

# **Evaluation of a proper measurement environment to determine directivity characteristics of the singing voice**

Toningenieur-Projekt

Larissa Kocher

Supervisor: DI Manuel Brandner

Graz, Mai 2021



institut für elektronische musik und akustik





## **Abstract**

Voice directivity depends on many factors, e.g. morphology: head and body shape, oral posture and vocal tract configuration for different phonemes. In this project, a proper measurement environment for the determination of sound radiation patterns of the singing voice was implemented and later evaluated.

As a measurement setup served the double circular microphone array (DCMA). For the measurement procedure, the "glissando-method" was used. Optical tracking sensors were used in order to measure the oral posture and center position of the singer. This allowed to validate the measurement and gave valuable information of the used mouth opening for the given phoneme. The video recording was used to validate the mouth tracking performance of the developed approach. The whole measurement environment was implemented in Puredata including audio and video recording as well as head position and mouth tracking, ensuring a reproducible and reliable measurement routine.

By visualizing measurements in form of sound-pressure maps and analyzing the beam width and main direction of the energy vector as well as the directivity index, it could be shown with the current measurement setup that the directivity of the singing voice is affected by the change of mouth configurations.

## **Zusammenfassung**

Die Richtwirkung der Stimme hängt von zahlreichen Faktoren, wie der Morphologie, Kopf und Körperbau, der Mundstellung, sowie der Formung des Vokaltrakts bei unterschiedlichen Phonemen ab. In diesem Projekt wurde eine geeignete Messumgebung zur Bestimmung von Richtwirkung bei Gesang implementiert und später mittels Messungen evaluiert.

Als Messsetup diente das double circle microphone array (DCMA), das Messverfahren bediente sich der "Glissando-Methode". Optische Tracking-Sensoren wurden angebracht, um die Mundformung sowie die zentrale Positionierung der Sänger\*innen zu messen. Dies sollte die Messung validieren und Information bezüglich der Mundöffnung für gegebene Phoneme geben. Die Videoaufzeichnung diente als Verifikation des hierfür entwickelten Trackingvorgangs. Die gesamte Messumgebung wurde in Puredata implementiert und enthält sowohl Audio- und Videoaufnahmen als auch das Head- und Mouth-Tracking, um eine reproduzierbare und zuverlässige Messroutine gewährleisten zu können.

Die Messdaten wurden mittels Schalldruckkarten visualisiert und anhand der Breite des Hauptschallstrahls, der Hauptrichtung des Energievektors, sowie des Bündlungsmaßes analysiert. Dadurch konnte schließlich gezeigt werden, dass die Richtwirkung bei Gesang stark von Änderungen der Mundöffnung abhängt.



# Contents

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Introduction</b>                              | <b>8</b>  |
| <b>2</b> | <b>Fundamentals</b>                              | <b>9</b>  |
| 2.1      | Far and near field . . . . .                     | 9         |
| 2.2      | Directivity . . . . .                            | 9         |
| 2.3      | Sound radiation from the mouth . . . . .         | 11        |
| <b>3</b> | <b>Methods</b>                                   | <b>13</b> |
| 3.1      | Measurement setup . . . . .                      | 13        |
| 3.1.1    | Double circle microphone array (DCMA) . . . . .  | 13        |
| 3.1.2    | Motion and mouth tracking . . . . .              | 14        |
| 3.1.3    | Measurement software . . . . .                   | 15        |
| 3.2      | Calibration . . . . .                            | 17        |
| 3.3      | Reference microphone and HATS offset . . . . .   | 17        |
| 3.4      | Metrics . . . . .                                | 19        |
| 3.4.1    | Directivity index for a single plane . . . . .   | 19        |
| 3.4.2    | Energy vector . . . . .                          | 19        |
| 3.5      | Variables . . . . .                              | 20        |
| 3.6      | Verification of the measurement method . . . . . | 20        |
| 3.7      | Glissando measurement . . . . .                  | 22        |
| 3.7.1    | Definition of mouth openings . . . . .           | 22        |
| 3.7.2    | Spectrogram of a sung glissando . . . . .        | 23        |
| <b>4</b> | <b>Results</b>                                   | <b>24</b> |
| 4.1      | HATS . . . . .                                   | 24        |
| 4.1.1    | Impact of an offset . . . . .                    | 24        |
| 4.1.2    | Summary . . . . .                                | 27        |
| 4.2      | Classical singer . . . . .                       | 27        |
| 4.2.1    | Consistency of the mouth position . . . . .      | 28        |
| 4.2.2    | Summary . . . . .                                | 31        |

|          |  |           |
|----------|--|-----------|
| 4.2.3    | Comparison of different mouth openings . . . . . | 31        |
| 4.2.4    | Summary . . . . .                                | 33        |
| <b>5</b> | <b>Conclusion</b>                                | <b>34</b> |
| <b>A</b> | <b>Puredata patches</b>                          | <b>35</b> |
| <b>B</b> | <b>MATLAB scripts</b>                            | <b>37</b> |



# 1 Introduction

It is already known that the directivity of the singing voice is affected by the change of mouth configurations [3]. It is assumed that bel canto singers use special techniques to form their vocal tract in order to achieve narrower and more focussed radiation of their singing voice.

The aim of this project was to measure the sound radiation patterns of the singing voice to investigate the impact of an assumed vocal tract forming technique.

Furthermore, this project focussed on the evaluation of a proper measurement environment for the determination of directivity characteristics of the singing voice. In order to provide a most flexible measurement environment the user interface and its signal processing was implemented in Puredata (PD). The output should be a time synchronous recording of 63 audio microphone signals in combination with a video camera signal and the data of the motion tracking system.

Testings of the implementation of the measurement environment as well as the verification of the measurement method should been done by an imitation of the DCMA where microphones on stands were placed at reference points of the DCMA. A head and torso simulator (HATS) generating an exponential sweep was used for the simulation of a human singer making use of the "glissando method". Since the mouth position of the dummy is static, the accuracy of the mouth tracking was tested on a human before passing over to the actual measurements. Deeper investigations on those prior testings are not included in this document.

For the validation of the measurement method the microphones had to be calibrated first, the offsets of the microphone positions had to be evaluated and the mouth-tracking algorithm had to be analyzed.

After a successful verification of the measurement method conclusions towards the measurement accuracy could be drawn to which the actual measurements of sound radiation of singing voice could be realized with the DCMA. The determination of directivity characteristics included the observation of the consistency of the mouth positions as well as the comparison of the different mouth positions. Directivity patterns were analyzed and visualized with adapted tools of the iem-DirPat repository<sup>1</sup>. Therefore, the directivity index and the energy vector were useful metrics to describe the sound source radiation. Sound-pressure maps and the visualization of the beam width and main direction of the energy vector should help to get an impression of the directivity characteristics.

---

1. <https://opendata.iem.at/projects/dirpat/>

## 2 Fundamentals

This chapter is about the essential theoretical background for measurements of voice directivity using the DCMA including the definition of far and near field as well as the concept of directivity.

### 2.1 Far and near field

As one may expect, the acoustic terms *far field* and *near field* [7] are defined by the physical distance from the sound source. The acoustic energy produced by the sound source will behave quite differently, depending on how far away the observer is from the sound emitting object.

The acoustic far field is defined as beginning at a distance of two wavelengths away from the sound source, and extends outward to infinity. As wavelength is a function of frequency, the start of the far field is also a function of frequency. In the far field, the source is far enough away to essentially appear as a point in the distance, with no discernable dimension or size. At this distance, the spherical shape of the sound waves have grown to a large enough radius that one can reasonably approximate the wave front as a plane-wave, with no curvature. At this distance, sound pressure level is governed by the inverse square law, and a single microphone sound recording will give reliable predictable results. For each doubling of distance away from the source, the sound pressure will drop 6 dB in the far field. The transition from near to far field is gradual in the transition region.

In many acoustic standards, measurements are often specified at a distance of at least one meter from the sound emitting object to ensure that the measurement is taken in the far field for the most critical frequencies.

When close to a sound emitting object, the sound waves behave in a much more complex fashion, and there is no fixed relationship between pressure and distance. Very close to the source, the sound energy circulates back and forth with the vibrating surface of the source, never escaping or propagating away. These are sometimes called evanescent waves. As we move out away from the source, some of the sound field continues to circulate, and some propagates away from the object. This transition from circulating to propagating continues in an unpredictable fashion until we reach the threshold distance of two wavelengths, where the sound field strictly propagates (the far field). This mix of circulating and propagating waves means that there is no fixed relationship between distance and sound pressure in the near field, and making measurements with a single microphone can be troublesome and unrepeatable. Typically, measuring in the near field requires the use of more than one microphone in order to accurately capture the energy borne by the circulating and propagating waves.

### 2.2 Directivity

The directivity of sound radiation is demonstrated by polar patterns, where the dependency on the angle of the sound pressure level in the free field is shown for singular planes as well as spatially.

The differences in phase with which the absolute values of different sectors of the source

overlap can be described with a function of the angle, in which the observer, seen from the center of the source, is located. With this given, the directivity of the source can be described by the directivity pattern  $p(\varphi, \delta)$  as sound pressure level  $p$  depending on a horizontal and vertical inclination [7]. Fig. 1 shows a generic plot of a directivity pattern of a cardioid.

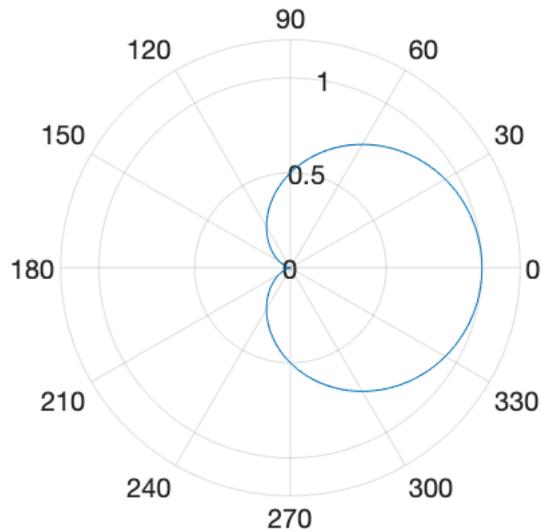


Figure 1: Generic plot of a directivity pattern.

Inside the shift from far field to near field, sound pressure levels overlap in complex ways, so they cannot be described by a directivity depending on an angle. Inside the near field, the line source approximately acts like an ideal line source with an attenuation in distance of 3 dB per doubled distance. In the far field it approximates the performance of a spherical wave with an attenuation in distance of 3 dB per doubled distance. That's why those directivity patterns are measured almost only in the far field since the sound field close to the sound source can diverge.

An important characteristic of directivity patterns is their frequency dependence independent of the produced pitch, which means that the directivity pattern for frequencies around 1000 Hz for the fundamental  $h''$  is the same as for the octave partial tone  $h'$  or the third partial of  $e'$ . Further, directivity is independent from the dynamics. Since in higher dynamic stages the parts with higher frequent components in the sound spectrum increase, directivity gains with increasing amplitude for higher frequencies importance for the sonic impression in different directions.

Directivity with its frequency dependence has highly complex structures, thus a collective simplification for practical applications is unavoidable. Therefore, the directions of strong sound radiation are interesting. Still, sometimes it is useful to know the directions of very low sound radiation. For the definition of areas with strong sound radiation, angles with a sound pressure level less than 3 dB gain loss or 10 dB below the maximum value have to be evaluated.

An omnidirectional characteristic theoretically doesn't lead to an angle dependent deviation of the sound pressure level. In practice, it is useful to take directivity patterns with

a deviation less than 3 dB as omnidirectional. The direction factor is important for the distance to the microphone as well as for spatial acoustic considerations. It describes the relation of the sound pressure level in a particular direction to the sound pressure level of an omnidirectional source with the same sound power, in the same distance.

$$\Gamma(f, \Theta, \phi) = \frac{p(f, \Theta, \phi)}{p(f, 0, 0)} \quad (1)$$

Values larger than 1 show an outstanding strong and values below 1 indicate a relatively low sound radiation. At the threshold to the above mentioned 3 dB area, the directivity factor decreases to the 0.7-fold, at the 10 dB threshold to the 0.3-fold of its maximum value.

## 2.3 Sound radiation from the mouth

The mouth opening of a human head can be approximated by a spherical source on a spherical baffle, as mentioned by Brandner in [4]. The spherical source can even be simplified by a small piston. The influence of mouth openings corresponding to different radii of a piston can be described by a simple calculation. Considering that directionality increases for  $k\frac{d}{2} \gg 1$ , it can be assumed that for a mouth opening with a diameter  $d = 3$  cm and  $d = 8$  cm this effect will appear at different frequencies. The following equations in (2) and (3) state these frequencies  $f_\gamma$  for the two diameters, with the wave number  $k = \frac{2\pi f_\gamma}{c}$ . The ratio of frequency  $f_\gamma$  is given in radians and the speed of sound  $c$  in meter per seconds.

$$f_\gamma = \frac{343 \frac{m}{s}}{\pi} \frac{1}{0.03m} \approx 3640 \text{Hz} \quad (2)$$

$$f_\gamma = \frac{343 \frac{m}{s}}{\pi} \frac{1}{0.08m} \approx 1365 \text{Hz} \quad (3)$$

Taking this into consideration, a remarkable change in directivity for large mouth openings around  $kd = 2$ , meaning at double the frequency of  $f_\gamma$ , seems promising. If we designate the far-field pressure for a circular baffled piston with the radius  $r$ , directivity  $D$  at specific spherical coordinates  $\phi$  and  $\Theta$ , azimuth and zenith angle, can be defined as following equation:

$$D(f, \Theta, \phi) = 2\pi r^2 \frac{J_1(kr \sin\Theta)}{kr \sin\Theta}. \quad (4)$$

Whereas  $J_1$  stands for the bessel function of the first kind. Fig. 2 shows the results for various diameter displayed at 2, 3, 4 and 5 kHz in the horizontal plane. It can be seen that a shift in frequency corresponds to a change in diameter, whereas the shift in frequency shows a stronger effect in regards of sound power level at lateral positions.

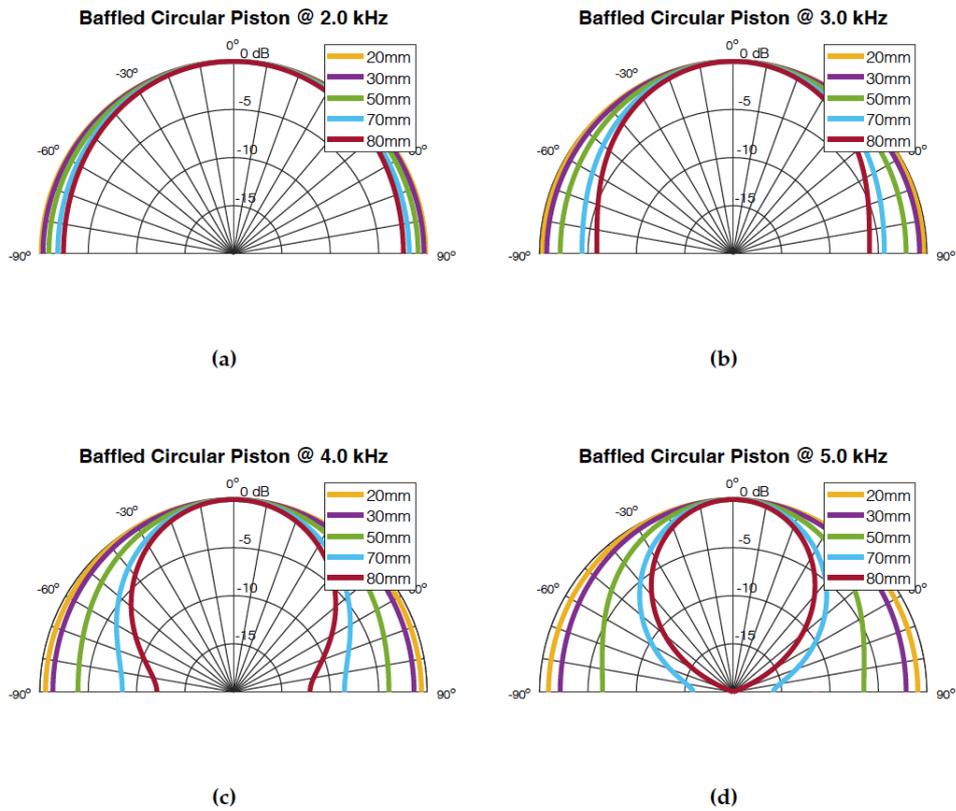


Figure 2: Different diameters (20, 30, 50, 70 and 80mm) of a baffled circular piston at the frequencies a) 2 kHz, b) 3 kHz, c) 4 kHz and d) 5 kHz. A shift towards higher frequencies or a larger diameter provoke higher directivity. A shift in frequency has a stronger impact on reducing lateral sound radiation. [4]

Slight deviations in the spectral structure can be perceived quite well, since the high energy of the singer's formant is within the frequency region where the human auditory system is most sensitive.

### 3 Methods

The whole measurement procedure includes several steps as a detailed cause-and-effect analysis of voice directivity is quite difficult due to the influences like posture, vowel tract geometry, and subject differences. Therefore, the measurement comprises of hardware and software that gathers all information i.e: all audio channels, video, and motion and mouth tracking at once.

#### 3.1 Measurement setup

The measurement chain for the setup starts with audio signals recorded by using the DCMA and the tracking signals recorded by nine cameras. The audio and video signals are processed in puredata (PD) and stored as a 64 channel Waveform Audio File Format. Fig. 3 shows the setup of the measurement chain in detail.

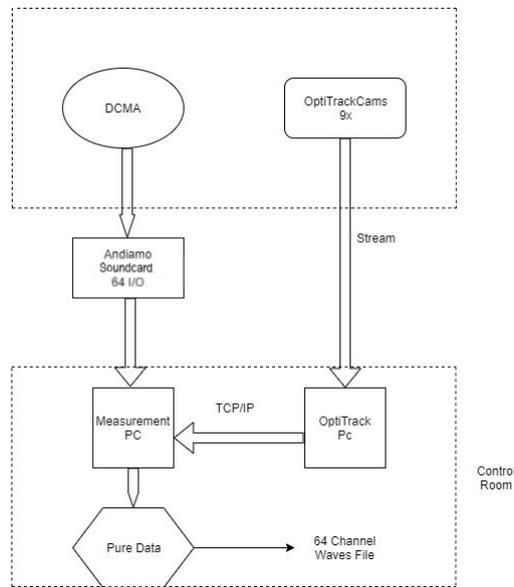


Figure 3: Measurement chain

##### 3.1.1 Double circle microphone array (DCMA)

The DCMA itself consists of two circular rings, one placed in the horizontal, the other one in the vertical plane. The two rings can hold together a maximum number of 62 microphones, as both rings intersect in the front and back of the array. With a radius of 1 m the microphones are spaced with  $11.25^\circ$ . At its widest point the apparatus has got a diameter of 2.56 m. To achieve a reduction of reflections, the 21 mm thick center-facing side of the rings is beveled. Due to the adjustable loudspeaker stands, the center of the array can be lifted to any height between 1.3 m and 2 m. The microphones mounted to the array are NTI MA 2230 and they are connected to an Andiamo.MC Directout Technologies microphone preamplifier.

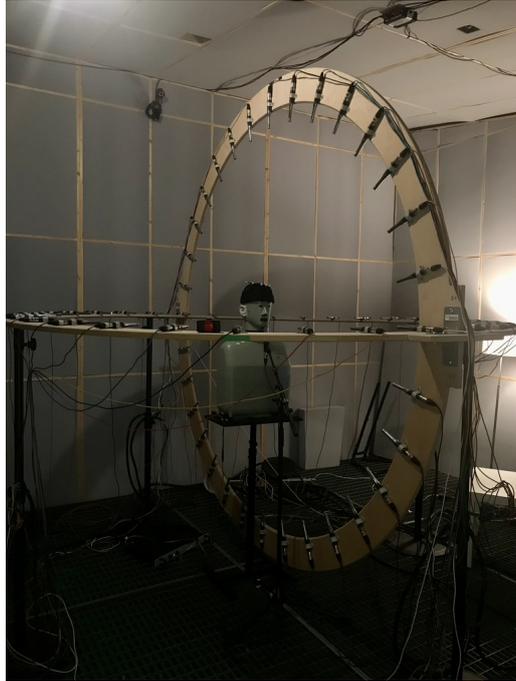


Figure 4: DCMA setup

The DCMA has a radius of 1 m, which means that the microphones positioned on it are placed in the far field. In the near field, disproportionate increase of sound particle velocity and phase shift of sound particle velocity and sound pressure would result in the proximity effect, leading to overexposure of lower frequencies. Since measurements are taking place in the far field, they shouldn't be affected by those artefacts.

### 3.1.2 Motion and mouth tracking

The setup for our measurement procedure consisted of nine OptiTrack motion tracking cameras. Six of those cameras were placed in the upper edge of the  $x/y$  plane (see Fig.5), which were covering well the  $x/y$  plane the head tracker is placed in. Three cameras were placed in the front  $y/z$  plane on microphone stands for the tracking of the eight facial markers placed around the dummy's or human's mouth. USB hubs were used to merge the USB signals and amplify them, a tracking computer, a computer for measurements and we needed internet connection.

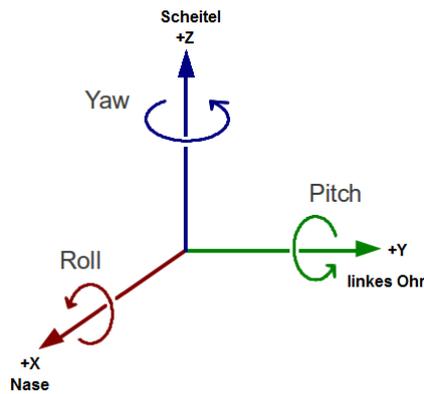


Figure 5: 3D coordinate system and quaternions transformed to yaw-roll-pitch.

### 3.1.3 Measurement software

Basically, two main patches are needed in order to record and process the tracking data as well as the acoustical measurements. These two patches are the tracking patch and the audio recording patch.

#### Tracking patch

`Pd_Data_Patch Transform_Coordinates = track_singer_data_v18.pd`

The tracking patch does all sufficient computations of the tracking data. Detailed information on the pd tracking patches can be found in appendix A.

#### Audio recording patch

`rec_singer_audio_v18.pd` is the second important PD patch operating the audio recordings as well as visualizing the head and mouth position of the singer. Detailed information on the pd audio patches as well as a screenshot of the GUI can be found in appendix A.

#### WAV file channels

The WAV file consists of 64 channels. Those are used for audio data as well as tracking data.

Channel 63 consists the tracking data, which is divided in several sequences:

|                           |                            |
|---------------------------|----------------------------|
| sequence of start samples | 0, 1, 2, (+0.99, -0.99, 0) |
| rigidbody_1_xyz           | 3, 4, 5                    |
| rigidbody_1_ypr           | 6, 7, 8 (yaw, pitch, roll) |
| othermarker_1             | 9, 10, 11                  |
| othermarker_1_xyz         | 12, 13, 14                 |
| othermarker_2_xyz         | 15, 16, 17                 |
| othermarker_3_xyz         | 18, 19, 20                 |
| othermarker_4_xyz         | 21, 22, 23                 |
| othermarker_5_xyz         | 24, 25, 26                 |
| othermarker_6_xyz         | 27, 28, 29                 |
| othermarker_7_xyz         | 30, 31, 32                 |
| othermarker_8_xyz         | 33, 34, 35                 |
| othermarker_9_xyz         | 36, 37, 38                 |
| othermarker_10_xyz        | 39, 40, 41                 |
| othermarker_11_xyz        | 42, 43, 44                 |
| othermarker_12_xyz        | 45, 46, 47                 |
| othermarker_13_xyz        | 48, 49, 50                 |
| othermarker_14_xyz        | 51, 52, 53                 |
| othermarker_15_xyz        | 54, 55, 56                 |
| rigidbody_2_xyz           | 57, 58, 59                 |
| rigidbody_3_xyz           | 60, 61, 62                 |
| sample index 63           | is free                    |

Table 1: Tracking data of channel 63.

Fig. 6 shows an exemplary visualization of the tracking data see Tab. 2.

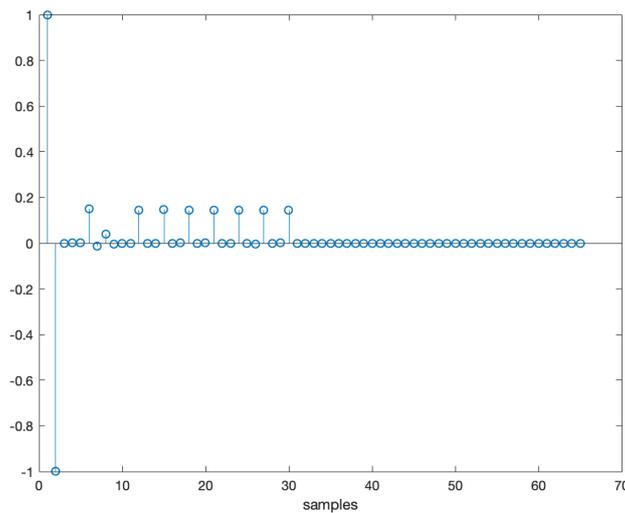


Figure 6: One segment of the tracking data of channel 63.

Fig. 7 shows a schematic of how the 62 microphone channels on the DCMA were arranged.

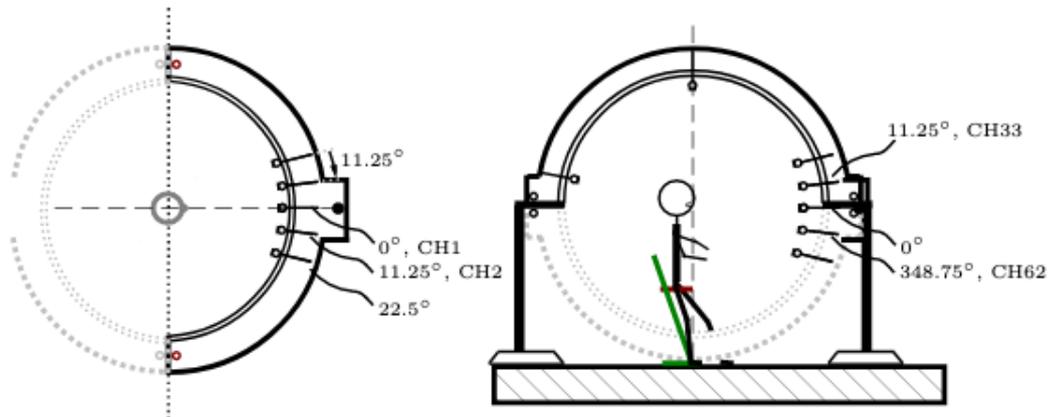


Figure 7: Schematic on the 62 channels of the DCMA.

## 3.2 Calibration

A calibration of all microphones is necessary to guarantee correct results from the measurements. For the microphones the calibration is done by a calibrator providing a 1 kHz sine with a level of 94 dB. The absolute sound level is recorded in PD in order to receive the same level as for the intended measurements. In Reaper, the 64 channel wav-file containing the calibration levels of each microphone is split into single mono channels in order to be cut to the same length and afterwards being rendered as one 64 channel wav-file. Additionally, the microphones are calibrated to each other to ensure that all of them have got the same sensitivity.

The calibration is done in MATLAB using the MATLAB scripts `calc_calibration.m`, which calibrates the microphones to each other and `calibration.m`, which calibrates each channel to 94 dB. Both MATLAB scripts are attached to B.

## 3.3 Reference microphone and HATS offset

The measurements took place in the sound treated measurement room with absorptive material on the walls as well as the floor and ceiling at the IEM.

The reference microphone is put in the center of the DCMA with the acoustical center of the HATS, meaning the HATS' mouth, in front of the microphone, but with a 5 cm offset in the negative z direction in the horizontal plane.

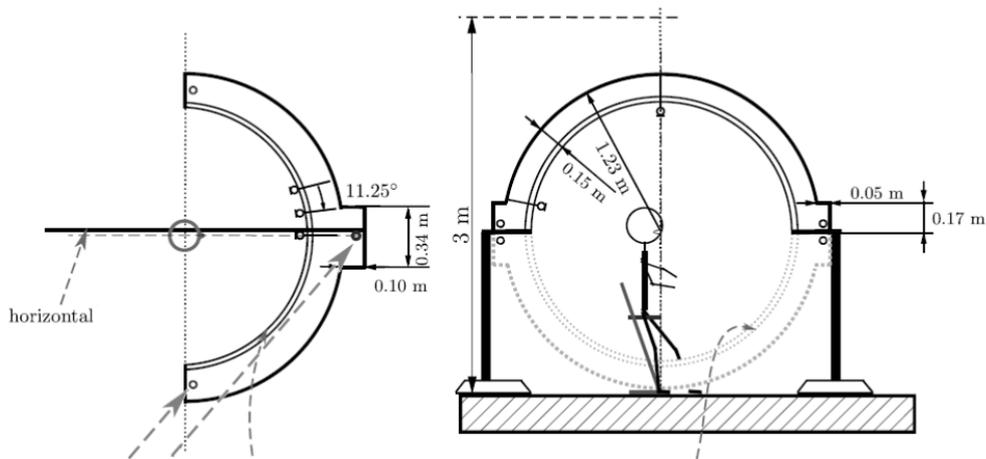


Figure 8: Schematic of the DCMA setup

This method will be also used for the test subjects because it ensures a constant distance from the test subject to the microphones of the DCMA throughout the entire measurement. This way, larger deviations due to positioning errors on the directivity pattern are averted. The positioning of the reference microphone and the HATS has been executed by laser measurements to avoid any inaccuracies and unwanted offsets.

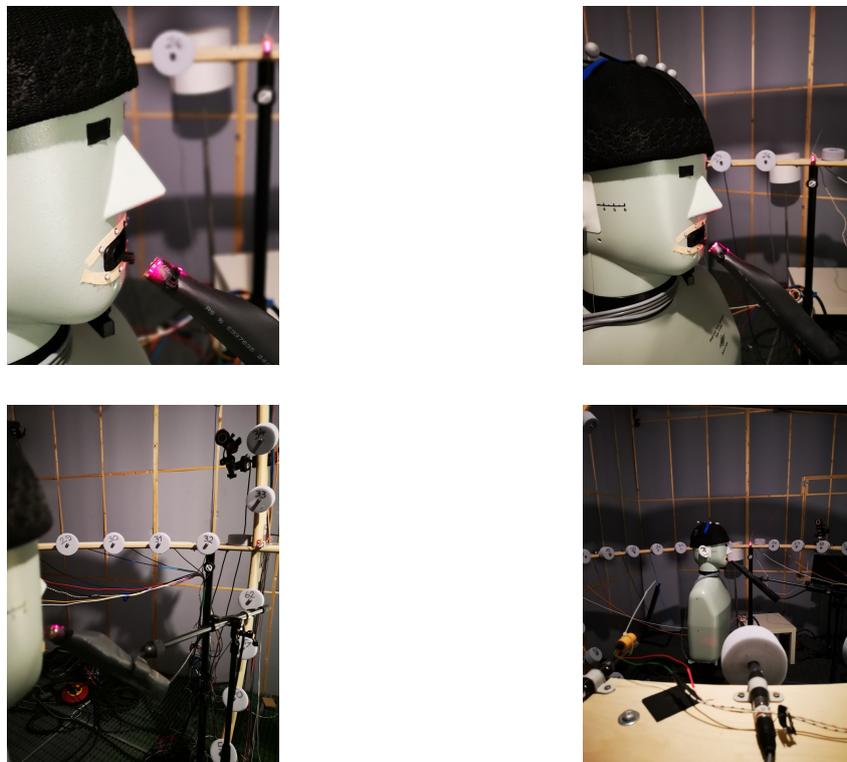


Figure 9: Centering of the reference microphone in front of the HATS and the HATS itself with an offset of 5 cm to the center in 0° direction in the DCMA with a laser.

An exponential sweep generated by the HATS is recorded and de-convoluted in MATLAB using deconvolution (see appendix B) with the original sweep and then again with the recorded sweep of the reference microphone. This way it could be shown that moving the HATS' mouth away from the center with a maximum offset of about 15 cm doesn't have any decisive impact on the directivity results, as in 3.6.1 results have shown.

### 3.4 Metrics

The following metrics were used to evaluate differences between measurements and were calculated from third-octave smoothed data.

#### 3.4.1 Directivity index for a single plane

The degree of focus can be calculated by integrating the direction factor (1) over a spherical surface as in [7]

$$\gamma(f) = \frac{4\pi}{\oint \Gamma^2(\phi_i, \delta, f) dS} \quad (5)$$

This degree of focus can be converted to the directivity index by using the logarithm of it

$$DI(f) = 10 \log \gamma(f) \text{ dB} \quad (6)$$

describes the difference of the sound pressure level from the affected area to an undirected source.

For our case, we define it for the horizontal and the vertical plane separately, as Brandner in [3]. It describes the ratio of the power on-axis  $P_{on-axis}$  to the mean of the power  $P_{mean}$  of all sampling positions on the respective plane.

$$\gamma_P(\omega) = \frac{P_{on-axis}}{P_{mean}} \quad (7)$$

So the HDI and VDI computed at an angular frequency  $\omega$  for the horizontal and the vertical plane can be specified in dB as:

$$DI(\omega) = 10 \log_{10}(\gamma_P(\omega)) \quad (8)$$

#### 3.4.2 Energy vector

With the energy vector  $\mathbf{r}_E$  the direction and the width of the main lobe of an acoustic source can be described. Using this measure it is possible to describe the properties of any arbitrary sound source radiation:

$$\mathbf{r}_E = \frac{\sum_{i=1}^L |H(\omega, \phi_i)|^2 \cdot \mathbf{m}_i}{\sum_{i=1}^L |H(\omega, \phi_i)|^2} \quad (9)$$

The frequency dependent magnitudes  $H(\omega, \phi_i)$  need to be multiplied by the vectors  $\mathbf{m}_i = [\cos(\phi_i), \sin(\phi_i)]^T$  of each measurement position  $i=1,2, \dots, L$  in each respective plane.

From the normalization of the vector by the sum of the energy values between 0 (omnidirectional) and 1 (maximum focus to one direction) are obtained. For the analysis in this project the following two metrics will be used:

- (i) the main beam width in each plane  $\Theta_w$ , which is two times the arc cosine of the norm of the energy vector:  $\Theta_w = 2\arccos\|\mathbf{r}_E\|$
- (ii) the main direction in the vertical plane  $\Theta_s = [\cos(\phi_s), \sin(\phi_s)]^T$ .

### 3.5 Variables

Several variables are needed to make statements about sound radiation. The ones included in this project are sound-pressure depicted in so-called sound-pressure maps, the beam width and the main direction of the energy vector.

With sound pressure maps the amplitude of the acoustic pressure normalized by its maximum over the angular position for each frequency can be shown. A dynamic limitation of 20 dB is applied for the visualization. In order to receive all necessary details for the comparison, in Fig. 11 no smoothing is applied to the frequency resolution. For further visualization,  $\frac{1}{3}$ -octave smoothing is applied.

### 3.6 Verification of the measurement method

For the verification of the measurement method several steps had to be taken into consideration.



Figure 10: Setup of the DCMA simulation

For first test measurements the DCMA was simulated by a setup with five microphones at a height of 1.4 m placed at  $\pm 11.25^\circ$ ,  $\pm 45^\circ$ ,  $180^\circ$ , three microphones at  $0^\circ$  with a

height of 1.4 m +/- 13.5 cm and 15 cm above the ground. The microphones had a distance of 1 m to the HATS placed in the center at a distance of 1 m to the microphones.

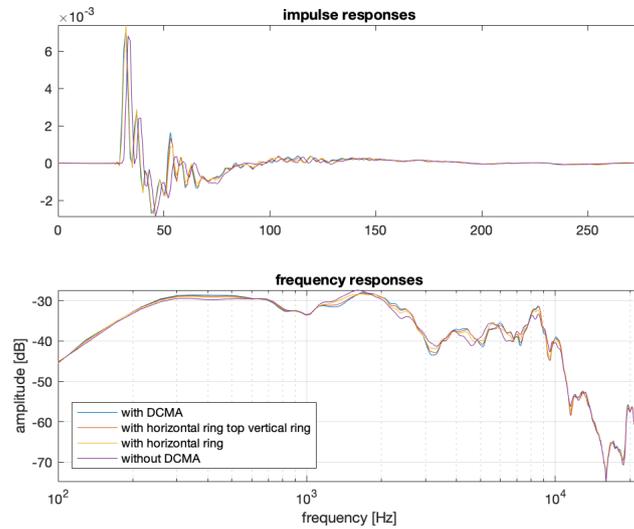


Figure 11: Comparison of different DCMA setups.

Fig. 11 shows exemplarily the influence of the wooden rings of the DCMA in  $0^\circ$  direction. By comparing impulse responses as well as frequency responses, almost identical developments could be observed below 1 kHz respecting both, the horizontal and the vertical ring, just one ring, and no rings, meaning the simulation by microphones on their stands. The reason for almost identical behaviour is diffraction. The biggest deviation can be observed at 2 kHz with a delta of 3 dB. Above 10 kHz the developments are rather identical again. Thus, it could be proven that the wooden rings of the DCMA have a negligible impact on the recordings.

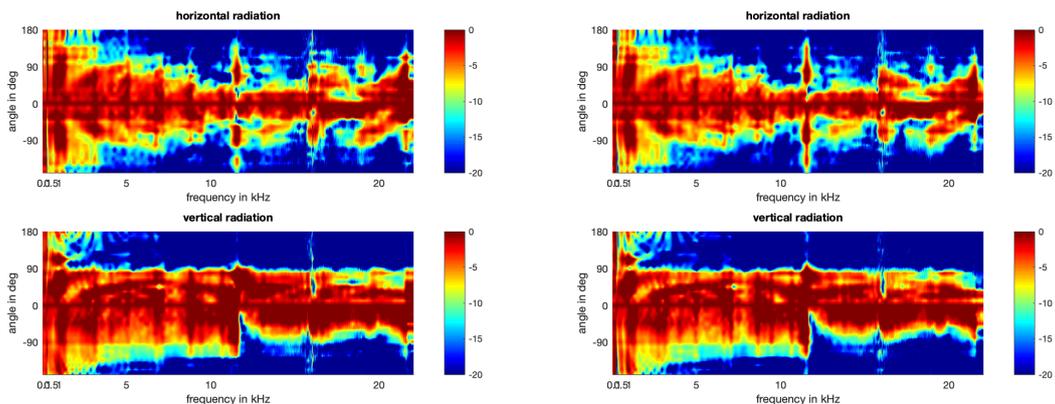


Figure 12: Left: Results after deconvolution with the analytical sweep and right: after deconvolution with the recorded signal from the reference microphone in front of the mouth opening of the HATS.

In Fig. 12 the horizontal and vertical sound pressure maps of the analytical sweep as well as the recorded sweep with the HATS in centered position are depicted. To investigate the differences between the convoluted analytical sweep and the recorded sweep, both produced by the HATS, the difference between the directivity pattern of the analytical sweep and the directivity pattern of the recorded sweep has been computed and visualized, as can be seen in Fig. 13.

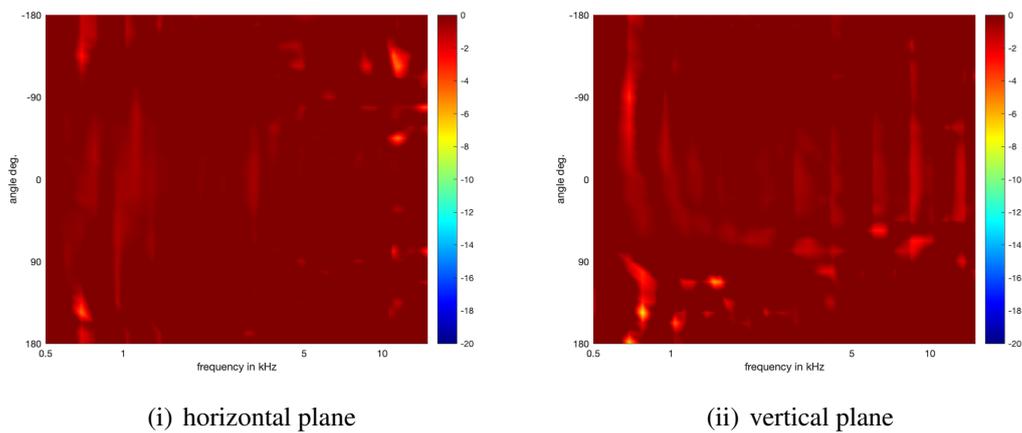


Figure 13: Difference between the directivity pattern of the analytical sweep and the directivity pattern of the recorded sweep in the horizontal (i) and vertical (ii) plane.

The biggest differences reaching up to 6 dB could be detected in the side lobes around 700 Hz in both, the horizontal and the vertical plane, and 11 kHz in the horizontal plane. All other differences lie below 4 dB and only affect smaller spots in the side lobes. Thus, it can be said that hardly any differences could be computed, which implicates that the glissando method could be verified successfully.

### 3.7 Glissando measurement

After the measurement method had been verified successfully, acoustical measurements with a singer could be realised. For the glissando method, where the singer was asked to sit in the center of the DCMA setup and sing the vowel /a:/ with four different mouth openings while starting at a low pitch (e) and raising the pitch over one octave. This way a wide range of frequencies can be captured.

#### 3.7.1 Definition of mouth openings

The four different mouth openings chosen for this procedure include "normal" like in speech and for a closed vowel, "long" with the jaw lowered, "open" with the jaw lowered and a large horizontal mouth width and "wide" mouth opening [2].



Figure 14: Sketch of the four proposed mouth openings. Left to right: "normal", "long", "open" and "wide". [4]

### 3.7.2 Spectrogram of a sung glissando

Fig. 15 shows exemplarily the spectrogram of a sung glissando on the vowel /a:/ starting at e with open mouth configuration. As can be seen, a wide range of frequencies could be captured.

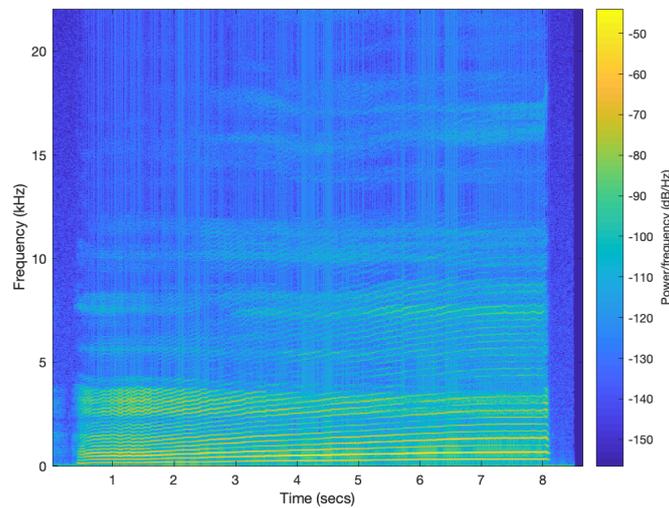


Figure 15: Glissando starting at e with open mouth configuration.

## 4 Results

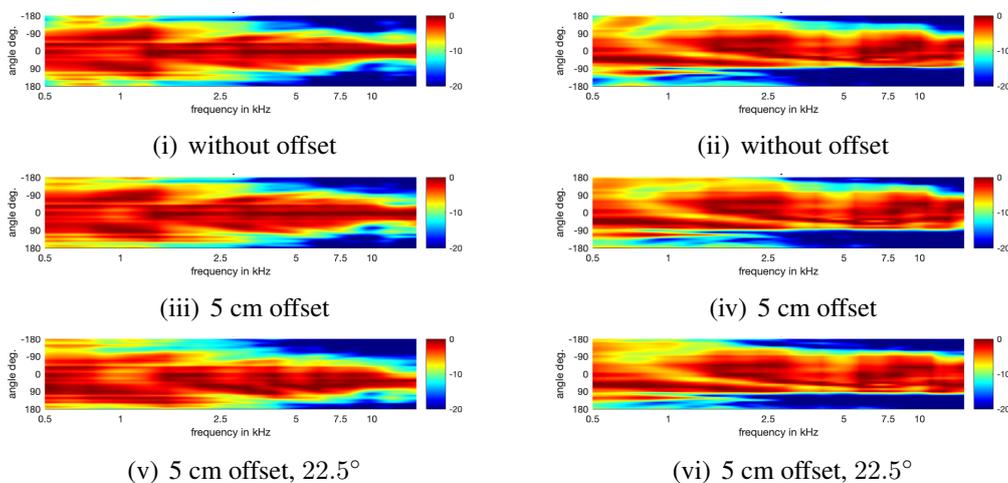
The measurements are divided into two main parts. First, measurements were carried out with the HATS to investigate the impact of an offset position on the singing voice directivity metrics. Afterwards, glissando measurements with a classical singer were carried out to give an insight on the consistency of the singer's mouth position and the impact of different mouth configurations on voice directivity.

### 4.1 HATS

#### 4.1.1 Impact of an offset

In Fig.16 one can see the horizontal and vertical normalized sound pressure maps of the recorded sweep of the HATS at the centered position (Fig.16 (i) and (ii)). Further, subfigures show the HATS being placed 5 cm off center with its head on axis (Fig. 16 (iii) and (iv)),  $22.5^\circ$  off axis (Fig. 16 (v) and (vi)) and  $45^\circ$  off axis (Fig. 16 (vii) and (viii)) and the same head rotations for a 15 cm offset (Fig. 16 (ix)-(xiv)).

In regard to the horizontal radiation, slight changes can be observed in the normalized sound pressure maps when increasing the offset up to 5 cm, see Fig. 16 (i) and (iii). An increase in the side lobes from 1 kHz to 2.5 kHz as well as above 10 kHz can be detected. Those side lobes around 1 kHz are produced by reflections of the torso. Increasing the offset to 15 cm (Fig. 16 (ix)) results in a larger increase at the side lobes up to 3 kHz. Further, a decrease of amplitude of the main lobe between 1 and 2 kHz occurs when increasing the offset to 15 cm and rotating the head towards  $45^\circ$  (Fig. 16 (xiii)). By turning the head  $22.5^\circ$  (Fig. 16 (xi)) or  $45^\circ$  (Fig. 16 (xiii)) off-axis, a change in distance from 5 cm offset to 15 cm offset results in a decrease of about 5 dB in the side lobes between 2 kHz and 5 kHz. Moving the head off axis provokes a shift of the main lobe and side lobes in the respective direction.



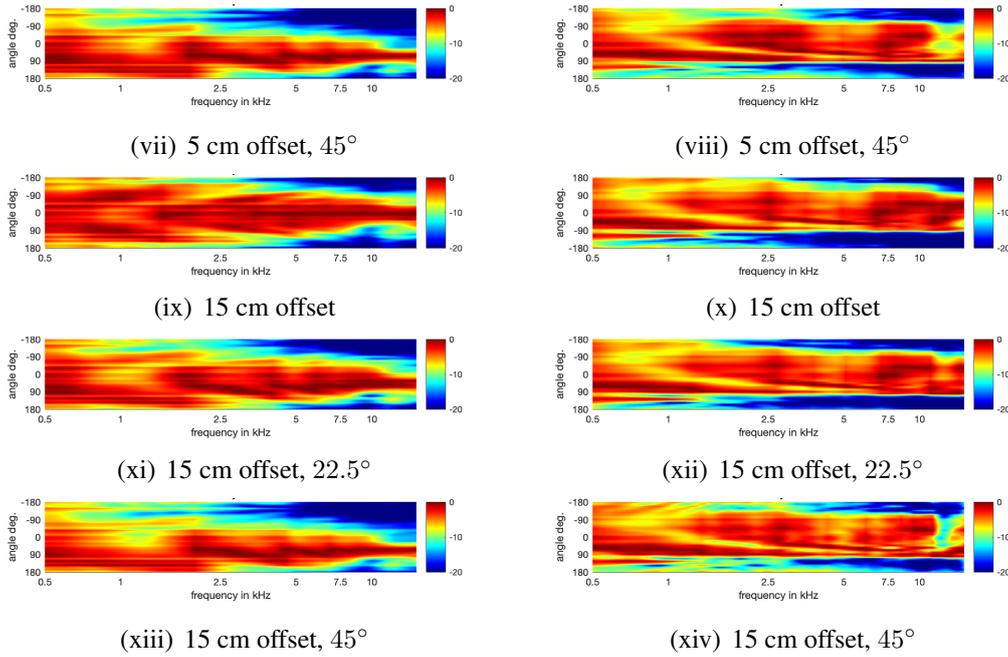


Figure 16: Normalized sound pressure maps of the HATS in the horizontal plane (left) and in the vertical plane (right).

In the vertical plane, similar phenomena as in the horizontal plane can be detected, whereas in the vertical plane a shift from the 5 cm offset position (Fig. 16 (iv)) to the 15 cm offset position (Fig. 16 (x)) or a rotation of the head position from on axis (Fig. 16 (iv) and (x)) to 45° off axis (Fig. 16 (viii) and (xiv)) shows less prominent influences in the acoustic radiation patterns. Again, the normalized sound pressure maps show increasing side lobes. Only this time this effect occurs when rotating the head from 0° on axis (Fig. 16 (iv)) to 45° (Fig. 16 (xiv)). In contrast to the horizontal plane, moving the head off axis doesn't provoke a shift of the main lobe and side lobes in the respective direction. The sound radiation patterns show a decay of energy in the main lobe between 1.5 kHz and 10 kHz throughout several positions starting at Fig. 16 (ii) and reaching Fig. 16 (xiv) with partly 5 dB less. Above 12 kHz a decrease of about 10 dB can be detected in the 5 cm offset (Fig. 16 (iv)-(viii)) and even 15 dB in the 15 cm offset positions (Fig. 16 (x)-(xiv)), when rotating the head up to 45° off axis.

Fig. 17 now shows the beam width of the energy vector at the different offset positions in the horizontal (see Fig. 17 (i) and (iii)) and the vertical (see Fig. 17 (ii) and (iv)) plane. Each graph shows the comparison of three head positions 0°, 22.5° and 45°.

In the horizontal plane as well as in the vertical plane, the beam width is decreasing rapidly above 1 kHz. Increasing the distance to the microphone up to 15 cm doesn't have any decisive impacts on the beam width. Head rotations, on the contrary, seem to have a larger effect on the beam width. In the horizontal plane, a 45° rotation of the head can lead to a decrease of 30° of beam width, as can be seen at about 1.2 kHz in Fig. 17 (iii). In the vertical plane, rotating the head up to 45° provokes a flatter decrease of the beam width above 1.3 kHz than keeping the head on axis. At 11 kHz, a difference of 40° of the beam width between 0° and 45° can be detected.

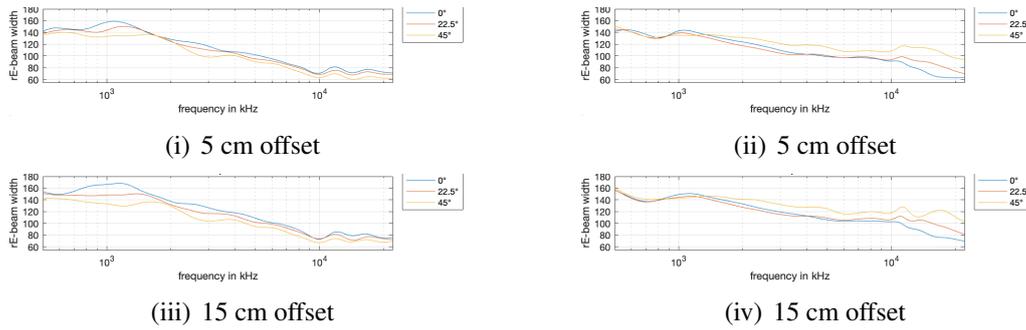


Figure 17: Beam width of the  $r_E$  vector of the recorded sweep with HATS in the horizontal plane for an offset of (i) 5 cm and (iii) 15 cm, and in the vertical plane for an offset of (ii) 5 cm and (iv) 15 cm.

Fig. 18 depicts the main direction of the energy vector at the different offset positions in horizontal (i), (iii) and vertical (ii), (iv) plane. As expected, in the horizontal plane the direction changes with rotation of the head. Still, in the horizontal plane it can be observed, the a  $22.5^\circ$  head rotation provokes a main direction of the energy vector of about  $40^\circ$ . A  $45^\circ$  head rotation results in a main direction of about  $50^\circ$ . The reason is a different distance from the mouth to the torso. Further, the sound source is out of the array center when the head is rotated. In the vertical plane, rotating the head doesn't have a strong impact on the main direction of the energy vector. The main direction of the energy vector is rather consistent, independent from any head rotations.

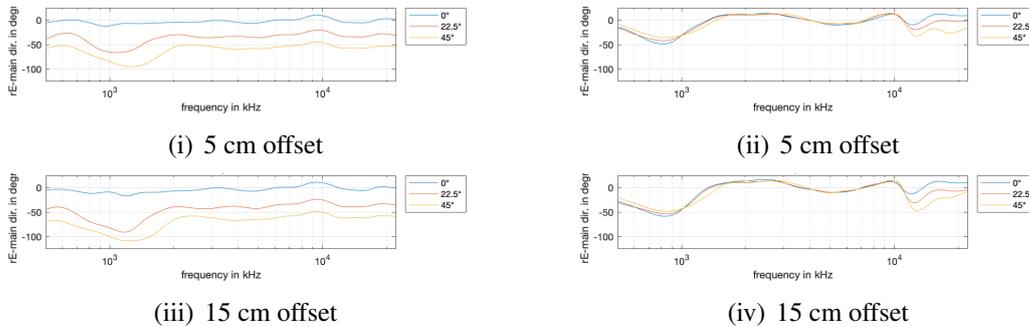


Figure 18: Main direction of the  $r_E$  vector of the recorded sweep with HATS in the horizontal plane (left) and in the vertical plane (right).

The functionality of the horizontal directivity index (HDI) and vertical directivity index (VDI) as shown in Fig. 19 is to quantify the directionality in the corresponding plane. Regarding the horizontal and vertical directivity index, Fig. 19 shows that increasing the offset position from 5 cm to 15 cm doesn't provoke any differences in the developments. Head rotations, conversely, lead to larger deviations in the HDI (Fig. 19 (i) and (iii)) as well as the VDI (Fig. 19 (ii) and (iv)). Independent from the offset as well as the respective plane, they lead to similar developments of the directivity index. As already seen in Fig. 16, strong side lobes provoked by the torso occurred around 1 kHz. Those

side lobes are destructive interferences, which now can be detected as a decrease of the directivity index decreasing to negative values around 1 kHz. Also, the drops of the HDI and VDI values represent the ripples visible in the normalized sound pressure maps at the same frequencies. The diverging developments above 10 kHz leading to a difference of up to 10 dB between  $0^\circ$  and  $45^\circ$  head rotations are caused by the head and mouth opening of the HATS. Above 10 kHz the head acts as a baffle and the mouth opening is larger than the wave length. This provokes resonances in the mouth of the HATS and is visible in inconsistencies of the measurement curves.

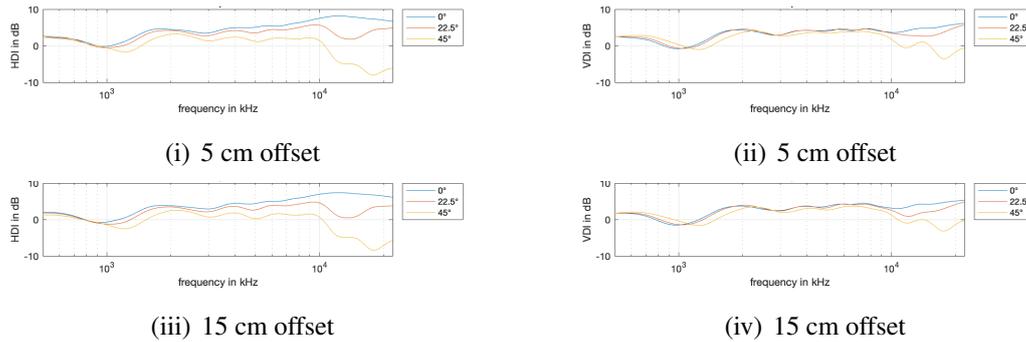


Figure 19: HDI and VDI of the recorded sweep with HATS in the horizontal plane (left) and in the vertical plane (right).

#### 4.1.2 Summary

Having analyzed normalized sound pressure maps, the beam width and main direction of the energy vector and the HDI and VDI, it can be said that an offset of 5 cm as well as potential rotation of the head up to  $22.5^\circ$  off axis hardly show any significant deviations to the head on axis. Since the singer is asked to keep his head still, it should be possible to keep an offset of 5 cm. Then, the error shows minimal influence on the sound radiation and therefore seems to be neglectable. On the contrary, an offset of 15 cm would already be too large and therefore could not be accepted for this measurement method as a maximum of tolerance limit.

## 4.2 Classical singer

The glissando measurements have been executed holding the same mouth position for three times on the vowel /a:/. Four different mouth openings as already defined in 3.7.1 and depicted in Fig.14 were investigated. This way not only the difference between each of the four mouth opening positions, but also the capability of reproducing the same vowel holding the same mouth position can be showed.

Below 500 Hz, no interpretations are attempted, since the singer started his glissando at around 300 Hz. Thus, all plots start at a frequency of 500 Hz.

### 4.2.1 Consistency of the mouth position

The following graphs show the sound radiation measured by using the glissando method (3.7). They show the comparison of three recordings of the vowel /a:/ for each of the mouth positions in the horizontal plane as well as the vertical plane.

As can be seen in Fig. 20 (ii) and (iv), a deviation of about  $10^\circ$  in the relevant frequency range between 1 kHz and 5 kHz could be measured with normal mouth opening.

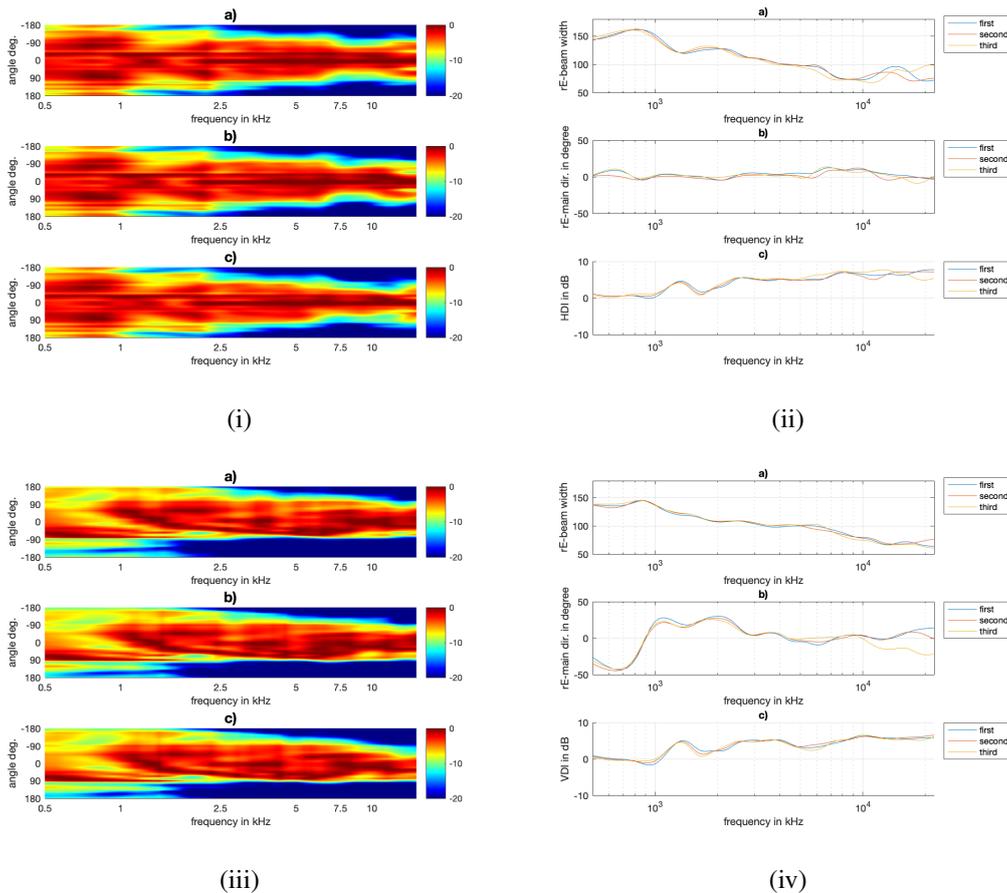
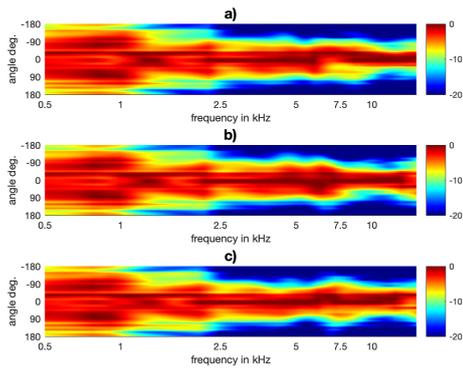
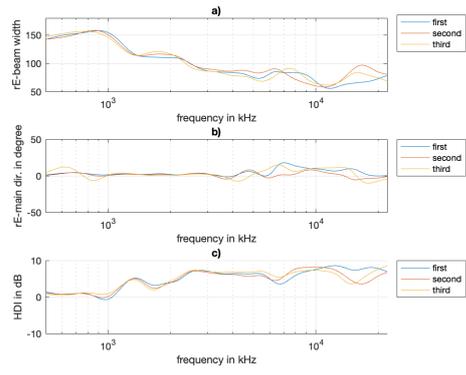


Figure 20: Glissando on the vowel a recorded three times in the horizontal (a), (b) and vertical (c), (d) plane. Normal mouth position.

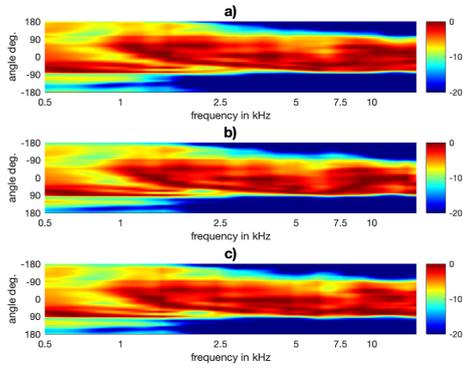
Fig. 21 (ii) and (iv) again shows a deviation of about  $10^\circ$  in the relevant frequency range between 1 kHz and 5 kHz, this time with wide mouth opening.



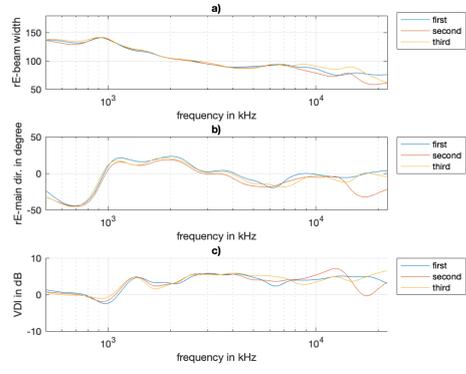
(i)



(ii)



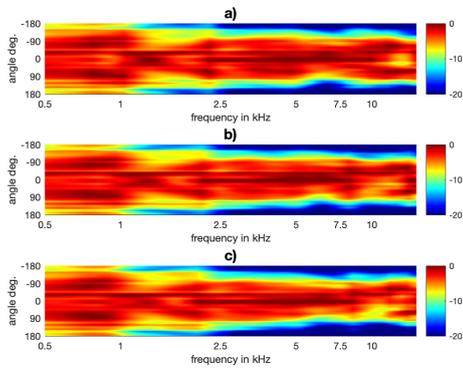
(iii)



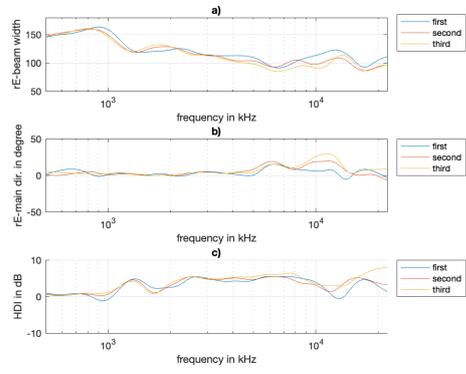
(iv)

Figure 21: Glissando on the vowel a recorded three times in the horizontal (a), (b) and vertical (c), (d) plane. Wide mouth position.

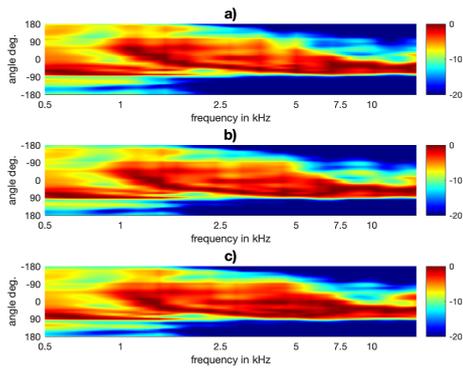
As for the past two mouth positions, with long mouth opening (Fig. 22 (ii) and (iv)) again no larger deviations than  $10^\circ$  could be detected in the relevant frequency range between 1 kHz and 5 kHz.



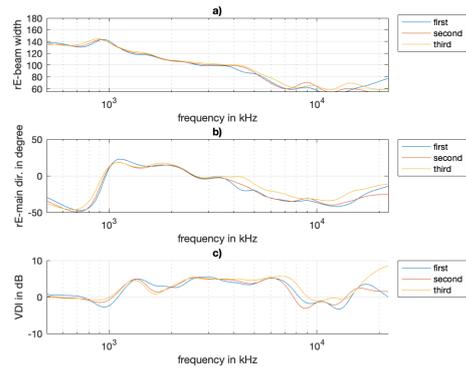
(i)



(ii)



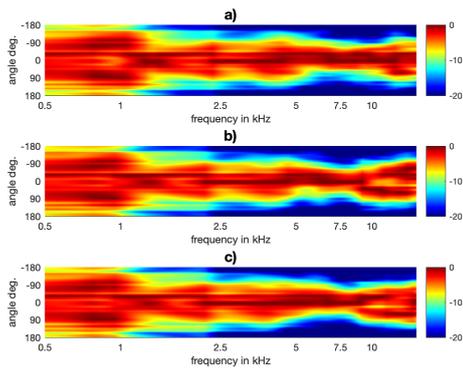
(iii)



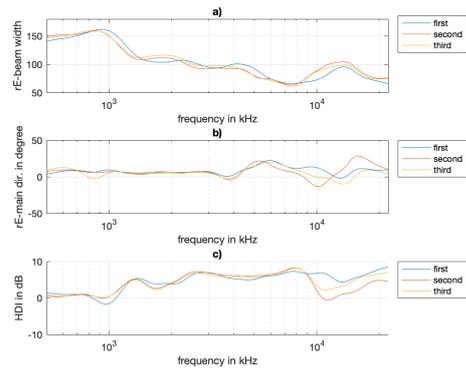
(iv)

Figure 22: Glissando on the vowel a recorded three times in the horizontal (a), (b) and vertical (c), (d) plane. Long mouth position.

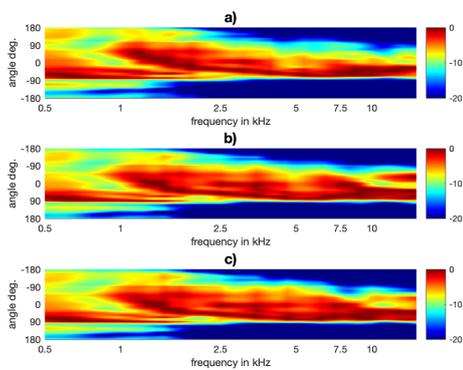
Fig. 23 (ii) and (iv) shows the glissando with open mouth position. Differences between the three recordings don't reach beyond  $10^\circ$  in the frequency range between 1 kHz and 5 kHz.



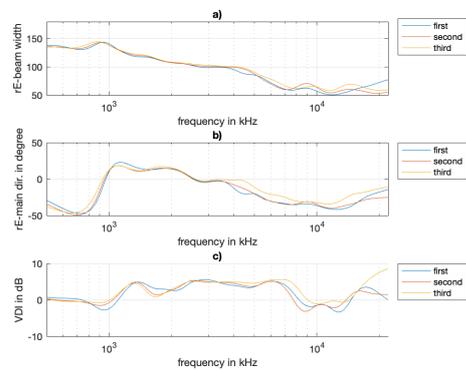
(i)



(ii)



(iii)



(iv)

Figure 23: Glissando on the vowel a recorded three times in the horizontal (a), (b) and vertical (c), (d) plane. Open mouth position.

#### 4.2.2 Summary

As can be seen in Fig.20 - 23, the singer was able to keep his mouth opening stable for each of the four positions. No larger deviations than  $10^\circ$  in beam width and directivity of the energy vector can be detected, meaning the singer could reproduce the same quality of sound radiation three times. The main direction of the energy vector in the horizontal plane shows the largest deviations between the three recordings, independent from the mouth opening. Short breaks of a few seconds between each of the three recorded glissandi might have led the singer to slightly move his head. Thus, minimal movements of his head might have provoked implied differences, which can be described as measurement inaccuracies or the slight inconsistencies due to repetitions of the recording. The shifts of the measurement curves around 1 kHz are provoked by the torso, which leads to reflections. Around 10 kHz the same effect is caused by the head and mouth opening of the singer. As already explained in 4.1.1, above 10 kHz the head acts as a baffle and the mouth opening is larger than the wave length. This provokes resonances in the singer's mouth and is visible in inconsistencies of the measurement curves.

#### 4.2.3 Comparison of different mouth openings

Fig. 24 and Fig. 25 depict the four mouth openings recorded at one take compared to the HATS as reference and to each other once in the horizontal and once in the vertical plane. As measurements have showed, for all mouth opening positions, an increase of the main beam width in the horizontal plane can be recognised in the frequency range up to 1 kHz (Fig. 24 (i) a)). Above 1 kHz it's decreasing steadily. Above 1.5 kHz the main beam widths of wide and long configuration diverge and build the two extremes of the different mouth openings. Long mouth configuration is reaching its minimum earlier than all other configurations, between 6 kHz and 7 kHz. Wide mouth opening reaches the smallest beam width above 10 kHz. The long mouth opening shows in contrast to all other mouth configurations a broader main beam width around 10 kHz. Also, above 2 kHz the wide and open configuration have got a narrower main beam width than the other mouth openings. Above 3 kHz, the difference to the open configuration is larger as expected. In the vertical plane the beam width acts the opposite way (Fig. 25 (i) a)). There, long and open configuration reach their minima last at above 10 kHz. The reason for the decrease of beam width of the long mouth configuration is similar as for the behaviour above 10 kHz. Because of the lowered jaw the torso is closer to the sound source and therefore the influence of the torso on reflections towards the front persists. In contrast to that, with wide configuration the mouth has got a bigger distance to the torso. Thus, torso reflections don't have a comparable impact on the main beam width of wide opening in this frequency range. This is why wide configuration has a rather omnidirectional behaviour in the vertical plane.

Above 2 kHz all configurations imply rather complex sound radiation. Especially for the long and wide mouth configurations several local minima occur in the horizontal plane (Fig. 24 (i) b)). For the long configuration this happens at 5 kHz and around 10.8 kHz, for the wide configuration at 6 kHz and 10.5 kHz.

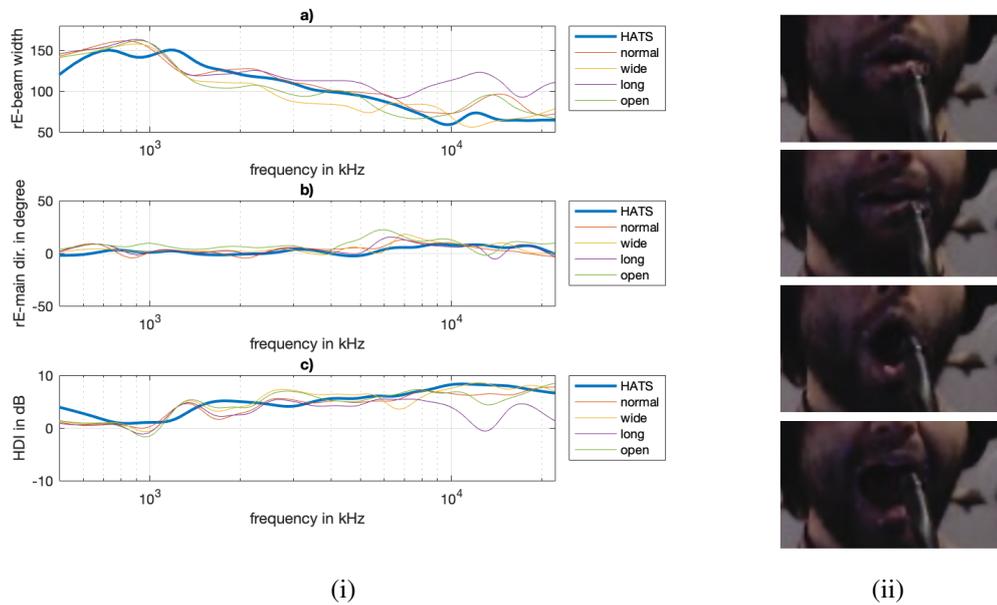


Figure 24: Glissando on the vowel a (i) with four different mouth openings (ii) in the horizontal plane.

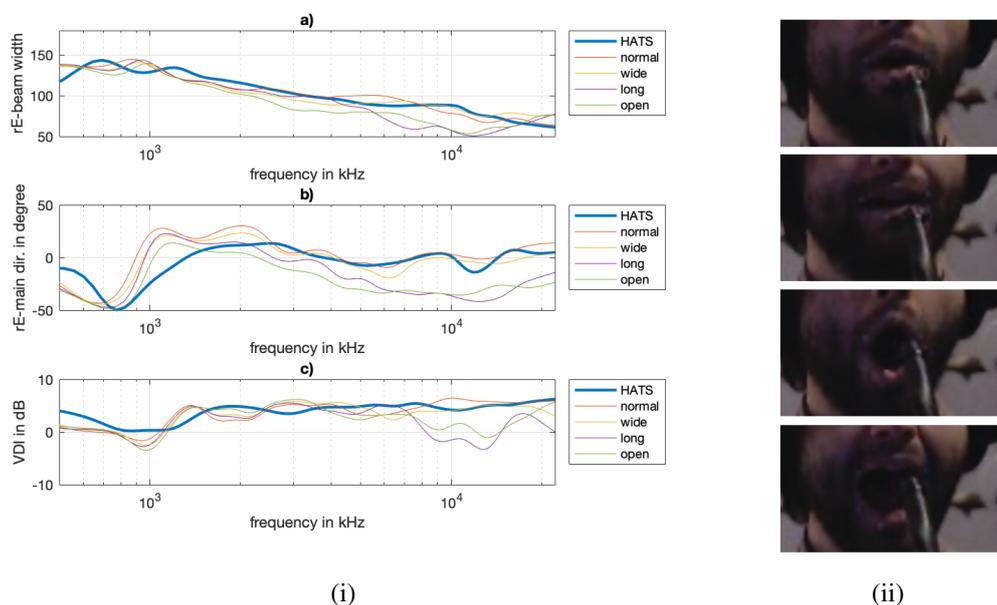


Figure 25: Glissando on the vowel a (i) with four different mouth openings (ii) in the vertical plane.

Further, it can be seen that larger mouth openings provoke an increase in directivity. Thus, the HDI of long configuration decreases above 8 kHz, whereas the HDI of the other configurations rather increases (Fig. 24 (i) c)). In contrast to that, a decrease of the VDI above

8 kHz can be observed with open and wide configuration. This is leading to a difference of about 9 dB in HDI and VDI.

#### 4.2.4 Summary

Measurements have shown that changing the mouth configuration to a larger extent provokes differences in sound radiation. Interesting findings are the asymmetries along the horizontal and vertical planes, meaning an increase or decrease in the main beam width as well as a shift of the main direction. In the horizontal plane, the direction of the main beam is not quite stable, a remarkable change of direction of the main beam happens around 4 kHz with all configurations. In the vertical plane, the direction of the main beam is quite similar with all mouth configurations, with a split at 6 kHz. There, normal and wide configuration maintain stable, whereas long and open configuration are decreasing. This can be explained as an effect of lowering the jaw for these mouth configurations.

In general, measurements have shown that all mouth configurations show rather complex patterns. As already discussed in 4.2.2, the shifts of the measurements curves around 1 kHz are provoked by the torso, which leads to reflections. Around 10 kHz the same effect is caused by the head and mouth opening of the singer. Above 10 kHz the head acts as a baffle and the mouth opening is larger than the wave length. This provokes resonances in the mouth of the singer and is visible in inconsistencies of the measurement curves. Nevertheless, it can be said that the biggest changes in sound radiation come with a change to wide and long mouth configurations above 1 kHz. In the frequency range between 1 kHz and 5 kHz deviations of almost  $30^\circ$  in beam width and main direction of the energy vector as well as deviations of almost 10 dB in HDI and VDI can be recognized.

Thus, the change of mouth openings has a greater influence on the directivity characteristics of the singing voice than measurement inaccuracies or the slight inconsistencies due to repetitions of the recording.

## 5 Conclusion

The aim of this project was to measure the sound radiation patterns of the singing voice to investigate the impact of an assumed vocal tract forming technique.

Furthermore, this project focussed on the evaluation of a proper measurement environment for the determination of directivity characteristics of the singing voice. In order to provide a most flexible measurement environment the user interface and its signal processing was implemented in Puredata (PD). The output was a time synchronous recording of 63 audio microphone signals in combination with a video camera signal and the data of the motion tracking system. The optical tracking sensors used for these measurements did not only ensure a reliable measurement routine of the oral posture and center position of the singer. The GUI with the visualization of the tracking points was of great help for the singer to maintain head and mouth positions stable throughout the entire measurements.

Testings of the implementation of the measurement environment as well as the verification of the measurement method had been done by an imitation of the DCMA where microphones on stands were placed at reference points of the DCMA. This way the influence of the wooden rings of the DCMA could be investigated. A head and torso simulator (HATS) generating an exponential sweep was used for the simulation of a human singer making use of the "glissando method". For the validation of the measurement method the microphones had to be calibrated first, the offsets of the microphone positions had to be evaluated and the mouth-tracking algorithm had to be analyzed.

The determination of directivity characteristics included the observation of the consistency of the mouth positions as well as the comparison of the different mouth positions. Directivity patterns were analyzed and visualized. Therefore, the directivity index and the energy vector were useful metrics to describe the sound source radiation. Sound pressure maps and the visualization of the beam width and main direction of the energy vector helped to get an impression of the directivity characteristics.

The measurements have shown that the directivity of singing voice is affected significantly by the change of mouth configurations. Special techniques used by classical singers to efficiently form their vocal tract for a wider radiation can be approved by our measurements as it could be shown that a larger mouth opening provokes an increase in directivity. Still, not only different mouth openings, but the torso as well, shows impacts on the directivity of singing voice.

Further, measurements have proven that professional singers perform rather precisely and therefore can reproduce what they are singing for several tries without any significant deviations.

## A Puredata patches

### Tracking patch

```
Pd_Data_Patch Transform_Coordinates = track_singer_data_v18.pd
```

```
Pd Optitrack_Interface
```

The data of the coordinates is being sent via user data protocol (UDP) internet by Motive via the NatNat bridge of UDP receive. A data gate converts the data in unpackOSC from binary to ASCII. The list then is converted to a selector message.

In Motive the head tracker is defined as rigid body 1 and at each side of the flap another rigid body was positioned, which is defined as rigid body 2 and rigid body 3. Those predefined rigid bodies are being routed in different paths.

After /rigidbody/1 the list is splitted with seven float elements in three cartesian coordinates in meter and four quaternion values for rotation. Further, the coordinates +y+z-x are being checked of inequality to zero. The quaternions are also being checked of inequality to zero and afterwards transformed to +yaw-roll-pitch. Of /rigidbody/2 and /3 just the first three values are used and checked of inequality to zero.

The facial markers are defined as /othermarker/0 reaching up to /othermarker/15. Their index comes in front of their three coordinates in the list, as in+y+z-x. With the pd functions XYZ-swap-directions and XYZ-mirror-directions the coordinates of /rigidbody/1 are transformed to the format of +x+y+z.

With the pd function ROT\_swap\_directions, ROT\_mirror\_directions and rad → deg the rotations of /rigidbody/1 are transformed to the format +yaw+pitch+roll with degree as its unit. With the pd functions POINT-swap-directions and POINT\_mirror\_directions the coordinates of the facial markers /othermarker/ are transformed to the format i+x+y+z. With the last two groups of XYZ\_swap\_directions and XYZ\_mirror\_directions the coordinates of /rigidbody/2 and /3 are transformed to the format of +x+y+z.

In parallelize\_data the already transformed values are being parallelized with the following order:

|                            |                                |
|----------------------------|--------------------------------|
| rigidbody_1_xyz            | 0 1 2                          |
| rigidbody_1_yaw_pitch_roll | 3 4 5                          |
| othermarker_0_xyz          | 6 7 8 (9 10 11, ..., 48 49 50) |
| othermarker_15_xyz         | 51 52 53                       |
| rigidbody_2_xyz            | 54 55 56                       |
| rigidbody_3_xyz            | 57 58 59                       |

Table 2: Parallelization of +x+y+z values

For the area, where the mouth is being detected, a cuboid is defined as a discriminator.

## Audio recording patch

rec\_singer\_audio\_v18.pd is the pd script for visualizing the head and face tracking. The following picture demonstrates the GUI of this script.

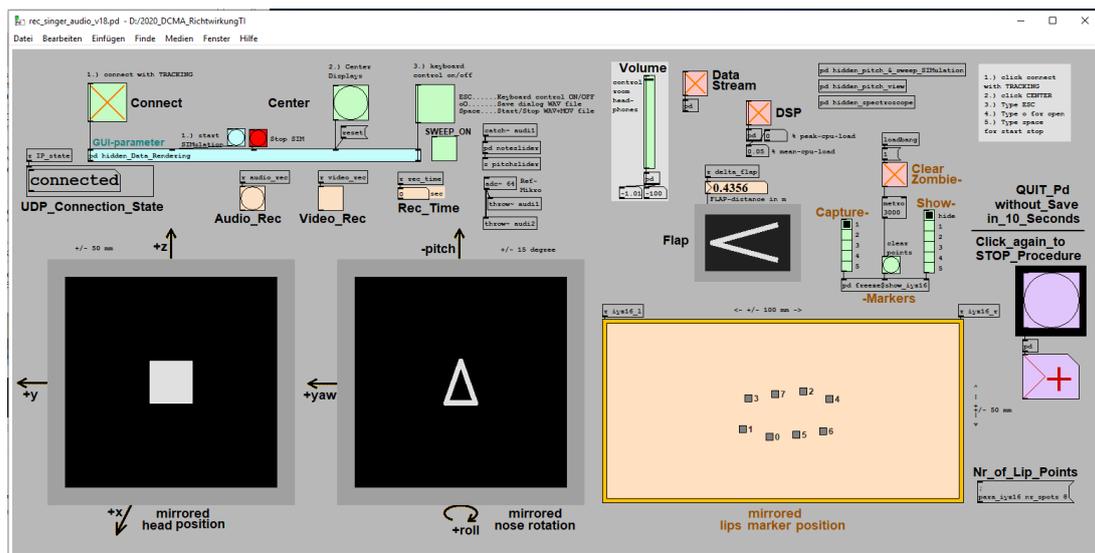


Figure 26: GUI of the pd patch for head and face tracking.

## B MATLAB scripts

calc\_calibration.m

```
1 function c = calc_calibration
2 % merges cut calibration files into one 64 channel file and calculates
3 % db to be added to recorded signal
4
5 list=dir('*chan*.wav');
6
7 for ii = 1:size(list,1)+1
8     ii
9
10     if ii == 63
11         X(:,ii) = zeros(size(x,1),1);
12     elseif ii == 64
13         fname=list(63).name;
14         [x fs]=audioread(fname);
15         X(:,ii) = x;
16     else
17         fname=list(ii).name;
18         [x fs]=audioread(fname);
19         X(:,ii) = x;
20     end
21 end
22 audiowrite('../calibrationFiles/calib.wav', X, 44100);
23
24 calib_db = db(rms(X))+abs(max(db(rms(X))));
25 calib_db(63) = 0;
26 calib_db(64) = calib_db(64) + 10;
27
28 %save(['../matlabCodes/cal'],'Xc');
29
30 %calibrate to 94dB
31 cal = 10.^(calib_db/20);
32 Pref = 20e-6;
33 RMS = sqrt(mean(cal.^2));
34 SPLmat = 20*log10(RMS/Pref);
35 c = 10.^(94-SPLmat)/20;
36
37 [y fs] = audioread('../Aufnahmen/HATS_sweep.wav');
38 yc = c.*y;
39 audiowrite('../kalibrierteAufnahmen/HATS_sweep.wav', yc, 44100);
40 end
```

calibration.m

```
1 clear all;
2 close all;
3
4 [cal fs] = audioread('../calibrationFiles/calib.wav');
5 [data fs] = audioread('../Aufnahmen/HATS_sweep.wav');
6
7 calib_db = db(rms(cal))+abs(max(db(rms(cal))));
8 calib_db(63) = 0;
9 calib_db(64) = calib_db(64) + 10;
10
11 %calibrated audio record
12 dc = db(rms(data))+calib_db;
```

## Device directory

|                              |   |
|------------------------------|---|
| 63 Measurement microphones   | NTI MA 2230                             |
| Calibrator                   | Brüel & Kjør Sound Calibrator Type 4231 |
| microphone preamplifier      | Andiamo.MC Directout Technologies       |
| HATS                         | Brüel & Kjør                            |
| DCMA                         |   |
| Cap with head-tracking array |   |
| Mouth-tracking points        |   |
| Tracking Software            | OptiTrack                               |

## List of Figures

|    |  |    |
|----|--|----|
| 1  | Generic plot of a directivity pattern. . . . .   | 10 |
| 2  | Different diameters (20, 30, 50, 70 and 80mm) of a baffled circular piston at the frequencies a) 2 kHz, b) 3 kHz, c) 4 kHz and d) 5 kHz. A shift towards higher frequencies or a larger diameter provoke higher directivity. A shift in frequency has a stronger impact on reducing lateral sound radiation. [4] . . . . . | 12 |
| 3  | Measurement chain . . . . .  | 13 |
| 4  | DCMA setup . . . . .   | 14 |
| 5  | 3D coordinate system and quaternions transformed to yaw-roll-pitch. . .  | 15 |
| 6  | One segment of the tracking data of channel 63. . . . .  | 16 |
| 7  | Schematic on the 62 channels of the DCMA. . . . .  | 17 |
| 8  | Schematic of the DCMA setup . . . . .  | 18 |
| 9  | Centering of the reference microphone in front of the HATS and the HATS itself with an offset of 5 cm to the center in 0° direction in the DCMA with a laser. . . . .  | 18 |
| 10 | Setup of the DCMA simulation . . . . .   | 20 |
| 11 | Comparison of different DCMA setups. . . . .   | 21 |
| 12 | Left: Results after deconvolution with the analytical sweep and right: after deconvolution with the recorded signal from the reference microphone in front of the mouth opening of the HATS. . . . .   | 21 |
| 13 | Difference between the directivity pattern of the analytical sweep and the directivity pattern of the recorded sweep in the horizontal (i) and vertical (ii) plane. . . . .  | 22 |
| 14 | Sketch of the four proposed mouth openings. Left to right: "normal", "long", "open" and "wide". [4] . . . . .  | 23 |
| 15 | Glissando starting at e with open mouth configuration. . . . .   | 23 |

|    |  |    |
|----|--|----|
| 16 | Normalized sound pressure maps of the HATS in the horizontal plane (left) and in the vertical plane (right). . . . .   | 25 |
| 17 | Beam width of the $r_E$ vector of the recorded sweep with HATS in the horizontal plane for an offset of (i) 5 cm and (iii) 15 cm, and in the vertical plane for an offset of (ii) 5 cm and (iv) 15 cm. . . . . | 26 |
| 18 | Main direction of the $r_E$ vector of the recorded sweep with HATS in the horizontal plane (left) and in the vertical plane (right). . . . .   | 26 |
| 19 | HDI and VDI of the recorded sweep with HATS in the horizontal plane (left) and in the vertical plane (right). . . . .  | 27 |
| 20 | Glissando on the vowel a recorded three times in the horizontal (a), (b) and vertical (c), (d) plane. Normal mouth position. . . . .   | 28 |
| 21 | Glissando on the vowel a recorded three times in the horizontal (a), (b) and vertical (c), (d) plane. Wide mouth position. . . . .   | 29 |
| 22 | Glissando on the vowel a recorded three times in the horizontal (a), (b) and vertical (c), (d) plane. Long mouth position. . . . .   | 30 |
| 23 | Glissando on the vowel a recorded three times in the horizontal (a), (b) and vertical (c), (d) plane. Open mouth position. . . . .   | 30 |
| 24 | Glissando on the vowel a (i) with four different mouth openings (ii) in the horizontal plane. . . . .  | 32 |
| 25 | Glissando on the vowel a (i) with four different mouth openings (ii) in the vertical plane. . . . .  | 32 |
| 26 | GUI of the pd patch for head and face tracking. . . . .  | 36 |

## List of Tables

|   |  |    |
|---|--|----|
| 1 | Tracking data of channel 63. . . . .       | 16 |
| 2 | Parallelization of +x+y+z values . . . . . | 36 |

## References

- [1] Manuel Brandner, Matthias Frank, Daniel Rudrich. *DirPat - Database and Viewer of 2D/3D Directivity Patterns of Sound Sources and Receivers*. AES Convention e-Brief 425, 2018.
- [2] Manuel Brandner, Alois Sontacchi, Matthias Frank. *Influences of Oral Posture on the Voice Directivity of a Classical Singer*. Journal of Voice, 2018.
- [3] Manuel Brandner, Remi Blandin, Matthias Frank, Alois Sontacchi. *A pilot study on the influence of mouth configuration and torso on singing voice directivity*. The Journal of the Acoustical Society of America 148, 2020.
- [4] Manuel Brandner. *Progress Report PHD - Analysis of Quality and Efficiency in Classical Singing with Objective Parameters*.
- [5] Matthias Frank. *Phantom sources using multiple loudspeakers in the horizontal plane*. Ph.D. thesis, Institute of Electronic Music and Acoustics, University of Music and Performing Arts, Graz, 2013.  
<http://phaidra.kug.ac.at/o:7008> (last viewed April 2021)
- [6] Michael Möser. *Technische Akustik*. Springer-Verlag Berlin Heidelberg, 2015.
- [7] Stefan Weinzierl. *Handbuch der Audiotechnik*. Springer-Verlag Berlin Heidelberg, 2008, 35-39, 487-488.