications for elite singers, Journal of Voice, Vol 30 (1), 2016, pp 128.e13–128.e32.

[11] Kob, M.; Jers, H. Directivity measurement of a singer, The Journal of the Acoustical Society of America, 1999, Vol 105 (2), pp 1003-1003.

[12] Brandner, M.; Sontacchi, A.; Frank, M.; Real-Time Calculation of Frequency-Dependent Directivity Indexes in Singing, DAGA 2019 Rostock.

[13] Aalto, D.; Aaltonen, O.; Happonen, RP.; Jääsaari, P,; Kivelä, A.; Kuortti, J.; Luukinen, JM.; Malinen, J.; Murtola, T.; Parkkola, R.; others Large scale data acquisition of simultaneous MRI and speech, Applied Acoustics, Vol 83, 2014, pp 64-75.

[14] Arnela, M.; Dabbaghchian, S.; Blandin, R.; Guasch, O.; Engwall, O.; Van Hirtum, A.; Pelorson, X. Influence of vocal tract geometry simplifications on the numerical simulation of vowel sounds, The Journal of the Acoustical Society of America, Vol 140 (3), 2016, pp 1707-1718.

[15] Xue, SA.; Hao, JG. Normative standards for vocal tract dimensions by race as measured by acoustic pharyngometry, Journal of Voice, Vol 20 (3), 2006, pp 391-400.

[16] Blandin, R.; Arnela, M.; Laboissière, R.; Pelorson, X.; Guasch, O.; Van Hirtum, A.; Laval, X. Effects of higher order propagation modes in vocal tract like geometries, The Journal of the Acoustical Society of America, Vol 137 (2), 2015, pp 832-843.

[17] Blandin, R.; Van Hirtum, A.; Pelorson, X.; Laboissière, R. Influence of higher order acoustical propagation modes on variable section waveguide directivity: Application to vowel [α], Acta Acustica united with Acustica, Vol 102 (5), 2016, pp 918-929.

[18] Monson, BB.; Lotto, AJ.; Story, BH. Detection of high-frequency energy level changes in speech and singing, The Journal of the Acoustical Society of America, Vol 135 (1), 2014, pp 400-406.

[19] Postma, BNJ.; Demontis, H.; Katz, BFG. Subjective evaluation of dynamic voice directivity for auralizations, Acta Acustica united with Acustica, Vol 103 (2), 2017, pp 181-184.

[20] Wendt, F.; Sharma, GK,; Frank, M.; Zotter, F.; Hoeldrich, R. Perception of spatial sound phenomena created by the icosahedral loudspeaker, Computer Music Journal, Vol 41 (1), 2017, pp 76-88.

[21] Frank, M.; Brandner, M. Perceptual Evaluation of Spatial Resolution in Directivity Patterns, DAGA 2019 Rostock.