Acta Acustica united with Acustica
Vol. 102 (2016) 918 – 929

DOI 10.3813/AAA.919006

Influence of Higher Order Acoustical Propagation Modes on Variable Section Waveguide Directivity: Application to Vowel [a]

R. Blandin[1)], A. Van Hirtum[1)], X. Pelorson[1)], R. Laboissière[2)]

[1)] GIPSA-Lab, UMR CNRS 5216, Grenoble University, 11 rue des Mathématiques (BP46), 38402 Grenoble, France. annemie.van-hirtum@gipsa-lab.grenoble-inp.fr

[2)] LPNC, UMR CNRS 5105, Université Pierre Mendès-France (BP47), 38040 Grenoble, France

Summary

The radiation of sound from the mouth of a speaker is often described using a plane piston assumption. This is satisfactory as long as the wavelength is longer than the largest transverse dimension of the vocal tract. For shorter wavelengths, higher order acoustical modes can propagate and the particle velocity field in the mouth exit plane can be nonuniform. As a consequence, the plane piston assumption does not hold and a more accurate description of the particle velocity distribution at the mouth exit is necessary. This can be achieved using the multimodal theory which allows one to account for higher order mode propagation. In order to study the influence of higher order modes on the sound radiated from the mouth, four approximations of the vocal tract shape for the vowel [a] are studied. The eccentricity of the junctions between the different portions of these geometries is used to control the propagation of a higher order mode. The radiated sound is simulated and compared with the measurements performed on mechanical replicas at various angles in the horizontal plane. It is observed that the propagation of a higher order mode can strongly modify the directivity. Varying the degree of eccentricity of the junction between the different sections shows that a very small eccentricity can induce the propagation of a higher order mode. Reducing the size of the open end shows that higher order modes can affect the radiated sound even if the radiating surface is small compared to the largest transverse dimension of the geometry.

PACS no. 43.70.Bk

1. Introduction

The directivity of the sound radiated by a speaker has been studied for various purposes, including microphone placement optimisation [1], telephony [2], vocal performance practice [3, 4, 5], experimental validation of acoustic theories [6], architectural acoustics [7], auralization and 3D sound synthesis [8, 9]. The observation on real subjects by several authors [1, 4, 3, 5] showed that if a speaker issuance is almost omnidirectional at low frequency, the directivity patterns become more and more complex at high frequency.

The variation of the amplitude and phase of the sound radiated by a speaker with the direction can be due to several physical phenomena including:

1. The diffraction of the radiated waves by the head of the speaker, which is expected to be significant for the wavelengths of the same order of magnitude as the head.
2. The reflections on the torso.
3. The presence of higher order acoustical modes (HOM) inside the vocal tract, which is expected to be significant for the wavelengths of the same order of magnitude as the transverse dimensions of the vocal tract.

The physical models and applications related to speech directivity are often limited to a frequency range of 0 kHz to about 5 kHz because, though the sound quality in not optimal, this is sufficient to ensure speech recognition and it allows one to perform simulations with a low computational cost. In this range, a plane piston assumption and accounting for the diffraction by the head and the reflections on the torso is sufficient to give a good description of the directivity of the radiated sound [6, 10] and the HOM can be neglected without loss of accuracy.

However the progress of the technology, the development of new applications such as wideband telephony or augmentative hearing, the increase of quality requirements for 3D sound synthesis and the latest findings in speech perception [11] leads one to focus more on frequencies beyond 5 kHz. In this frequency range, the wavelength becomes of the same order of magnitude as the transverse dimensions of the vocal tract. Thus, HOM propagation is expected and has been observed on mechanical replicas [12, 13]. As a consequence, the particle velocity

Received 11 February 2016,
accepted 27 July 2016.

distribution on the mouth exit plane can become nonuniform and the directivity of the radiated sound can be affected.

The aim of this work is to understand how the physical phenomenon of HOM propagation can potentially affect the speech directivity using a theoretical model validated against experiments.

In order to take into account the HOM in the computation of the radiated sound, a more precise description of the particle velocity distribution on the mouth exit plane is needed. This can be achieved with any simulation method which allows one to compute the 3D acoustic field inside a variable section waveguide. However, the multimodal (MM) theory [14, 15, 16, 17], with a lower computational cost compared to other methods (such as finite elements or finite differences), has the advantage of providing directly all the information related to the propagation modes (eigen functions, cut-on frequencies[1], mode amplitudes and coupling terms). This is a valuable aid to understand the physical phenomena occurring inside the waveguide, and their potential influence on the radiated sound.

Though a realistic vocal tract geometry would be interesting to study, it has the disadvantage to be subject to inter- and intra-subject variations and has a complex shape which makes the understanding of the HOM propagation occurring inside more difficult. Thus, it has been chosen to study simplified geometries created from two approximations of the vocal tract shape for vowel [a] which allows one to identify clearly the HOM involved. The first approximation is constituted of two cylindrical tubes [18] and the second one of 44 cylindrical tubes [19]. Knowing that, for a circular cross-section shape, the excitation of the first HOM depends on the eccentricity of the junction between the cylinders, the use of concentric and eccentric junctions for both approximations has been chosen in order to control the propagation of this HOM. The directivity patterns have been experimentally studied and compared with the outcomes of the MM simulations for frequencies up to 10 kHz. In particular the impact of the eccentricity and the exit diameter on the directivity has been investigated.

The paper is structured as follows: some aspects of the MM theory and its implementation important for this study are detailed in section 2, then the experimental setup is described in section 3 and the simulation and measurement results are presented and discussed in section 4.

2. Outline of the multimodal acoustic theory

The MM acoustic theory used to perform the simulations is detailed in [13]. The implementation used for this work is limited to discrete straight waveguides with arbitrary cross-sections set in an infinite baffle. Since the geometries used for this work are described with piecewise constant cross-section, it is the more adapted method for this study. However, a continuous approach allowing the curvature to be taken into account [20] could be useful to simulate more accurate approximations of the vocal tract shape in the future.

According to the MM theory, the pressure field inside a constant cross-section waveguide can be described by the expression

$$p(x_1, x_2, x_3, t) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \psi_{mn}(x_1, x_2) P_{mn}(x_3)\, \mathrm{e}^{\mathrm{j}\omega t}, \qquad (1)$$

where (x_1, x_2, x_3) is a generalized coordinate system, x_3 is the propagation direction, $P_{mn}(x_3)$ is the modal amplitude, $\omega = 2\pi f$ is the angular frequency, and t is time.

The functions ψ_{mn} are the so called propagation modes. In the case of a circular cross-section of radius R, they can be expressed using polar coordinates (r, θ) in the plane (x_1, x_2) as a product of Bessel functions of the first kind and sine and cosine functions

$$\psi_{mn}(r, \theta) = \frac{J_n(r\gamma_{mn}/R)}{K_{mn}} \begin{cases} \sin(n\theta) \\ \cos(n\theta) \end{cases}, \qquad (2)$$

where K_{mn} is a normalization factor, γ_{mn} is the order m zero of the first derivative of the Bessel function J_n. The cut-on frequency f_c of each mode ψ_{mn} is given by

$$f_c = \frac{c\gamma_{mn}}{2\pi R}, \qquad (3)$$

where c is the speed of sound. In the case of a more complex cross-sectional shape, the functions ψ_{mn} can be estimated using a numerical method such as finite differences.

The coupling of the modes ψ_{mn} at a discontinuity in the cross-section is described with the matrix $\boldsymbol{F}$ which links the modal amplitude on both sides, a and b, of the junction

$$\boldsymbol{P_a} = \boldsymbol{F}\boldsymbol{P_b} \qquad (4)$$

where $\boldsymbol{P_a}$ and $\boldsymbol{P_b}$ are two vectors containing the modal amplitudes P_{mna} and P_{mnb} corresponding to each side of the discontinuity. This equation is true only if the cross-sectional surface of side a is smaller than the one of side b and lies inside it. The small tube can be placed either on the right or left side so that convergent and divergent shapes can be handled. Even if this case does not appear in the following, it is worth noticing that the problem of a junction whose cross-sections are not enclosed in each other can be solved by inserting an intermediate zero-length waveguide portion whose cross-section is the intersection of the original two neighboring cross-sections.

The terms of the matrix $\boldsymbol{F}$ can be evaluated with the expression

$$F_{mn_a,mn_b} = \frac{1}{S_a} \int_{S_a} \psi^*_{mn_a} \psi_{mn_b}\, \mathrm{d}S, \qquad (5)$$

with S_a the smallest cross-sectional surface.

[1] The term cut-on frequency refers to the fact that the HOM begin to propagate above their cut-on frequency, it is the same as the cutoff frequency which refers to the fact that the HOM stop propagating under their cutoff frequency.

An analytical expression can be found in the particular case of two circular cross-sections sharing the same central axis. However, as long as it is not the case, finding an analytical expression is very difficult and numerical integration has been used. On the other hand, this allows one to handle the case of the junction between two waveguide portions having different cross-sectional shape.

A grid of N_p points regularly spaced is generated on S_a and the value of ψ_{mn_a} and ψ_{mn_b} are evaluated at each point of the grid. The coupling coefficient can then be evaluated by computing the sum

$$F_{mn_a,mn_b} = \frac{1}{N_p} \sum_{i=1}^{N_p} \psi^*_{mn_a}(x_{1i}, x_{2i}) \psi_{mn_b}(x_{1i}, x_{2i}). \quad (6)$$

For the simulations performed in this study, the spacing between the points has been set to 0.5 mm (for a diameter of 30 mm this yields $N_p \approx 650$). This resolution has shown a good agreement with experimental data [13], and convergence tests confirmed the reliability of the simulations[2].

The radiated pressure can be computed with the Rayleigh-Sommerfeld integral

$$P(x_1, x_2, x_3, k) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{\mathrm{j}\omega\rho}{2\pi S} \int_S U_{mn} \psi_{mn}(x'_1, x'_2) \frac{\mathrm{e}^{\mathrm{j}kh}}{h} \mathrm{d}S, \quad (7)$$

with U_{mn} the modal amplitude of the particle velocity, $h = \sqrt{(x_1 - x'_1)^2 + (x_2 - x'_2)^2 + (x_3 - x'_3)^2}$, S being the surface at the open end of the waveguide and (x'_1, x'_2, x'_3) the coordinates of the points on S. This integral can be approximated considering a finite summation over $(M + 1) \times (N + 1)$ modes and replacing the integral by a finite summation evaluated on a grid of N_p points

$$P(x_1, x_2, x_3, k) = \sum_{m=0}^{M} \sum_{n=0}^{N} \frac{j\omega\rho}{2\pi N_p} \sum_{i=1}^{N_p} U_{mn} \psi_{mn}(x'_{1i}, x'_{2i}) \frac{\mathrm{e}^{\mathrm{j}kh_i}}{h_i}, \quad (8)$$

with

$$h_i = \sqrt{(x_1 - x'_{1i})^2 + (x_2 - x'_{2i})^2 + (x_3 - x'_{3i})^2},$$

$(x'_{1i}, x'_{2i}, x'_{3i})$ being the coordinates of the points of the grid. Note that one does not divide by the area S, since the area elements dS have not been taken into account in the summation above. On the other hand, one must divide by the number of points N_p. This expression does not allow one to compute the acoustic pressure at the exact location of a point of the grid: in this case $h_i = 0$ and a division by zero is introduced. However, it is possible to compute the pressure at these locations with Equation (1).

[2] In the range 0 kHz to 10 kHz the mean amplitude difference between a simulation performed with a resolution of 0.5 mm and a simulation performed with a resolution of 0.3 mm is lower than 0.7 dB. The mean phase difference is lower than 0.2 rad.

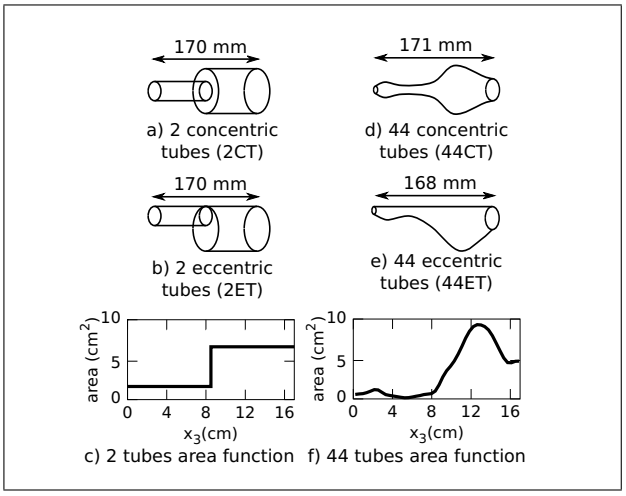

Figure 1. Vocal tract geometries approximating vowel [ɑ], the glottis is located at $x_3 = 0$.

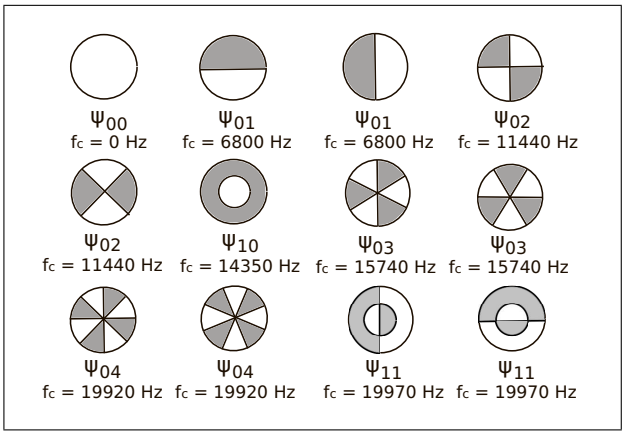

Figure 2. Mode-shapes and cut-on frequencies (f_c) of the propagation modes of a circular cross-section whose cut-on frequency is below 20 kHz for a diameter of 29.5 mm and a temperature of 26.5 °C. The lines are zero amplitude nodal lines and the colour indicates the phase change.

An experimental setup (Figure 3) has been designed to measure the directivity patterns of different vocal tract replicas (Figure 1) in order to quantify the directivity patterns and to validate the above theory.

3. Measurement method

3.1. Replicas

For this study four vocal tract replicas which can be considered as a vowel [ɑ] vocal tract approximation [18, 19] have been used:

1. Two replicas composed of two cylindrical tubes (85 mm long with diameters 14.5 mm and 29.5 mm) [18] sharing either the same central axis (Figure 1a) or a common line on the edge (Figure 1b) which corresponds to a distance between the centers of 7 mm. In what follows, they are referred to as 2CT (Two Concentric Tubes) and 2ET (Two Eccentric Tubes) respectively. For both replicas the cross-section area function is depicted in Figure 1c. The lowest cut-on frequency based on the exit diameter yields 6.8 kHz (at 26.5 °C).

2. Two replicas with 44 circular cross-sections whose area function is taken from [19] with all sections sharing either the same central axis (Figure 1d) or a common line on the edge (Figure 1e). In what follows, they are referred to as 44CT (44 Concentric Tubes) and 44ET (44 Eccentric Tubes) respectively. For both replicas the cross-section area function is depicted in Figure 1f. The lowest cut-on frequency yields 5.9 kHz (at 26.5 °C) based on the maximal diameter.

The multimodal approach outlined in section 2 is applied to the described geometries. As an example, the mode-shapes and the cut-on frequencies of the propagation modes corresponding to the largest tube of the two tube geometry are presented in Figure 2 up to 20 kHz. The modes ψ_{01}, ψ_{02}, ψ_{03}, ψ_{04} and ψ_{11} are degenerate: they have two mode-shapes similar by rotation of a certain angle having the same cut-on frequency. Note that the first subscript index corresponds to the radial dimension and the second to the angular dimension in the cross-section plane (x_1, x_2).

3.2. Experimental setup

A schematic diagram of the experimental setup is presented in Figure 3. A compression driver (Eminence PSD:2002S-8) is used to generate a linear sweep signal with a frequency varying from 2 kHz to 10 kHz, which is transmitted to the replicas through an adaptation part which features a 2 mm diameter hole. The compression driver and the adaptation part are rotationally symmetric with respect to the center of the entrance of the replicas.

The open end of the replica is set inside an insulated room (1.92 x 1.95 x 1.99 m^3, Vol = 7.45 m^3). The replica is equidistant from the wall of the room in the plane (x_1,x_3). Even though this room cannot be considered as perfectly anechoic, its acoustic performances in the frequency range of interest in this study (2 kHz to 10 kHz) are sufficient so that the influence of the external noise can be neglected and the free field assumption holds: the direct field is higher than the reverberated field up to 0.94 m from the open end of the replica and the attenuation of the external noise is greater than 25 dB SPL [21].

A baffle of dimensions (365 x 360 mm^2) is attached to the outlet of the replica in order to mimic a human face. It corresponds to the flanged boundary condition of the theory. The sound source is placed outside of this room to avoid interferences.

The radiated sound is measured at 48 cm from the exit of the replica at 13 different angles from the axis of the replicas (every 15° from −90° up to 90° as indicated in Figure 3) using a microphone (B&K 2669 L). For the eccentric geometries the common edge corresponds to the angle 90°.

The output of the microphone and the excitation signal are amplified respectively with a microphone conditioner (B&K 5935 L) and an amplifier (Onkyo a-807). The generation of the excitation signal and the recording of the signals from the microphone and the input of the compression driver are managed by a data acquisition card (NI PCI-MIO 16 XE). The whole measurement process is controlled with Labview.

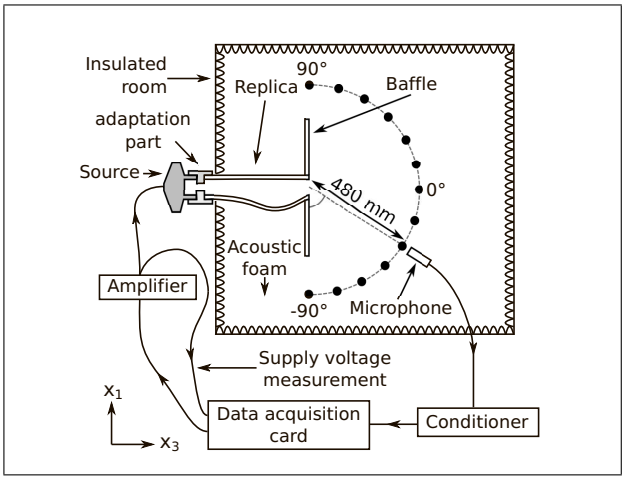

Figure 3. Experimental setup used to measure the directivity patterns of vocal tract replicas. The orientation of the eccentric part of the replica corresponds to the convention used for the experiment and the simulations.

In order to enhance the change of pressure amplitude with respect to the angular position, unless stated differently, the amplitude at the different positions has been normalized by the amplitude at the central positions (which corresponds to the point located in front of the replica at an angle of 0°). The fact that the sound source amplitude depends on the frequency could disturb the interpretation of the measurements and their comparison with the simulations. However, its amplitude being the same for all the positions, dividing by the amplitude at one position cancels its influence, and the results can be compared with the simulations.

To get a more general view of the directivity effects, the maximal sound pressure level difference (MSPLD) between the different positions has been computed for all the frequencies. To do so, for each frequency, the minimum of the sound pressure level (in dB SPL) has been subtracted from the maximum with respect to the angular position.

On the experimental data the amplitude at −90° and 90° is lower for every measured frequency. Since it is not the case for the simulated data and that this happens even at low frequencies, this can be attributed to a measurement artifact probably related to the fact that the microphone was close to the wall of the insulated room at these positions. To avoid the perturbations of this artifact in the analysis, the MSPLD without the measurements of the edges (at −90° and 90°). However the observed overall tendency remains similar.

4. Results and discussions

In section 4.1 the measured and simulated data are described, and an analysis is provided in section 4.2.

4.1. Results

The measured and simulated normalized pressures and the MSPLD are presented in Figures 4 and 5 for respectively

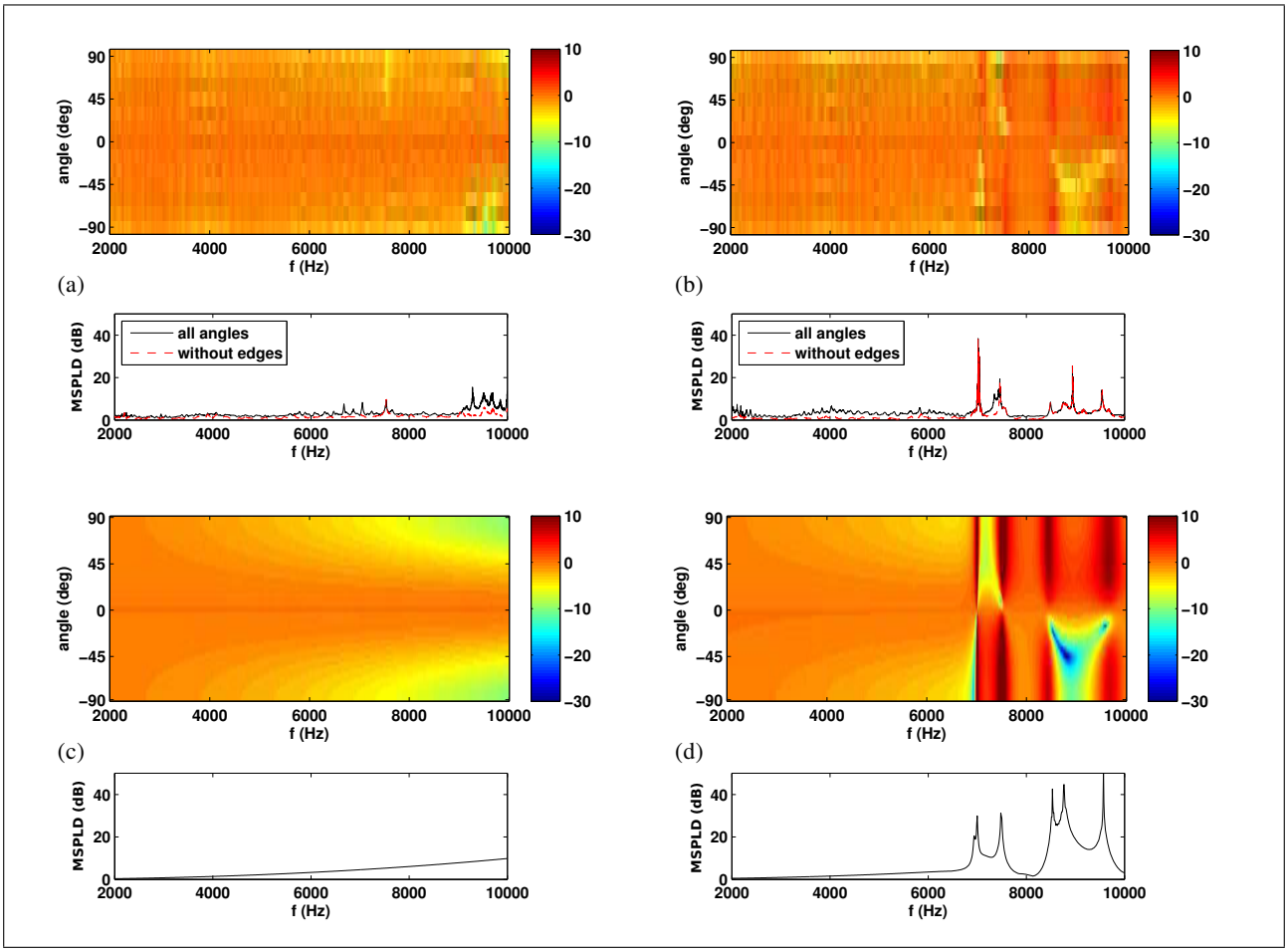

Figure 4. (Colour online) Normalized amplitude (dB) and maximal sound pressure level difference MSPLD with respect to the angular position of the pressure radiated from two tubes vowel [ɑ] replicas with concentric and eccentric junction measured and simulated at 48 cm from the exit between 2 kHz and 10 kHz every 15 ° and 3 ° for respectively the experiments and the simulations. To avoid perturbations linked to experimental artifacts present on the extremal positions, the maximal sound pressure level difference has been computed with and without the edge positions (−90° and 90°). a) 2 concentric tubes (2CT) experiment, b) 2 eccentric tubes (2ET) experiment, c) 2 concentric tubes (2CT) MM simulation, d) 2 eccentric tubes (2ET) MM simulation.

the two and 44 tubes geometries. Directivity patterns obtained for selected frequencies are presented in Figure 6.

4.1.1. Directivity of concentric configurations (2CT and 44CT)

The concentric configurations generate a one lobe symmetric directivity pattern with a maximum in the center (at 0°, see Figures 4a, 4c, 5a and 5c). The amplitude of this lobe increases in the center and decreases towards the edges (−90° and 90°) when the frequency is increased. The MSPLD increases progressively and reaches 10 dB and 7 dB at 10 kHz for the 2CT and 44CT respectively. In the experimental data, this pattern is perturbed by measurement noise which can be due to insufficiently damped reflections on the walls of the insulated room, or external noise.

In addition to this noise, one can see more important localized discrepancies between the experiments and the simulations. For the 2CT, a peak of MSPLD of 10 dB can be seen at 7.52 kHz. The corresponding directivity pattern presented in Figure 6a is asymmetric with a lower amplitude between 0°and 90°. Other less pronounced discrepancies can be found at 7 kHz and 9.3 kHz. On the experimental data obtained with the 44CT, peaks of MSPLD can be seen at 8.1 kHz (36 dB), 8.3 kHz (34 dB), 8.8 kHz (14 dB) and 9.1 kHz (57 dB). Inside a narrow frequency band centered on these frequencies, the directivity patterns is constituted of two asymmetric lobes separated by a low amplitude direction (see Figure 6c). These discrepancies can be related to the differences observed between the measured and simulated internal pressure field observed in [13] (see Figure 8 and 11 of [13]). It can be hypothesized that they are generated by small asymmetries of the replicas, this is looked into later. Another explanation could be that a small asymmetry in the placement of the replica inside the insulated room would induce an asymmetry of the reflections. However, the direct field is higher than the reverberated field at 48 cm from the end of the replica, and the effect of reflections would be expected in the whole frequency range. Apart from these particular narrow frequency bands, the same patterns are observed in the experimental and the simulated data.

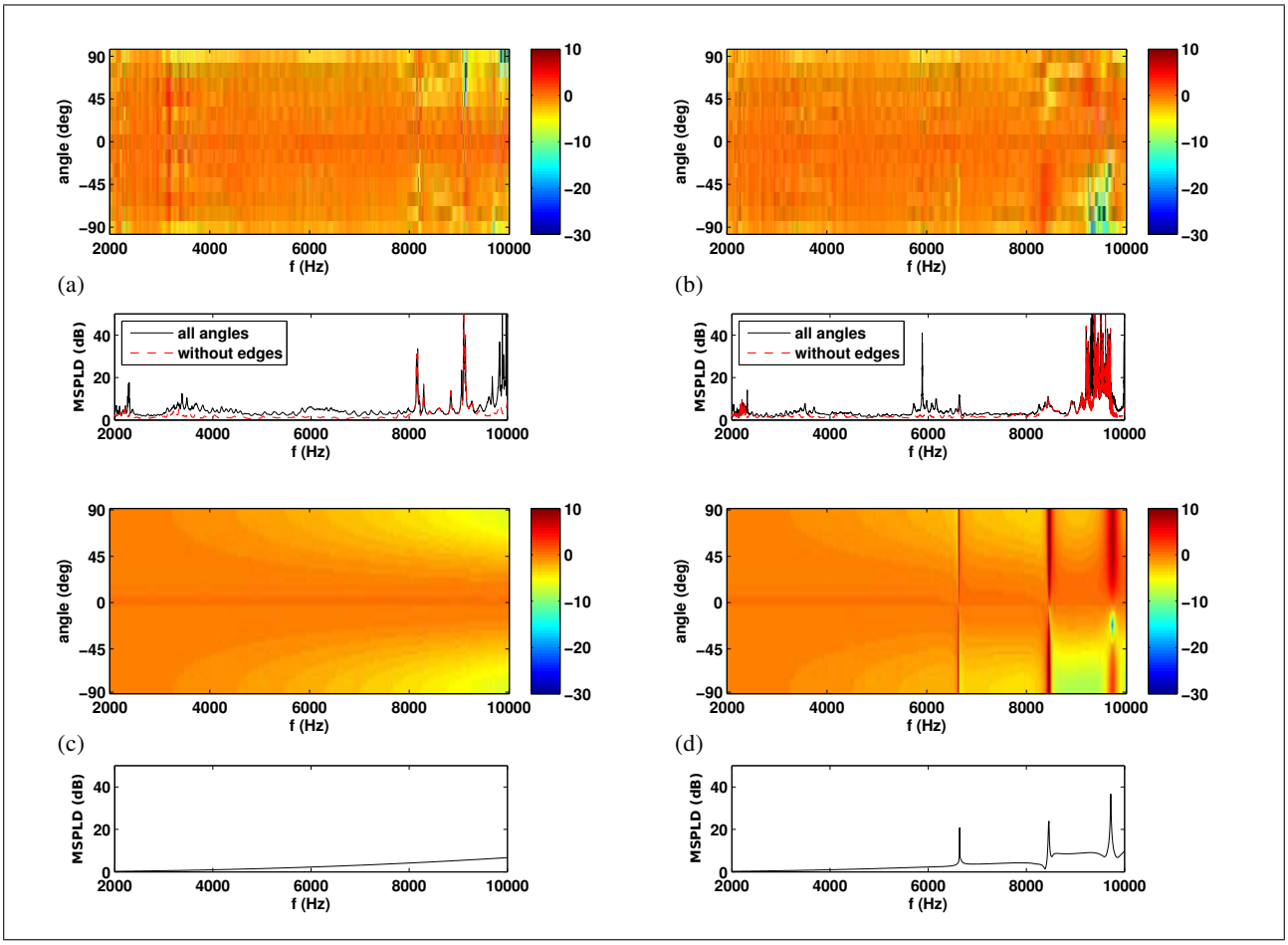

Figure 5. (Colour online) Normalized amplitude (dB) and maximal sound pressure level difference MSPLD with respect to the angular position of the pressure radiated from 44 tubes vowel [ɑ] replicas with concentric and eccentric junction measured and simulated at 48 cm from the exit between 2 kHz and 10 kHz every 15 ° and 3 ° for respectively the experiments and the simulations. To avoid perturbations linked to experimental artifacts present on the extremal positions, the maximal sound pressure level difference has been computed with and without the edge positions (−90° and 90°). a) 44 concentric tubes (44CT) experiment, b) 44 eccentric tubes (44ET) experiment, c) 44 concentric tubes (44CT) MM simulation, d) 44 eccentric tubes (44ET) MM simulation.

4.1.2. Directivity of eccentric configurations (2ET and 44ET)

For the eccentric geometries, a one lobe pattern which becomes more pronounced for increased frequency can be seen up to 6.5 kHz. Above this limit, more complex asymmetric patterns with one or two lobes are visible (see Figures 4b, 4d, 5b, 5d, 6b and 6d). The variations of the directivity are important within short frequency intervals: the MSPLD can vary up to 50 dB within 100 Hz intervals.

In order to see better the appearance and the evolution of the two lobe pattern, the angle corresponding to the minimum between the lobes has been plotted in Figure 7. This minimum is typically of the order of 30 dB lower than the maximum and is limited to a narrow angular region of about 30°. Its direction can vary a lot (up to 60°) within a short frequency interval (of the order of 100 Hz). For both 2ET and 44ET, the minimum of amplitudes tends to be located between -90° and 0°, which corresponds to the side opposite to the common edge.

In the case of the 2ET, the patterns obtained by simulation can be recognized in the experimental data. For both experiment and simulation, peaks of MSPLD can be seen at 7 kHz, 7.5 kHz, 8.5 kHz, 8.8 kHz and 9.6 kHz. Except between 7.9 kHz and 8.1 kHz, two lobes can be seen at all the frequencies from 6.94 kHz (see Figures 6b and 7a).

In the case of the 44ET, similar patterns can be recognized in the experiment and the simulation, but there are more differences above 9 kHz. Peaks of MSPLD can be seen at 6.6 kHz, 8.5 kHz and 9.7 kHz (see Figure 6d) for both experiment and simulation. A two lobes pattern appears inside more reduced and isolated frequency intervals: 6.63 kHz to 6.65 kHz, 8.38 kHz to 8.53 kHz and 9.58 kHz to 9.89 kHz (see Figure 7b). Between 6.6 kHz and 8.5 kHz a one lobe symmetric pattern can be seen and between 8.5 kHz and 9.7 kHz, there is a one lobe asymmetric pattern with lower amplitude between 0°and 90°.

4.1.3. Input impedances

In order to characterize more globally the studied vocal tract geometries, the input impedance Z has been computed as the ratio of the pressure over the particle velocity

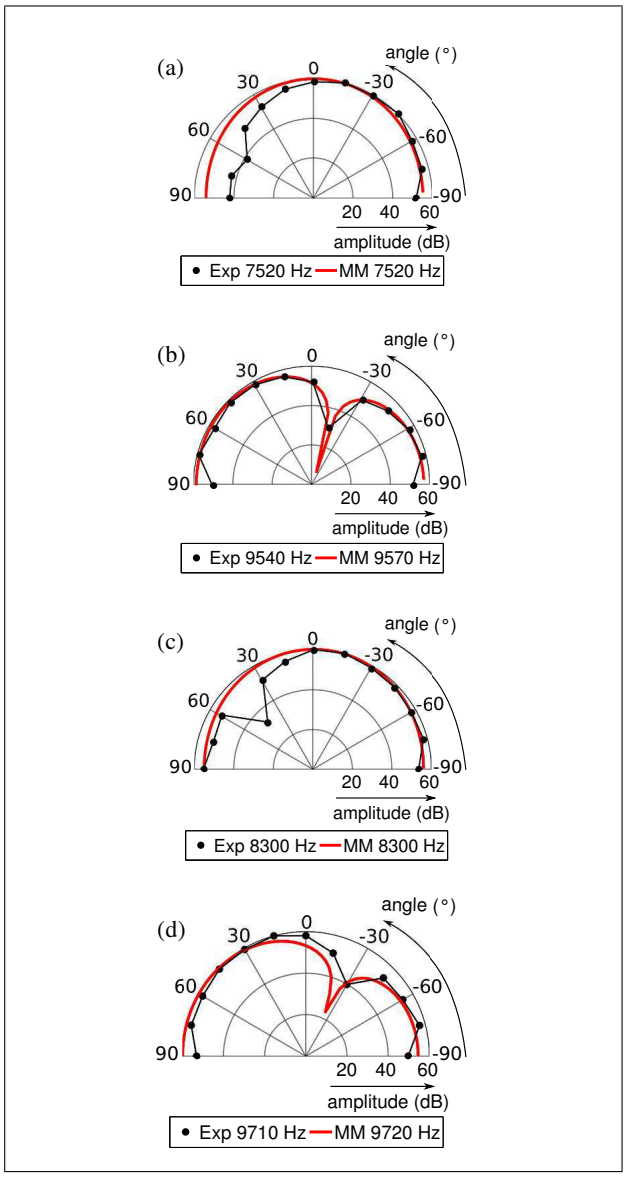

Figure 6. (Colour online) Measured (dotted lines) and simulated (full lines) acoustic pressure amplitude (dB) as a function of angle at selected frequencies. a) 2 concentric tubes (2CT), b) 2 eccentric tubes (2ET), c) 44 concentric (44CT), d) 44 eccentric tubes (44ET).

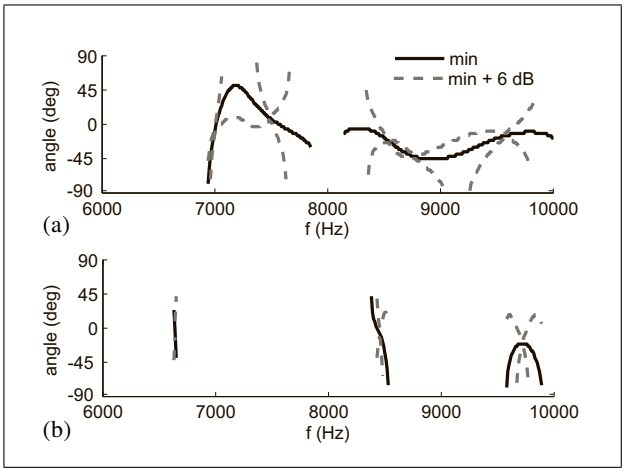

Figure 7. Angle corresponding to the minimum (for maximal sound pressure level difference MSPLD values greater than 3 dB) of the radiated pressure as a function of the frequency extracted from the simulations of the eccentric geometries. The angles corresponding to the pressure 6 dB higher than the minimum are presented in dashed lines. a) two eccentric tubes (2ET), b) 44 eccentric tubes (44ET).

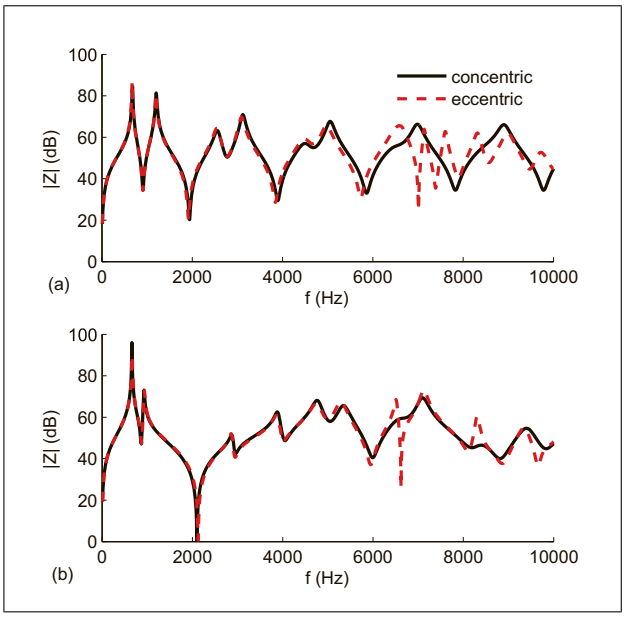

Figure 8. (Colour online) Modulus of the input impedance Z corresponding to four vocal tract approximations based on [18] and [19] with either concentric or eccentric junctions between the different sections. a) 2 tubes geometries (2CT and 2ET), b) 44 tubes geomeries (44CT and 44ET).

at the center of the communication hole. The amplitude of Z is presented for the four studied geometries in Figure 8.

The peaks of input impedance correspond to resonances inside the geometries. The two first peaks correspond to the two first formants, which characterize the vowels. In the case of the two tube geometries (see Figure 8a), their frequency are 670 Hz and 1200 Hz, and in the case of the 44 tube geometries (see Figure 8b) they are 660 Hz and 930 Hz (similar values are given in [19]). These values do not depend on the junction type.

Up to 6.5 kHz the input impedance curves of the concentric and eccentric configurations are almost exactly similar. Above this frequency, the variations of Z are more complex for the eccentric configurations. The input impedance of the 2ET is almost completely different from the one of the 2CT, and the input impedance of the 44ET is different from the one of the 44CT inside three frequency intervals (6.2 kHz to 7 kHz, 7.9 kHz to 8.6 kHz and 9.3 kHz to 9.8 kHz). Additional maxima can be seen at 6.6 kHz, 7.1 khz, 7.6 kHz, 8.3 kHz 9 kHz and 9.7 kHz for the 2ET and at 6.5 kHz and 8.3 kHz for the 44ET. There are also additional minima at 7 kHz, 7.4 kHz, 7.9 kHz, 8.6 kHz and 9.5 kHz for the 2ET and at 6.6 kHz, 8.1 kHz, 8.9 kHz and 9.7 kHz for the 44ET.

4.1.4. Influence of the degree of eccentricity

In order to investigate the influence of the degree of eccentricity of a junction on the directivity, simulations have been performed varying the distance between the centers of the two cross-sections of the 2 tubes geometry from 0 mm to 7 mm by steps of 0.5 mm. This corresponds to the progressive transition from the 2CT (Figure 1a) to the 2ET (Figure 1b). The radiated pressure has been computed for each distance between the centers. The MSPLD corresponding to four distances between the centers is presented in Figure 9. These distances, 0 mm, 0.5 mm, 3.5 mm and 7 mm, corresponds respectively to the concentric configuration, the lowest possible eccentricity with the discretization of the cross-sectional surface used (0.5 mm), the intermediate configuration and the fully eccentric configuration. At 0.5 mm of distance, peaks of MSPLD up to 20 dB can be seen above 6.5 kHz and the MSPLD curve is completely separated from the 0 mm curve above 8 kHz. With a distance of 3.5 mm the MSPLD curve is very similar to the 7 mm curve.

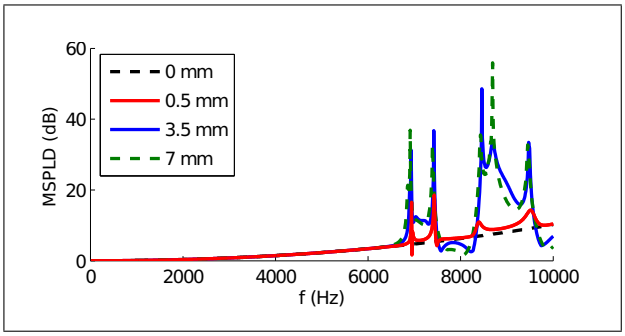

Figure 9. (Colour online) Maximal sound pressure level difference MSPLD as a function of the frequency computed for different spacing of the centers of both cross-sections of the two tubes geometries.

4.1.5. Convergent exit

During speech production, the vocal tract often has a cavity located before the mouth with larger transverse dimensions than the mouth exit, and thus has a convergent exit. This is what one can see in the area function of the 44 tubes approximations which decreases near the open end of the vocal tract (see Figure 1f).

However, this is not the case for the two tubes approximation. In order to investigate how a convergent exit can affect the directivity, a third section, of zero length, has been added at the end of the 2ET geometry. Its diameter has been reduced by steps of 0.5 mm from 29.5 mm (which corresponds to the diameter of the second section) to 10 mm. A convergent exit is thus gradually introduced. The radiated pressure has been computed for each diameter value. The MSPLD and the input impedance corresponding to 3 diameters values, 29.5 mm, 20 mm and 10 mm, are presented in Figures 10a and 10b respectively. The first one corresponds to the 2ET without convergent exit.

When the diameter of the third section is reduced, the MSPLD is globally decreasing. With a diameter of 10 mm it is almost zero up to 6.5 kHz. However, above 6.5 kHz the MSPLD still has locally important values (up to about 35 dB at 7 kHz, 7.3 kHz, 8.1 kHz and 9.3 kHz for an exit diameter of 10 mm), but between these peaks the MSPLD is reduced. The frequency of the peaks is decreased and one peak disappears at 10 mm. The peaks of input impedance have a greater amplitude and a reduced bandwidth and the minima have a reduced amplitude. The frequency of the peaks and minima of input impedance is also reduced.

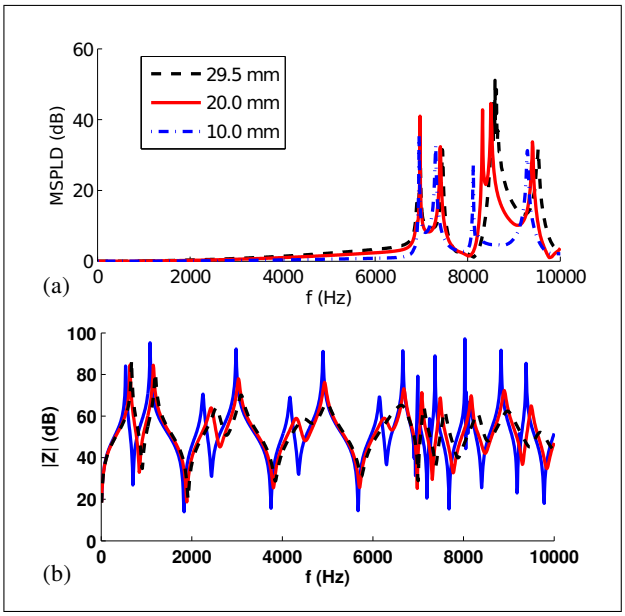

Figure 10. (Colour online) Maximal sound pressure level difference MSPLD and input impedance as a function of the frequency computed for different diameters of the exit of the two eccentric tubes geometry. a) Maximal sound pressure level difference, b) Input impedance.

4.1.6. Spectrum of the sound radiated by the 2ET

In order to get an idea of how the HOM effect can be perceived, the 2ET replica has been excited using a broadband noise. This can be considered as an imitation of whispered voice. The spectrum of the radiated sound is presented for each measurement position in Figure 11. Up to 6.5 kHz it is very similar for each position, except at the edges (-90° and 90°) where the amplitude is globally lower than at the other positions (this has the same origin as the artifact mentioned at the end of section 3.2). Above 6.5 kHz there is noticeable differences between the positions. As an example, at -75°, -60°and -45°the amplitude is lower than at the other positions between 8.5 kHz and 9.5 kHz. The pattern observed on the directivity maps of Figures 4b and 4d can be recognized in the variations of the spectrum with the position. The authors perceived an audible difference between the positions when listening to the recorded sounds.

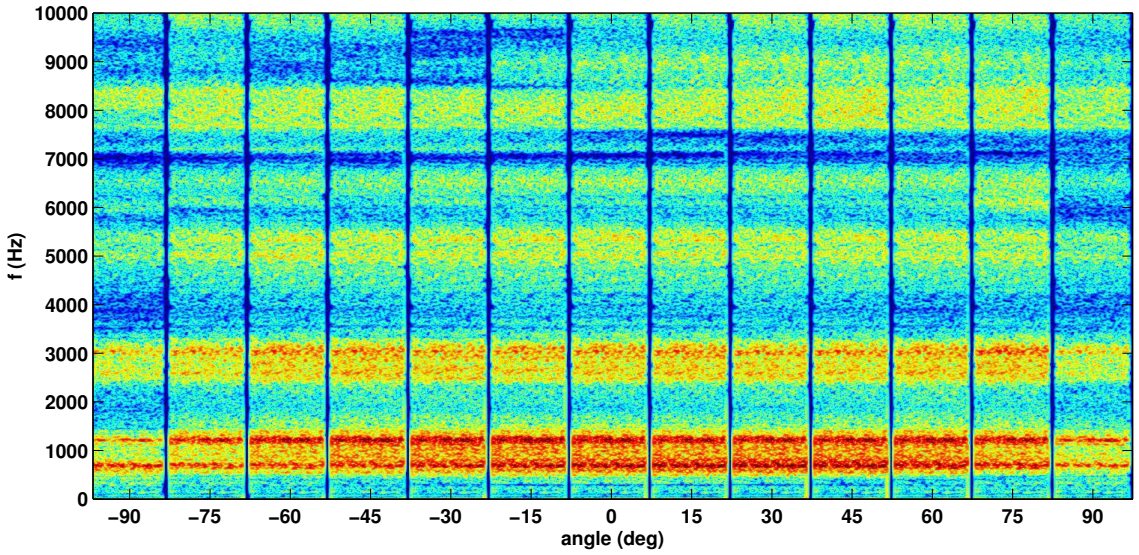

Figure 11. (Colour online) Spectrogram of the sound radiated by a vocal tract replica consisting in a two tube approximation with an eccentric junction (2ET) recorded at a distance of 48 cm from the exit every 15°. The excitation signal used is a broadband noise. Hanning windows of 46.4 ms (2048 samples) have been used with an overlap rate of 0.9.

4.2. Discussion

4.2.1. Propagation modes inside the geometries

Inside the small tube of the two tubes geometries the cut-on frequency of the first HOM (the only one propagating under 20 kHz) is 14.5 kHz. The discontinuity corresponding to the transition from the communication hole to the replica is likely to generate evanescent HOM[3] at the entrance of this tube. However, the length of this tube being more than five times greater than its diameter, the influence of evanescent HOM can be neglected at the junction between the small tube and the larger one. Thus, between 0 kHz and 10 kHz, only plane waves (mode ψ_{00}) arrive at this junction. Inside the larger tube, ψ_{01} which has a cut-on frequency of 6.8 kHz (see Figure 2), is the only HOM which can propagate under 10 kHz. The 44 tubes geometries have the same kind of shape as the two tubes ones: a narrow part (from $x_3 = 0$ mm to $x_3 = 80$ mm, see Figure 1) is followed by a wider one (from $x_3 = 80$ mm to $x_3 = 167$ mm). The same arguments as for the two tubes geometries lead to the conclusion that the only modes propagating inside these geometries under 10 kHz are ψ_{00} and ψ_{01}.

4.2.2. Coupling between ψ_{00} and ψ_{01}

The exit of the geometries being perfectly axisymmetric, the HOM ψ_{01} can be excited only if the junctions are asymmetric. Indeed, in Eq. (5) the contributions of ψ_{00} on either side of the nodal line of ψ_{01} compensate each other exactly if the junction is perfectly axisymmetric. In this case, the coupling term $F_{00,01}$ between ψ_{00} and ψ_{01} is zero and ψ_{01} cannot be excited. If an asymmetry is introduced, $F_{00,01}$ becomes to be non zero and ψ_{01} can be excited. When the eccentricity is increased $F_{00,01}$ increases. Thus ψ_{01} is not expected to be present in the axisymmetric configurations (2CT and 44CT).

4.2.3. Particle velocity distribution at the exit

In the case of the concentric configurations, the only propagation modes expected at the exit of the geometries are ψ_{00} and evanescent HOM which are generated by the discontinuity corresponding to the transition from the geometries to the exterior space. In the case of the eccentric configurations, ψ_{01} can be present and introduce the appearance of two areas with opposite phase on the exit surface.

The evanescent HOM can change the particle velocity distribution on the exit surface, and one could expect an influence on the directivity of the radiated sound. However, in the concentric configurations, a simulation performed with all modes whose cut-on frequency lies below 20 kHz and a simulation with plane mode only give exactly the same directivity patterns at 0.48 m from the exit. Thus, the influence of these modes can be neglected at this distance and the particle velocity distribution can be considered as uniform if ψ_{00} is the only mode propagating.

4.2.4. Directivity of concentric configurations (2CT and 44CT)

For the concentric configurations, the only propagation mode involved in the radiation is ψ_{00}. Thus, the only directivity effect which can be expected is the well known radiation from a baffled circular piston ([22] p226-227). Indeed, except for some localized discrepancies in the experimental data, the directivity pattern of these configurations is one symmetric lobe with higher amplitude in the center (at 0°) whose MSPLD increases progressively with the

[3] An HOM is evanescent when it is excited under its cut-on frequency. It is exponentially damped along the propagation axis from the excitation point.

frequency. This pattern is only due to the difference in the distance traveled by the waves coming from the different parts of the radiating surface to the reception point which induces a phase difference.

4.2.5. Directivity of eccentric configurations (2ET and 44ET)

In the case of the eccentric configurations, the baffled circular piston pattern can be seen under the cut-on frequency of ψ_{01} because up to this frequency ψ_{00} is the only mode propagating. Above this cut-on frequency the particle velocity distribution on the exit surface is no longer uniform. The phase difference between the waves coming from the different parts of the radiating surface is not only due to the differences in traveled distance but also to the amplitude and phase variations on the radiating surface. As a consequence, more complex directivity patterns are observed.

4.2.6. Two lobes pattern

The presence of ψ_{01} at the exit of the geometries can induce the appearance of two areas with opposite phase. In this case, the radiating surface acts as an acoustic dipole and a directivity pattern with two lobes separated by a low amplitude direction is generated.

The interference of the acoustic pressure radiated by ψ_{00} and ψ_{01} can reinforce the radiated acoustic pressure on one side and decreases it on the other. Asymmetric directivity patterns are thus obtained. The low amplitude direction is no longer orientated to the center (0°), but is shifted to the side having the lowest amplitude. The variations of the amplitudes and phase of ψ_{00} and ψ_{01} with the frequency changes the direction of the minimum of amplitude.

The propagation of a HOM other than ψ_{01} would generate different directivity patterns. As an example, a HOM having two parallel nodal lines would generate three lobes separated by two low amplitude directions. Thus other cross-sectional shapes (such as elliptical) would generate different directivity patterns.

4.2.7. Comparison between HOM and plane piston

At some frequencies the combination of ψ_{00} and ψ_{01} can result in a particle velocity distribution which compensates partially the phase difference due to the differences in traveled distances between the different parts of the radiating surface and the reception point. This generates a reduced directivity compared to the baffled circular piston at the same frequency. This can be seen between the second and the third peaks of MSPLD and after the fifth peak of MSPLD for the 2ET (see Figures 4b and 4d). For the 44ET it can be seen before or after the MSPLD peaks (see Figures5b and 5d). The effect of HOM is slightly noticeable below the cut-on frequency, from about 6.4 kHz, whereas the cut-on frequency of ψ_{01} is 6.8 kHz for the two tubes geometries. This can be explained by the fact that close to the cut-on frequency the reduction of amplitude of the evanescent modes generated at the junction becomes less important and they can have a significant amplitude at the exit.

4.2.8. Relation between input impedance and directivity

The peaks of MSPLD appear close to input impedance minima. For the 2ET the peaks of MSPLD at 7 kHz, 7.5 kHz, 8.5 kHz and 9.6 kHz can be related to minima of impedance located at 7 kHz, 7.4 kHz, 8.6 kHz and 9.5 kHz. Likewise for the 44ET the peaks of MSPLD at 6.6 kHz and 9.7 kHz can be related to minima of input impedance located at close frequencies. However, the frequencies are not exactly the same and some peaks of MSPLD cannot be related to maxima or minima of input impedance. As an example, for the 2ET, the peak of MSPLD at 8.8 kHz does not correspond to a maximum or minimum of input impedance. Likewise, for the 44ET, the peak of MSPLD at 8.5 kHz can not be related to a maximum or minimum of input impedance. One can also notice that for the 2ET, the peak of MSPLD located at 7.5 kHz is between a minimum (at 7.4 kHz) and a maximum (at 7.6 kHz) of input impedance. Thus, it is not possible to find a simple relation between the input impedance curve and the directivity phenomenon.

4.2.9. Influence of eccentricity degree

The simulations performed varying the eccentricity showed that even a very small eccentricity (0.5 mm) can induce significant changes of the directivity patterns with respect to the axisymmetric configuration (up to 15 dB of difference). Thus, it is not required for there to be a large eccentricity for ψ_{01} to be excited and propagate. This is in agreement with the assumption that the localized discrepancies between simulations and experiments can be explained by small asymmetries in the replicas. The fact that the frequencies of the discrepancies observed for the 2CT corresponds to peaks of MSPLD of the 2ET reinforces this hypothesis. However, the matching between the frequency of the discrepancies and the peaks of MSPLD of the eccentric configuration is less good for the 44 tubes geometries. The perfectly axisymmetric cases thus appear as theoretical cases difficult to reproduce accurately experimentally. The human vocal tract cannot be considered as axisymmetric and from these observations, one can conclude that axisymmetric geometrical approximations of it are not likely to reproduce a realistic directivity above their first cut-on frequency. Introducing asymmetries in such kind of approximations would allow one to simulate a qualitatively more realistic directivity above the first cut-on frequency.

4.2.10. Convergent divergent shape

In the case of a convergent exit, the radiation losses are reduced in comparison with a non-convergent exit. As it can be observed on the input impedance curves, this induces a higher amplitude and a reduced bandwidth of the resonances inside the vocal tract. This can be related to the fact that the effects of the HOM are more localized around the peaks of MSPLD for reduced exit diameters (see Figure 10). When the diameter of the exit is reduced, one gets closer to a closed end boundary condition and the resonance frequencies are reduced. This can be related to the

decrease of the frequency of the peaks of MSPLD and input impedance for reduced exit diameters. Thus, the fact that the 44ET has a convergent exit and not the 2ET can explain why the effects of HOM are more localized in frequency for the 44 eccentric tubes geometry.

The influence of HOM can be seen even with a diameter of 10 mm which is almost three times smaller than the diameter of the large tube. Thus, more generally, it can be expected to be significant even if the mouth aperture is small compared to the preceding cavity. One can expect that the influence of HOM on the directivity varies qualitatively with the phonemes. A wide mouth aperture would induce an effect visible at all the frequencies from the first cut-on frequency of the vocal tract and a narrow mouth aperture would induce an effect limited to localized narrow frequency intervals.

4.2.11. Comparison between 2 tubes and 44 tubes

The two and 44 tubes eccentric geometries generate significantly different directivity patterns above 6.5 kHz. There is not the same number of peaks of MSPLD and they occur at different frequencies. The effects of HOM are more localized in frequency for the 44ET than for the 2ET. This can be seen in both the MSPLD and input impedance curves, and in the appearance of the two lobes pattern. Since a convergent exit is closer to a real vocal tract shape, the 2ET may over-estimate the effects of HOM and be qualitatively less realistic than the 44ET. However, the 44 eccentric tubes, while being more realistic because of a more accurate area function, does not take the cross-section shape or the bending of a real vocal tract into account. Thus, it would be interesting to study a more realistic geometry to test this hypothesis.

In terms of input impedance, the two and 44 tubes geometries are also quite different. The frequencies of the second formants are very different (1200 Hz for the 2 tubes and 930 Hz for the 44 tubes). This can be explained by the fact that the two tubes shape has been designed so that the frequencies of the two first formants corresponds to the average values observed for the vowel [ɑ] [18], whereas the area function of the 44 tubes geometries has been determined from MRI measurements [19]. In this case, the perturbations related to the measurement process (supine position in a noisy environment, see [19] for a detailed discussion) can induce formant frequencies different from the average value, but which still corresponds to a vowel [ɑ].

5. Conclusion

The measurements and the simulations presented in this study showed that the propagation of the HOM ψ_{01} inside vocal tract approximations of the vowel [ɑ] induces great variations (up to 50 dB of MSPLD) of the directivity patterns within small frequency intervals (of the order of 100 Hz) above 6.5 kHz.

A particular effect related to the mode-shape of ψ_{01} has been highlighted. The division of the exit surface into two areas with opposite phase generates a directivity pattern composed of two lobes separated by a low amplitude direction (of the order of 30 dB lower than the maximum). The variations of amplitude and phase of ψ_{00} and ψ_{01} with the frequency induce variations of the size of the lobes and the direction of the minimal amplitude (up to 60° within 100 Hz). Another effect of the propagation of ψ_{01} is the reduction of the directivity in comparison with the baffled circular piston directivity in some frequency intervals.

The excitation of HOM being very sensitive to small asymmetries, the relevance of axisymmetric vocal tract approximation for the simulation of the speech directivity above the first cut-on frequency of the vocal tract can be questioned. Introducing asymmetries in these approximations would allow one to simulate qualitatively more realistic directivity patterns.

The effects of HOM can be seen even if the radiating surface is small compared to the largest dimension of the vocal tract. However, in this case, they tend to be localized in particular frequency bands.

Since the vocal tract geometry is subject to inter- and intra-speaker differences even for a single phoneme as considered in this study, care is needed when extrapolating these conclusions to human speech. The two vocal tract approximations studied generate very different directivity patterns above the first cut-on frequency. Therefore, it would be interesting to study a more realistic geometry in order to know better which one is the closest to realistic directivity patterns.

Because the directivity patterns become different from the baffled circular piston above the first cut-on frequency in the eccentric configurations, one can expect that the directivity is perceived differently above this frequency. This assumption is reinforced by the fact that differences related to the position are visible in the spectrogram of the sound emitted by the replicas excited with a broadband noise above 6.5 kHz. However, one can question how the great variations (up to 50 dB of MSPLD within 100 Hz) of the effects of the HOM on the directivity with the frequency are perceived or if they can be noticed by a listener. Thus, it would be interesting to carry out perceptual tests to further investigate this question. For the same reasons, classical spectral measures such as the octave/third-octave frequency band might mask the localized aspect of this effect. On the other hand, given the limited angular region of some phenomena, it would be interesting to increase the angular resolution in the future. Because of the high frequency energy present in fricative and plosive sounds, this effect is expected to have more perceptual impact. Thus, it would be interesting to investigate its consequences with vocal tract geometries corresponding to these phonemes.

Acknowledgements

This research was supported by EU-FET Grant No. EUNISON 308874. We acknowledge Romain Murgat for his contribution to the experiments and Stephane Bayle for his contribution to the simulations.

References

[1] H. Dunn, D. Farnsworth: Exploration of pressure field around the human head during speech. J. Acoust. Soc Am. **10** (1939) 184–199.

[2] T. Halkosaari, M. Vaalgamaa, M. Karjalainen: Directivity of artificial and human speech. J. Audio Eng. Soc. **53** (2005) 620–631.

[3] D. Cabrera, P. Davis, A. Connolly: Long-term horizontal vocal directivity of opera singers: Effects of singing projection and acoustic environment. J Voice **25** (2011) e291–e303.

[4] A. Marshall, J. Meyer: The directivity and auditory impressions of singers. Acta Acust united Ac **58** (1985) 130–140.

[5] B. Monson, E. Hunter, B. Story: Horizontal directivity of low-and high-frequency energy in speech and singing. J. Acoust. Soc Am. **132** (2012) 433–441.

[6] J. Flanagan: Analog measurements of sound radiation from the mouth. J. Acoust. Soc Am. **32** (1960) 1613–1620.

[7] F. McKendree: Directivity indices of human talkers in english speech. INTER-NOISE and NOISE-CON Congress and Conference Proceedings, 1986, Institute of Noise Control Engineering, 911–916.

[8] J. Peng, T. Wang, S. Wu: Investigation on the effects of source directivity of chinese speech intelligibility in real and virtual rooms. Appl Acoust **74** (2013) 1037–1043.

[9] L. Savioja, J. Huopaniemi, T. Lokki, R. Väänänen: Creating interactive virtual acoustic environments. J. Audio Eng. Soc. **47** (1999) 675–705.

[10] J. Huopaniemi, K. Kettunen, J. Rahkonen: Measurement and modeling techniques for directional sound radiation from the mouth. Applications of Signal Processing to Audio and Acoustics, 1999 IEEE Workshop on, 1999, IEEE, 183–186.

[11] B. Monson, E. Hunter, A. Lotto, B. Story: The perceptual significance of high-frequency energy in the human voice. Frontiers in psychology **5** (2014).

[12] K. Motoki, N. Miki, N. Nagai: Measurement of sound-pressure distribution in replicas of the oral cavity. J. Acoust. Soc Am. **92** (1992) 2577–2585.

[13] R. Blandin, M. Arnela, R. Laboissière, X. Pelorson, O. Guasch, A. Van Hirtum, X. Laval: Effects of higher order propagation modes in vocal tract like geometries. J. Acoust. Soc Am. **137** (2015) 832–843.

[14] A. Roure: Propagation guidée. etude des discontinuités (Guided propagation. Study of discontinuities). Dissertation. Université d'Aix-Marseille, 1976.

[15] J. Kergomard, A. Garcia, G. Tagui, J. Dalmont: Analysis of higher order mode effects in an expansion chamber using modal theory and equivalent electrical circuits. J. Sound Vib. **129** (1989) 457–475.

[16] V. Pagneux, N. Amir, J. Kergomard: A study of wave propagation in varying cross-section waveguides by modal decomposition. Part I. Theory and validation. J. Acoust. Soc Am. **100** (1996) 2034–2048.

[17] J. Kemp: Theoretical and experimental study of wave propagation in brass musical instruments. Dissertation. University of Edinburgh, 2002.

[18] G. Fant: Acoustic theory of speech production: with calculations based on x-ray studies of russian articulations. Vol. 2. Walter de Gruyter The Hague, 1971. Chap. 1.

[19] B. Story: Comparison of magnetic resonance imaging-based vocal tract area functions obtained from the same speaker in 1994 and 2002. J. Acoust. Soc Am. **123** (2008) 327.

[20] A. Maurel, J. Mercier, V. Pagneux: Improved multimodal admittance method in varying cross section waveguides. Proc. R. Soc. A, 2014, The Royal Society, 20130448.

[21] A. Van Hirtum, Y. Fujiso: Insulation room for aero-acoustic experiments at moderate Reynolds and low Mach numbers. Appl Acoust **73** (2012) 72–77.

[22] A. Pierce, et al.: Acoustics: an introduction to its physical principles and applications. Vol. 20. McGraw-Hill New York, 1981.