

Influence of nasal cavities on voice quality: Computer simulations and experiments

Tomáš Vampola,^{1,a)} Jaromír Horáček,² Vojtěch Radolf,² Jan G. Švec,³ and Anne-Maria Laukkanen⁴

¹Department of Mechanics, Biomechanics and Mechatronics, Faculty of Mechanical Engineering, Czech Technical University in Prague, Technická 4, 160 00 Prague 6, Czech Republic

²Institute of Thermomechanics, Academy of Science of the Czech Republic, Dolejškova 5, 182 00 Prague 8, Czech Republic

³Voice Research Lab, Department of Biophysics, Faculty of Science, Palacky University Olomouc, Tr. Svobody 26, 771 46 Olomouc, Czech Republic

⁴Speech and Voice Research Laboratory, Faculty of Social Sciences, Tampere University, Virta, Åkerlundinkatu 5, 33100 Tampere, Finland

ABSTRACT:

Nasal cavities are known to introduce antiresonances (dips) in the sound spectrum reducing the acoustic power of the voice. In this study, a three-dimensional (3D) finite element (FE) model of the vocal tract (VT) of one female subject was created for vowels [a:] and [i:] without and with a detailed model of nasal cavities based on CT (Computer Tomography) images. The 3D FE models were then used for analyzing the resonances, antiresonances and the acoustic pressure response spectra of the VT. The computed results were compared with the measurements of a VT model for the vowel [a:], obtained from the FE model by 3D printing. The nasality affects mainly the lowest formant frequency and decreases its peak level. The results confirm the main effect of nasalization, i.e., that sound pressure level decreases in the frequency region of the formants F_1 – F_2 and emphasizes the frequency region of the formants F_3 – F_5 around the singer's formant cluster. Additionally, many internal local resonances in the nasal and paranasal cavities were found in the 3D FE model. Their effect on the acoustic output was found to be minimal, but accelerometer measurements on the walls of the 3D-printed model suggested they could contribute to structure vibrations. © 2020 Acoustical Society of America. <https://doi.org/10.1121/10.0002487>

(Received 31 May 2020; revised 11 October 2020; accepted 20 October 2020; published online 30 November 2020)

[Editor: Zhaoyan Zhang]

Pages: 3218–3231

I. INTRODUCTION

Human voice is produced through self-oscillations of the vocal folds excited by air flowing from the lungs. The vibration of the vocal folds produces the primary sound of voice by modulating the air stream in a narrow oscillating constriction between the two vocal folds called glottis. This primary sound signal propagates through the supraglottal cavities from the vocal folds to the lips and the nostrils (i.e., through the VT) which modifies the voice quality through acoustic resonances. The acoustic resonances of the VT create the so-called formants. The formants are determined by the size and shape of the VT cavities and define vowels and the voice timbre. The final sound quality of the human voice is thus given both by the characteristics of the vocal fold vibration and by the VT properties (Fant, 1960; Sundberg, 1987; Titze, 1994; Titze, 2006).

The influence of the geometric configuration of the main channel of the VT on the vocal output has been studied rather extensively. The influence of side cavities of human VT has received less attention and their role for the resulting vocal intensity has been considered undesirable or unimportant. The effect of side cavities, such as the piriform sinuses

and valleculae in the epilaryngeal part of the human VT, has been studied by some authors (Dang and Honda, 1997; Motoki, 2002; Fujita and Honda, 2005; Takemoto, 2013) who reported anti-resonances in the resulting vocal spectrum decreasing the radiation of some of the spectral frequencies from the mouth, thus reducing the vocal sound pressure level.

However, the newest studies have revealed that, besides the antiresonances, there are also additional resonances in the VT that occur due to these side cavities enhancing the voice quality (Delvaux and Howard, 2014; Vampola *et al.*, 2015b). These specific resonances can contribute, e.g., to a speaker's formant cluster created in the frequency range of 3–5 kHz. Furthermore, the spectral analysis of singers' voices indicates that the formant structure around 3–5 kHz is more complex than expected due to the existence of the side branches (Švec *et al.*, 2008).

The nasal tract (NT) has also been known to introduce antiresonances (dips) in the acoustic spectrum of voice. These antiresonances contribute to the perceptual identification of nasal phonemes such as /m/ or /n/. A dull resonance around 250 Hz and an antiresonance at about 500 Hz belong to the principal characteristic features of nasalization of vowels, and each nasal formant is paired with an antiresonance observed at the mouth (Hattori *et al.*, 1958).

^{a)}Electronic mail: Tomas.Vampola@fs.cvut.cz, ORCID: 0000-0002-1382-8363.

The results obtained by [Dang and Honda \(1996\)](#) indicate that each of the three major nasal sinuses, i.e., the sphenoid, maxillary, and frontal sinuses, contribute with antiresonances to the transmission characteristics of the NT. The studies by [Havel *et al.* \(2016\)](#) and [Havel *et al.* \(2017\)](#) showed that, under normal conditions, the maxillary and sphenoidal sinuses are accessible to sound excitation, but they absorb rather than amplify sound. This contrasts with the traditional assumption of voice pedagogy that the paranasal sinuses are important amplifiers of the final sound.

While nasality may be considered as an acoustic disadvantage since the spectral zeros diminish the radiated power of voice, in practice there seem to be benefits from nasality. Speech and singing training involves the use of nasals to a large extent. Velopharyngeal opening was observed in classical singers at least in vowel [a:] ([Birch *et al.*, 2002a](#); [Austin, 1997](#)) and “nasal resonance” has been regarded as being important for voice quality in some vocal pedagogy schools. Scientifically this issue has been controversial, however. For instance, some singers may confuse the “nasal resonance” or “head resonance” with the singer’s formant. Furthermore, the results by [Vennard \(1964\)](#) suggested that resonances of the nose and nasal sinuses do not affect the acoustic output of trained singers. Some studies, nevertheless, suggest that dampening of the first formant region by velopharyngeal opening may increase the prominence of the singer’s formant cluster that may be beneficial for the perceptual voice quality ([Birch *et al.*, 2002b](#); [Sundberg *et al.*, 2007](#)).

[Vampola *et al.* \(2008b\)](#) studied the effects of velopharyngeal insufficiency and clefting of soft or hard palate in relation to the voice quality for the vowels [a:], [i:], and [u:]. Acoustic influence of clefting on the pronunciation of vowel [a:] was found to be smaller than for the vowels [i:] and [u:]. The models of human vocal and nasal tracts used in the [Vampola *et al.* \(2008b\)](#) study were largely simplified, however. The geometric configuration of the nasal tract included only central nasal cavity and disregarded the separate paranasal cavities. For proper prediction of the voice quality it is necessary to establish more accurate models, particularly for the nasal cavities.

The present study investigates the acoustic effects of nasalization in a human female subject using both computational and physical models of phonation. The research question is: can nasalization enhance the sound level at speaker’s/singer’s formant region? This question is addressed for the open vowel [a:] and closed vowel [i:].

These vowels were selected since it has been shown that operatic singers use slight velopharyngeal opening for [a:] but that it occurs rarely for [i:] ([Birch *et al.*, 2002a](#); [Bauer, 2002](#)) or at least to a lesser extent for [i:] and [u:] ([Tanner *et al.*, 2005](#)). Slight velopharyngeal opening (VPO) is known to lower the amplitude of the first formant (F_1) for [a:] ([Fant, 1960](#)). This could be beneficial for singing as it may increase the perceptual prominence of the singer’s formant ([Birch *et al.*, 2002b](#)). The modelling results by [Sundberg *et al.* \(2007\)](#) showed that VPO lowered the

amplitude of F_1 also in [i:] and [u:] but in these vowels the bandwidth of F_1 increased remarkably. This is likely to produce nasal quality since broadening of the bandwidth of F_1 was found to be one of the main acoustic correlates of perceived hypernasality in the voice [e.g., [House and Stevens \(1956\)](#), [Huffman \(1990\)](#), and [Kozaki-Yamaguchi *et al.* \(2005\)](#)].

This study focuses on the effects of nasalization on the lowest five formants. Nasality is modeled here by interconnecting the acoustic cavities of the NT with the VT at the level of the soft palate. If the interconnection is completely shut, the NT is closed.

II. METHODS

A. Computational models of the vocal tract with and without the nasal tract

Sophisticated three-dimensional (3D) finite element FE models of the VT for the vowels [a:] and [i:] were created from the Computer Tomography (CT) recordings of a female subject (co-author of the present paper) during phonation, see [Vampola *et al.* \(2015a\)](#). As specified further, the geometric resolution of the models was highly detailed and comparable to the VT models of [Arnela *et al.* \(2016a\)](#) and [Arnela *et al.* \(2016b\)](#) who evaluated the influence of geometric VT simplifications on numerical simulations.

The CT images of the VT did not include the nasal cavities (for health protection reasons). The 3D FE model of the acoustic nasal cavities was therefore created from CT images of a head of another subject of the same gender, similar age and body size, who underwent the CT investigation for another, voice-unrelated medical reason. After segmentation of the CT images, the volume models of the VT and NT were carefully manually interconnected. The complete 3D FE models for the nasalized vowels [a:] and [i:] were created from these volume models, see Fig. 1.

The resonance frequencies of the main nasal side cavities (maxillary and sphenoid sinuses) were estimated using the formula for the Helmholtz oscillator [see [Havel *et al.* \(2017\)](#)],

$$f = \frac{c_0}{2\pi} \sqrt{\frac{S}{LV}}, \quad (1)$$

where c_0 is the speed of sound, V is the volume of the sinus cavity, S is the cross-sectional area of the cavity ostium, and L is the effective neck/ostium length [set to 0.2 cm according to [Havel *et al.* \(2017\)](#)]. The area S was estimated from the measured diameter D of the neck of the maxillary sinus in the volume model as $S = \pi(D/2)^2$, where $D \cong 1$ mm.

It is known that the main acoustic energy losses in the VT are caused by radiation of the sound from the mouth and nose to free field. These losses were modelled by a circular plate vibrating like a piston in an infinite wall. The frequency-dependent acoustic impedance was derived for this case earlier ([Vampola *et al.*, 2008a](#)),

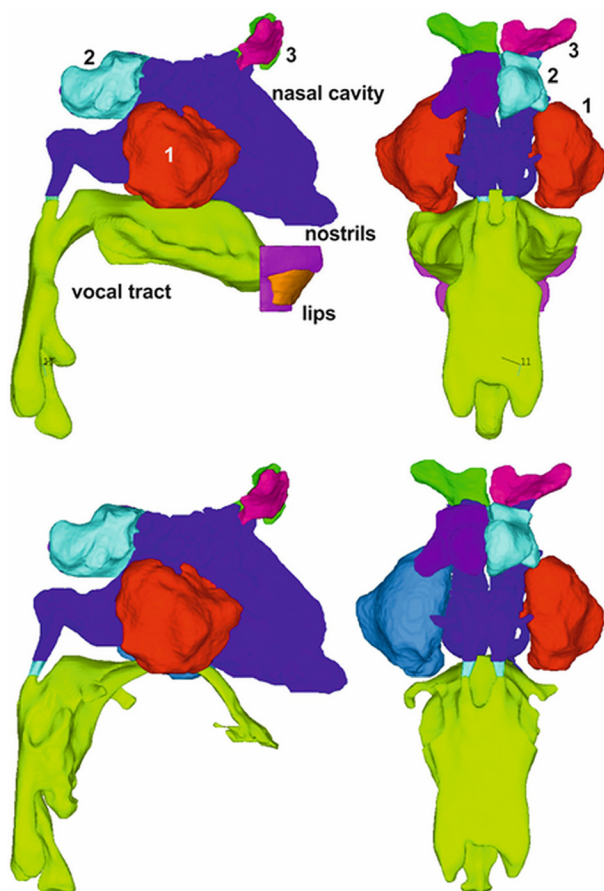


FIG. 1. (Color online) Computer volume models of the human VT for the nasalized vowel [a:] (top) and [i:] (bottom). The following left and right paranasal sinuses are attached to the nasal cavity: 1—maxillary sinuses, 2—sphenoid sinuses, 3—frontal sinuses.

$$Z_a = \frac{c_0 \rho_0}{S} (A + iB), \quad A = 1 - \frac{J_1(2kR)}{kR}, \quad B = \frac{H_1(2kR)}{kR}. \quad (2)$$

Here, c_0 is sound velocity, ρ_0 is air density, R is an equivalent radius of the vibrating plate calculated from the cross-section of the VT model at the level of the lips and nostrils, $k = \omega/c_0$ is the wave number, ω is angular frequency, and i is the imaginary unit. The Bessel J_1 and Struve H_1 functions can be calculated using series expansion. The acoustic energy losses at the inside VT walls were incorporated in the model via the boundary admittance coefficient $\mu = r/\rho_0 c_0$, where r is the real component of the specific acoustic impedance (resistance term).

The boundary admittance coefficient $\mu \in \langle 0, 1 \rangle$ is related to the dimensionless absorption coefficient of soft tissues α by the formula [see, e.g., Richards and Mead (1960)]

$$\alpha = \frac{1}{0.5 + 0.25(\mu + \mu^{-1})}. \quad (3)$$

Equation (2) can be used in harmonic analysis of the assembled system of differential equations describing the

acoustical system. Calculation of transient analysis requires converting the frequency dependent acoustic impedance to the time domain (Nieuwenhof and Coyette, 2001; Arnela et al., 2013; Takemoto, 2010).

Input data for the computations were considered as follows: air density $\rho_0 = 1.2 \text{ kg/m}^3$, speed of sound $c_0 = 347 \text{ m/s}$, dimensionless absorption coefficient $\alpha = 0.02$ ($\mu = 0.005$) (Dedouch et al., 2000). The speed of sound was targeted for the temperature at which the physical model experiments were made, i.e., 22–25 °C, using the relationship

$$c_0 = (331.8 + 0.61 t) \text{ m/s}, \quad (4)$$

where t is the temperature.

The governing equations for the acoustic pressure in the acoustic cavities of the full 3D FE model of the VT can be written in the FE formulation as

$$\mathbf{M}\ddot{\mathbf{P}} + \mathbf{B}\dot{\mathbf{P}} + \mathbf{K}\mathbf{P} = \mathbf{0}, \quad (5)$$

where $\mathbf{M}$, $\mathbf{B}$, and $\mathbf{K}$ represent global ($N \times N$) mass, damping and stiffness matrices, $\mathbf{P}$ is the vector of nodal pressures in N nodes inside the VT, and the dot and double dots above the pressure denote the first and second time derivatives, respectively. The total number of the nodes for the full 3D FE model was about $N \approx 500\,000$. Due to the highly detailed geometrical configuration of the acoustic cavities the total number of the elements for the full 3D FE model was about 2 000 000. The convergence criteria of the FE mesh were tested on models with different size of elements.

It is recommended to use about 8–10 elements per wavelength for accurate description of the oscillating process. Assuming the speed of sound $c_0 = 347 \text{ m/s}$ and frequency range up to 10 kHz the maximal element size can be computed by the formula

$$element_{size} = \frac{c_0}{10f} = 3.47 \text{ mm}. \quad (6)$$

Figure 2 reveals that all the elements used in the FE model were smaller and thus the FE mesh was able to describe oscillating process with sufficient accuracy. Figure 3 shows the volume of the individual acoustic cavities of the assembled FE model. The assembled differential equations from Eq. (5) were solved in the frequency and time domain by the commercial software ANSYS. The Newmark integration scheme was used with the time step 0.00001 s for transient analysis.

A broadband frequency flow velocity pulse at the level of glottis was used to obtain the acoustic resonant properties of the FE model [similar to Takemoto (2010) and Gully et al. (2018)]. The values of the acoustic pressure were computed at the position of the lips and nose. The excitation pulse of the VT in the time and frequency domain is shown in Fig. 4. The spectrum of the pulse ($0.8 \text{ dm}^3/\text{s}$) is practically flat in the frequency range up to about 10 kHz. At 10 kHz the spectral level of the pulse shows a decrease of only 1.38% from the initial value at 0 Hz.

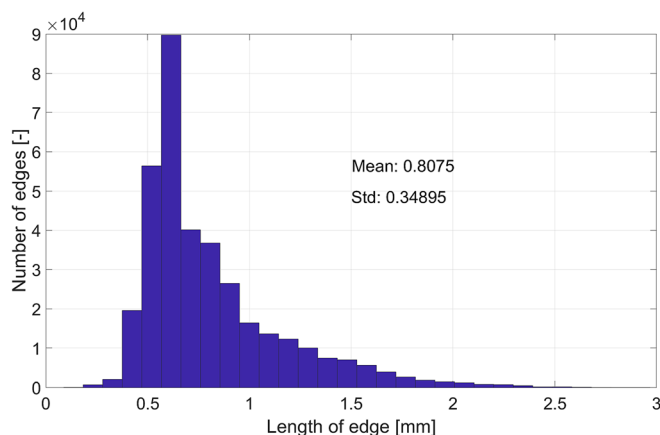


FIG. 2. (Color online) Histogram of the length of the edges of the tetrahedral finite elements. The average finite element size was about 0.8 mm.

B. Measurements on the physical models

In addition to the FEM model, a physical VT model corresponding to the nasalized vowel [a:] was 3D-printed, see Fig. 5. The characteristics of the material used can be seen in Table I.

A thin movable metal sheet was air-tightly positioned between the nose and VT cavities for simulating the velopharyngeal opening. Nasalization of the VT was realized by fully opening the velopharyngeal channel. The experiments with the model were performed in a large laboratory room with no reflecting surfaces in the neighbourhood of the model and with the background noise of ca. 30.5 dB to avoid acoustic contamination of the measured sound signals.

An earphone of diameter 15 mm was placed at the inlet of the VT at the position of the glottis and sealed. A unidirectional linear sine sweep signal in the range from 150 to 5155 Hz with the change of 715 Hz/s was used to acoustically excite the VT model twice during the 14 s recording time. The frequency response of the earphone measured in a free space (without VT model) was subtracted from the measured spectra. The frequency range of the earphone was

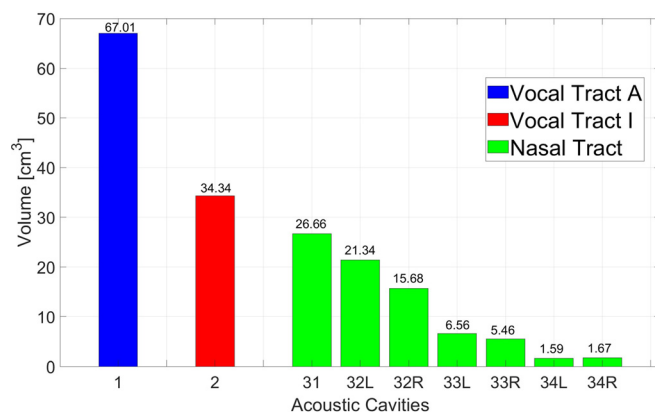


FIG. 3. (Color online) Volumes of the VTs for the vowels [a:] and [i:] and of the specific nasal cavities: 31—central nose cavity, 32L, 32R—left and right maxillary sinuses, 33L, 33R—left and right sphenoid sinuses, 34L, 34R—left and right frontal sinuses.

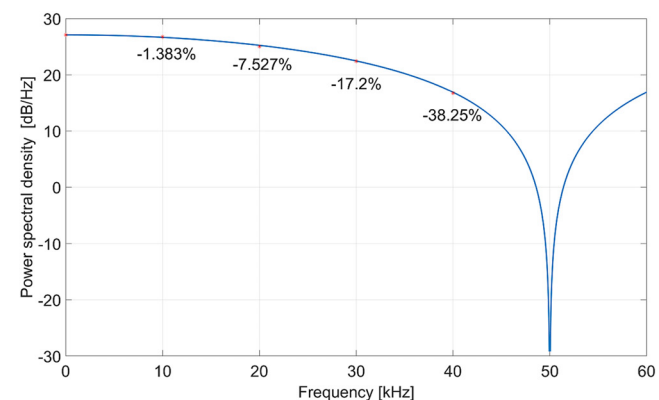
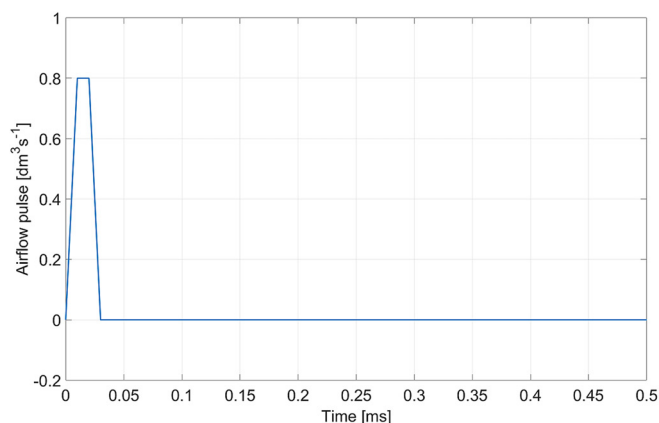


FIG. 4. (Color online) The excitation pulse of the VT in the time (top) and frequency (bottom) domain.

usable above 200 Hz. The acoustic signal was registered by the microphone of the sound level meter B&K 2239 positioned at the distance of 3 cm from the mouth of the model.

Acoustic pressure inside the model of the oral cavity was measured by the B&K 4138 miniature microphone (range 6.5 Hz–140 kHz) positioned 3 mm from the lips. A special B&K 4182 microphone probe (range 1 Hz–20 kHz) installed 3 mm inside the right nostril measured the sound in the nose of the model, see Fig. 5. Lower frequency limit of the outer microphone, i.e., of the sound level meter B&K 2239 installed at the distance of 3 cm from the mouth of the

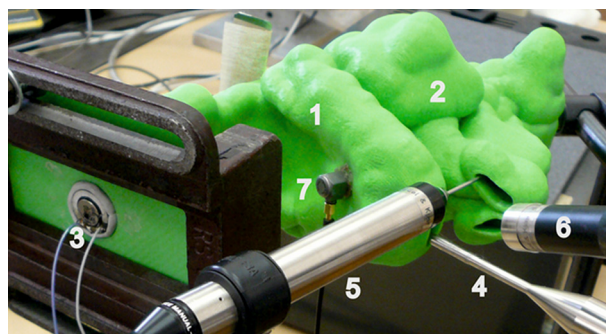


FIG. 5. (Color online) Measurement setup for the earphone-excited VT model for the nasalized vowel [a:]. (1—VT, 2—nasal tract, 3—earphone, 4—miniature microphone, 5—microphone probe, 6—sound level meter, 7—accelerometer).

TABLE I. Vero YellowV RGD83 - material characteristics of the 3D-printed model.

Material Property	Value
Tensile Strength	50–65 MPa
Elongation at Break	10%–25%
Modulus of Elasticity	2000–3000 MPa
Flexural Strength	75–110 MPa
Flexural Modulus	2200–3200 MPa
Shore Hardness (D)	83–86 (Scale D)
Polymerized Density	1.17–1.18 g/cm ³

model was set to 21.4 Hz and no disturbing sound originated in surrounding was observed in the frequency range above this limit. Any cross-talk between the microphone signals was excluded.

The vibrations of the model surface were monitored by the B&K 4393 accelerometer. All the measured signals were simultaneously sampled at the frequency of 16.4 kHz and registered by the measurement system Brüel & Kjaer PULSE type 3560 C with the Input/Output Controller Modules Type 7537A and 3109, controlled by a personal computer equipped by the SW PULSE LabShop Version 10.

C. Human experiment on nasalization

The same female subject from whom the vocal tract model was created (Vampola *et al.*, 2015a), was recorded in a sound-treated studio while sustaining vowels [a:] and [i:] at habitual speaking pitch and loudness with and without nasalization. The sound was captured and measured using the Brüel & Kjaer Mediator 2238 level meter/microphone, which was placed at 40 cm in front of the mouth. The signals were recorded on a PC using the sound card IFOCUSRITE and SOUNDFORGE software (44.1 kHz sampling rate, 16 bits amplitude quantization).

III. RESULTS

A. FE modelling results

The effect of interconnecting the oral and nasal cavities for the vowel [a:] is demonstrated in Fig. 6. Nine acoustic resonances (R_1 – R_9) are marked in the acoustic pressure response spectra computed for the non-nasalized and nasalized VTs. New resonances and anti-resonances (N_1 – N_{11}) appear in the acoustic pressure response spectra for the nasalized vowel.

The peak levels L_{R1} and L_{R2} of the VT resonances R_1 and R_2 were markedly lowered by the nasality, by ca. 14 and 6.4 dB, respectively. The peak levels of the higher resonances R_3 – R_5 were practically unchanged. The first resonance frequency $R_1 \cong 599$ Hz increased by ca. 26 Hz to $R_1 \cong 625$ Hz while the second resonance frequency $R_2 \cong 1265$ Hz remained practically unchanged. The third resonance frequency R_3 markedly increased from $R_3 \cong 2737$ Hz towards the frequencies R_4 and R_5 , which were in the frequency range of the singer's formant, i.e., between 3

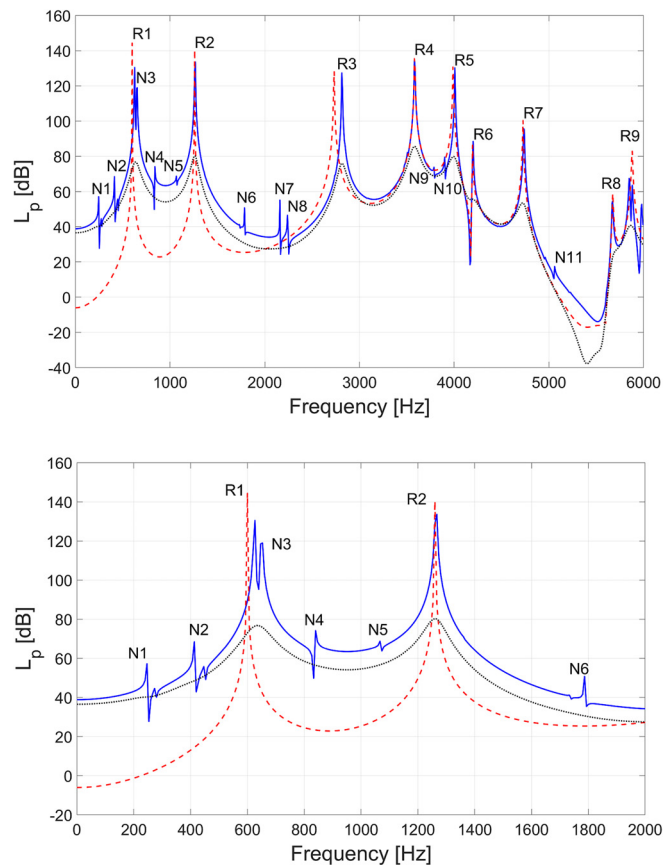


FIG. 6. (Color online) Acoustic response spectra of the 3D finite-element (FE) model of the VT for vowel [a:] computed at the lips. Upper panel: Comparison of the spectra without (dashed red line) and with nasality (full blue line) computed without damping ($\alpha=0$). Lower panel: The detail of the same spectra in the lower frequency range without (red dashed line) and with nasality (blue full line). Notice that the two lowest peaks N_1 , N_2 caused by the nasality are doubled and are accompanied by anti-resonance peaks. The effect of damping of the VT walls ($\alpha=0.02$) is shown by the black dotted lines for the model with nasality. The acoustic pressure level was computed as $L_p = 20 \log_{10}(|p|/p_0)$, where $p_0 = 20 \text{ e} - 6 \text{ Pa}$.

and 4 kHz. Nasality influenced the resonance frequencies R_4 and R_5 only minimally.

A detailed investigation revealed that the lowest nasal resonances N_1 – N_4 belonged to the paranasal cavities, i.e., to the maxillary, sphenoid and frontal sinuses (see Fig. 7) where the lowest acoustic mode shapes of vibration for the nasalized vowel [a:] occurred. Because the damping of the acoustic system was moderate, the acoustic mode shapes of vibration were approximated by modal analysis of the undamped system. For this the boundary condition of zero acoustic pressure at the outlet of the model was used.

The first two oro-nasal resonances N_1 and N_2 were doubled and accompanied by anti-resonances which followed the resonance peaks, see the detail in the lower frequency range in Fig. 6. The first two pairs of oro-nasal resonances N_1 , N_2 were below the first VT resonance frequency R_1 . The first nasal double resonance N_1 (two resonance-anti-resonance pairs) at 250 and 277 Hz belonged to the left and right maxillary sinuses coupled to the rest of the VT. Similarly, the second double resonance N_2 at 416 and 452 Hz belonged

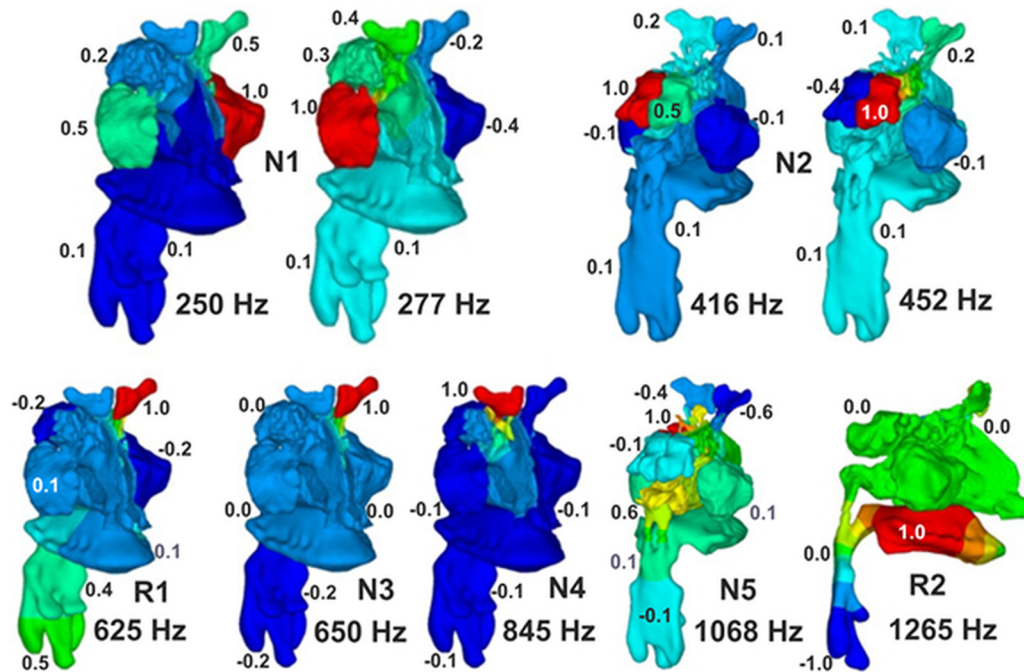


FIG. 7. (Color online) The lowest acoustic mode shapes of vibration computed for the nasalized vowel [a:]. The mode shapes correspond to the resonances marked in the computed spectra in Fig. 6. The resonances R₁ and R₂ are linked to the VT; the resonances N₁–N₅ are linked to the NT. The acoustic pressure eigenmodes are normalized to unity, maximal pressures are red and pressure minima dark blue. Small numbers also mark the pressure values from a minimum to the maximum value 1.0 around each eigenmode. Computed without damping ($\alpha = 0$).

to the left and right sphenoid sinuses, respectively (see Fig. 7). The excited maximal acoustic pressure amplitudes are, in both cases, at lower frequencies for the larger (left) cavities. The maxillary and sphenoid paranasal cavities behave as Helmholtz resonators [see, e.g., [Dang and Honda \(1996\)](#)]. The third and fourth nasal resonances N₃ and N₄ are between the VT resonances R₁ and R₂ at frequencies of 650 and 845 Hz, respectively. They are associated with the left and right frontal sinuses.

Eleven nasal resonances were detected at the lips in the frequency range up to 6 kHz (see Fig. 6). Here, all the nasal resonance peaks N₁–N₁₁ were small compared to the peaks

of the VT resonances R₁–R₅. Inside the nostrils, many more nasal resonances occurred, see Fig. 8. The resonance frequencies were slightly different when measured at the left or right nostril. Compared to the measurements inside the lips, the density of the inner nasal resonances was much higher especially in the range above 2.5 kHz when detected at the nostrils.

All the graphs presented in Figs. 6–8 were computed without considering the absorption of the VT and NT walls. Once the sound absorption was included, all the nasal resonances N₁–N₁₁ disappeared from the spectra computed inside the lips (see Fig. 9). Inside the lips, the main acoustic

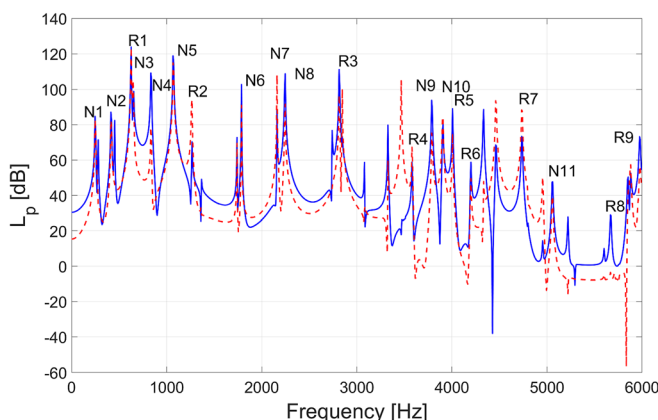


FIG. 8. (Color online) Spectra of the acoustic responses computed at the left (full blue line) and right (dashed red line) nostrils by using the 3D finite-element (FE) model of the VT for the nasalized vowel [a:]. Computed without damping ($\alpha = 0$).

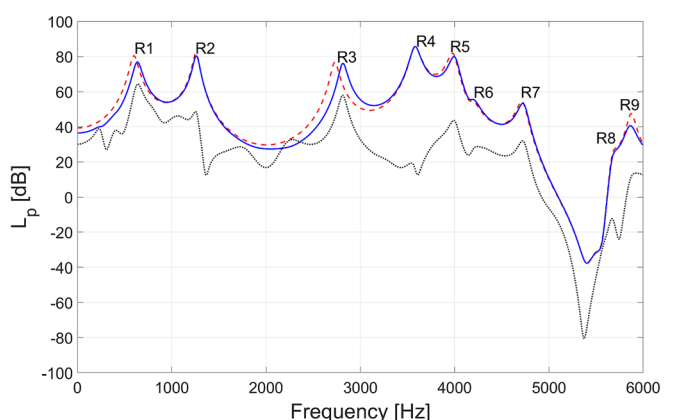


FIG. 9. (Color online) Comparison of the spectra without (dotted red line) and with nasality (full blue line) computed at the lips and acoustic response spectra of the left nostril (black dotted line) for the nasalized vowel [a:] including the damping of the VT and NT cavity walls ($\alpha = 0.02$).

TABLE II. Acoustic resonance frequencies for the nasalized and non-nasalized phonation on vowel [a:] evaluated from (a) the acoustic pressure response spectra computed by the finite element method (FEM) without damped walls, (b) measurements of a physical model, and (c) *in vivo* measurements of a female subject. The nasality effect for each resonance was calculated as the percentage difference. The symbolic notation for the formant and resonance frequencies is adopted from the recommendations of Titze *et al.* (2015).

Vowel [a:]—phonation	f_{N1} [Hz]	f_{N2} [Hz]	f_{R1}, f_{F1} [Hz]	f_{R2}, f_{F2} [Hz]	f_{R3}, f_{F3} [Hz]	f_{R4}, f_{F4} [Hz]	f_{R5}, f_{F5} [Hz]
FEM—without nose	250	416	625	1265	2820	3590	3990
Difference [%] (FEM-phys.model) / FEM	/	/	16.0	23.2	6.8	13.1	6.7
FEM—without nose	/	/	599	1260	2739	3597	3978
Difference [%] (comparing to earphone)	/	/	23.9	22.9	11.5	10.4	3.7
(FEM – phys. model) / FEM							
FEM—nasality effect [%] (nasal-normal) / normal			4.3	0.4	3.0	−0.2	0.3
Physical model—without nose	320		525	971	2627	3120	3723
Physical model—without nose	/		456	971	2425	3223	3831
Nasality effect [%] (nasal-normal) / normal			15.1	0.0	8.3	−3.2	−2.8
Female—nasalized phonation	522		701	1130	2952	3702	4162
Difference [%] (comparing to FEM)	/		−12.2	10.7	−4.7	−3.1	−4.3
(FEM- <i>in vivo</i>) / FEM							
Female—non-nasalized phonation	/		723	1172	2705	3837	4330
Difference [%] (comparing to FEM)	/		−20.7	7.0	1.2	−6.7	−8.8
(FEM- <i>in vivo</i>) / FEM							
Female—nasality effect [%] (nasal-normal) / normal			−3.0	−3.6	9.1	−3.5	−3.9

effect of nasality was the reduction of the peak level of the first resonance R_1 by ca. 2.4 dB, and an increase of the resonance frequencies R_1 and R_3 by ca. 36 and 83 Hz, respectively. It can be noted that these most important changes caused by the nasality are practically the same as observed in Fig. 6 for the undamped FE model.

Numerically, the results shown in Figs. 6–9 are provided in Table II summarizing the computed resonance frequencies for the vowel [a:]. According to the FE modelling, the nasality affects the main acoustic resonance frequencies f_{R1} – f_{R5} only slightly, the maximal increase is 4% for f_{R1} .

Similarly as done for the vowel [a:], the results for the vowel [i:] are shown in Figs. 10 and 11 for both the undamped and damped FE models. The influence of interconnection between the oral and nasal cavities on the resonances for the vowel [i:] is demonstrated in Fig. 10. Due to the nasality, the first resonance peak R_1 decreased its level by nearly 10 dB and increased its frequency from ca. 231 Hz to ca. 339 Hz. This was true for both the undamped and damped VT walls. The frequencies and peak levels of all the higher resonances R_2 – R_8 were negligibly influenced by the nasality. Additional nasal resonances N_1 and N_2 appeared at 240 Hz and 433 Hz below and above the first resonance frequency R_1 , respectively. Higher nasal resonance peaks were also detected, but showed very small amplitudes. When the damping of the VT surface was included, no apparent nasal resonance peaks appeared in the spectrum of the acoustic pressure response computed at the lips (Fig. 10, lower panel).

Figure 11 shows the lowest acoustic eigenmodes computed for the nasalized vowel [i:]. Similarly to the vowel [a:], the first nasal double resonances N_1 , N_2 belonged to the doubled maxillary and sphenoid sinuses. The resonances N_3 and N_4 were associated with the resonances of the left and right frontal sinuses at the frequencies of 642 and 835 Hz.

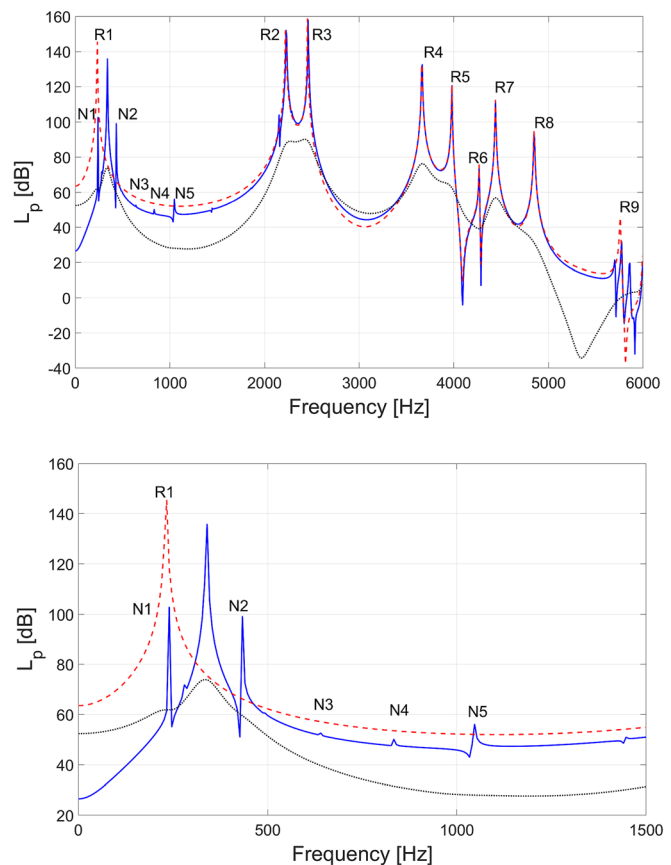


FIG. 10. (Color online) Acoustic response spectra of the 3D finite-element (FE) model of the VT computed at the lips for vowel [i:]. Upper panel: Comparison of the spectra without (dashed red line) and with nasality (full blue line). Computed without damping ($\alpha=0$). Lower panel: The detail of the same spectra in the lower frequency range without (red dashed line) and with nasality (blue full line). The effect of damping of the VT walls ($\alpha=0.02$) is shown by the black dotted lines for the model with nasality.

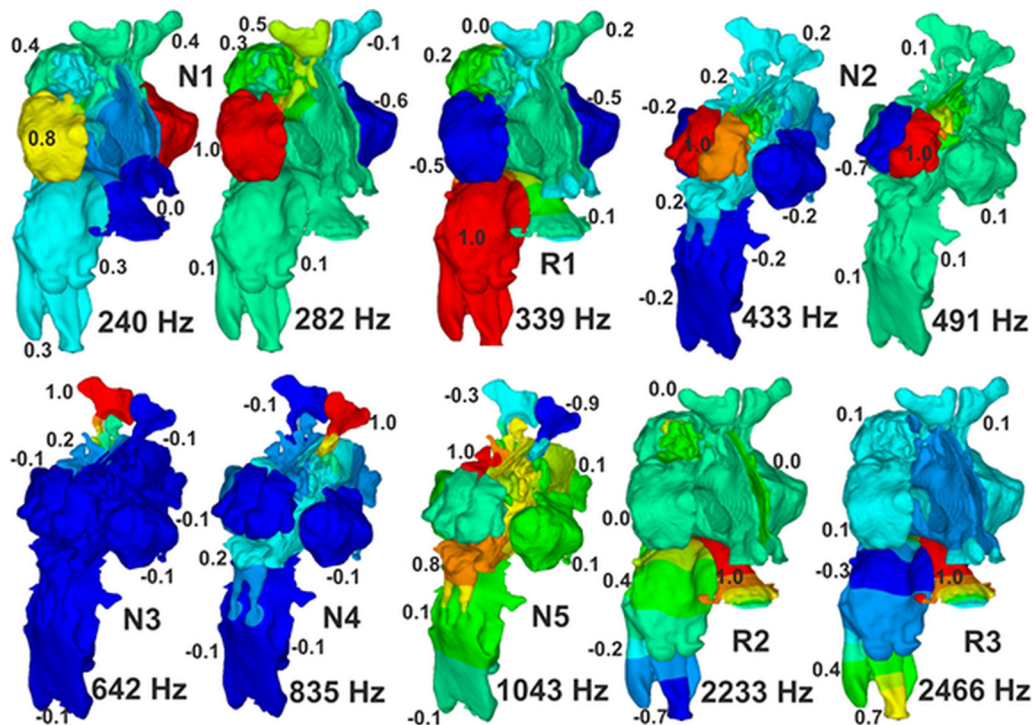


FIG. 11. (Color online) The lowest normalized acoustic pressure eigenmodes computed for the nasalized vowel [i:]. The mode shapes correspond to the resonances marked in the computed spectra in Fig. 9. Computed without damping ($\alpha = 0$). The colors and labels are the same as in Fig. 7.

These were excited only weakly, however, even if the spectrum shown in Fig. 9 was computed without the damping of the VT walls. The computed resonance frequencies for the vowel [i:] together with the results from physical measurements are summarized in Table III.

Comparison of the acoustic eigenmodes and resonance frequencies for the nasalized vowels [a:] and [i:] in Figs. 7 and 11, respectively, shows their similarity. Namely, the first two oro-nasal resonances that correspond to the eigenmodes with maximal amplitudes in the left and right maxillary sinus were around 250 and 280 Hz, respectively, in both cases. The resonances of the second oro-nasal-resonance pair N_2 , corresponding to the eigenmodes with maximal amplitudes in the left or right sphenoidal sinus, were also close, i.e., 416 and 452 Hz for the vowel [a:] and 433 and 491 Hz for the vowel [i:]. Also, the resonance frequencies of the oro-nasal resonances N_3 and N_4 corresponding to the

eigenmodes with the maximal amplitudes in the left or right frontal sinus were similar, i.e., the frequencies 650 and 845 Hz for the vowel [a:] were close to the frequencies 642 and 835 Hz for the vowel [i:]. Moreover, the eigenmodes corresponding to N_3 and N_4 for the two vowels were very similar in their appearance. Analogous similarities between the [a:] and [i:] vowels could be found also in the higher-order eigenmodes up to the upper limit of the frequency range. This indicates that the inner-nose acoustic modes of vibration depend only little on the geometry of the VT cavities, at least for the two vowels considered in this paper, and that these inner-nose modes are practically without any important interaction with ordinary acoustic eigenmodes of the non-nasalized vowels. Moreover, all the nasal resonances disappeared from the sound pressure computed at the lips when the damping of the walls of supraglottal spaces was included in the FE model. The only important acoustic

TABLE III. Acoustic resonance frequencies f_{R_n} and the corresponding formant frequencies f_{F_n} for the nasalized and non-nasalized vowel [i:] obtained from (a) the acoustic pressure response spectra computed by finite element method (FEM) without damped walls and (b) *in vivo* measurements of a female subject. Nasality effect was calculated as the percentage difference.

Vowel [i:]—phonation	f_{N1} [Hz]	f_{R1}, f_{F1} [Hz]	f_{N2} [Hz]	f_{R2}, f_{F2} [Hz]	f_{R3}, f_{F3} [Hz]	f_{R4}, f_{F4} [Hz]	f_{R5}, f_{F5} [Hz]
FEM - with nose	240 282	339	433 491	2233	2466	3680	3992
FEM—without nose	/	231	/	2209	2453	3670	4000
FEM—nasality effect [%] (nasal-normal) / normal		46.8		1.1	0.5	−0.4	−0.2
Female—nasalized phonation	270	320	470	2280	2600	2830	3640
Difference [%] (FEM- <i>in vivo</i>) / FEM		5.6		−2.1	−5.4	23.1	8.8
Female—non-nasalized phonation	/	282	/	2425	2810	3840	4525
Difference [%] (FEM- <i>in vivo</i>) / FEM	/	−22.1	/	−9.8	−14.6	−4.6	−13.1
Female—nasality effect [%] (nasal-normal) / normal		13.5		−6.0	−7.5	−26.3	−19.6

effect of the nasalization therefore occurred in the changes of the frequencies of the VT resonances R_1 and R_3 for the vowel [a:] and R_1 for the vowel [i:], and the decrease of the peak levels L_{R1} of the resonances R_1 for both vowels.

B. Results from the physical 3D-printed model experiments

Figure 12 shows the measurement results obtained from the physical models of nasalized and non-nasalized vowel [a:]. They compare the spectra of the pressure signals measured in nasal tract [Fig. 12(a)], the spectra of the pressure signals measured inside the mouth [Fig. 12(b)] and the

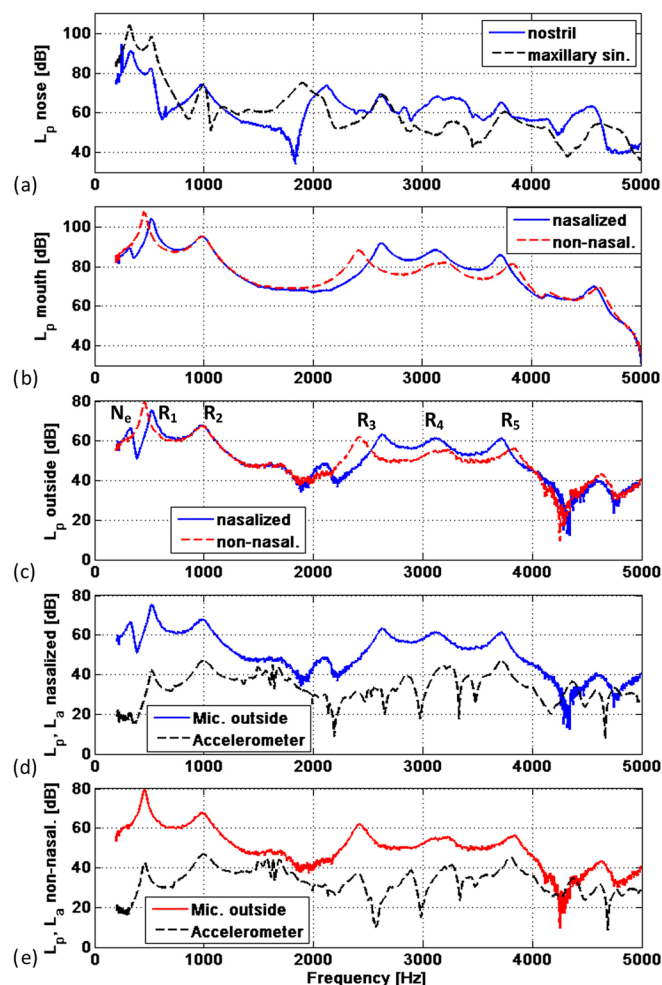


FIG. 12. (Color online) Spectra obtained from the physical model of the nasalized and non-nasalized VT for the vowel [a:] excited by the earphone: (a) measurements for nasalized VT in the nostril (solid line) and in the maxillary sinus (dashed line); (b) measurements inside the mouth for nasalized (solid line) and non-nasalized (dashed line) VT; (c) measurements by the microphone outside of nasalized (solid line) and non-nasalized (dashed line) VT; (d) measurements for nasalized VT done by the microphone (solid line) and by the accelerometer placed at a location of lower jaw of the model (dashed line); (e) measurements for non-nasalized VT obtained by the microphone outside (solid line) and by the accelerometer placed at a location of a lower jaw of the model (dashed line). The acoustic pressure level was computed as $L_p = 20 \log_{10}(|p|/p_0)$, where $p_0 = 20 \text{ e-}6 \text{ Pa}$. Acceleration level was computed from the autospectrum $a(f)$ of the measured acceleration as $L_a = 20 \log_{10}(|a(f)|/a_0)$, where $a_0 = 1 \text{ e-}6 \text{ ms}^{-2}$ is a reference value.

spectra measured 3 cm in front of the mouth [Fig. 12(c)]. Moreover, the spectra of the pressure signals measured outside the mouth are compared with the spectra of the vibrations measured by the accelerometer placed on the VT wall [Figs. 12(d) and 12(e), recall also Fig. 5]. It is obvious that the earphone excited both the acoustic and structural resonances. Among other resonances, the accelerometer showed clear peaks at all VT acoustic resonances R_1 – R_5 for the non-nasalized VT [Fig. 12(e)]. Several more structural resonances were visible in the structural vibrational response, particularly near the frequency 1.6 kHz, slightly above 2.8 kHz, and around 4.3 kHz [Figs. 12(d) and 12(e)]. These structural resonances did not show as peaks in the radiated sound and thus did not contribute to the acoustic output. Nasalization caused changes of the accelerometer spectra particularly in the frequency range of 2–2.7 kHz. A new antiresonance occurred here around 2.2 kHz [Fig. 12(d)], around the region of the nasal resonances N_7 – N_8 detected in the undamped FE model (recall Fig. 6, upper panel, and Fig. 8). Another antiresonance occurred in the accelerometer spectrum around 2.7 kHz obscuring the third VT resonance of the nasalized vowel [Fig. 12(d)].

The VT resonances R_1 – R_5 in Fig. 12 corresponded well to the resonances in Fig. 9 obtained from FE modelling when the damping on the walls of the VT and NT was considered. Numerically, the results are provided in Table II. Similarly, as in the FE modelling, the first resonance frequency f_{R1} increased by nasality while its peak level L_{R1} decreased. The second resonance frequency f_{R2} was without any noticeable change but the resonance frequencies f_{R3} – f_{R5} were closer together due to nasality. This was caused by an increase of the resonance frequency f_{R3} and decrease of f_{R5} . This effect also led to an increase of resonance peak levels L_{R4} and L_{R5} and promoted creating a singers' formant cluster around 3 kHz. In contrast to the four nasal resonances found in the undamped FE model below the resonance R_1 (see Figs. 6 and 7), only one cumulative nasal resonance N_e and one anti-resonance was found below R_1 in the experiments performed on the model and in the human subject, see Fig. 12 and Table II.

A low level acoustic resonance at ca. 2000 Hz was detected in Fig. 12(c) in front of the lips and also at the nostril [Fig. 12(a)]. Around this frequency we found the group of inner nose resonances N_6 – N_8 clearly visible in the computed acoustic pressure response spectra shown in Figs. 6 and 8 without damping of the VT walls. When the damping in FE model was included, these nasal resonances appeared only in the spectra computed at the nostril, see Fig. 9. However, no such inner acoustic nasal resonances are identifiable in the mouth [see Fig. 12(b)] nor in the acceleration measured on the lower jaw [see Fig. 12(d)].

Table IV compares the sound pressure levels (SPL) measured on the vowel [a:] model in two consecutive frequency intervals. The lower frequency interval (200–1500 Hz) includes the first two VT resonances/formants while the upper frequency interval (1500–5000 Hz) includes the rest of the resonances/formants that can still

TABLE IV. SPLs obtained from the non-nasalized and nasalized physical model of VT for the vowel [a:] using sweep excitation by the earphone. The SPLs were measured in front of the lips and are shown for two adjacent frequency intervals (0.2–1.5 kHz and 1.5–5 kHz).

Vowel [a:] phonation-model	SPL ₁ [dB] at 0.2–1.5 kHz	SPL ₂ [dB] at 1.5–5 kHz	SPL ₁ –SPL ₂ [dB]
Nasalized	95.47	89.73	5.74
Non-nasalized	97.54	86.70	10.84
Difference	–2.07	3.03	–5.10

potentially contribute towards the overall SPL in phonation. The level difference was calculated by subtracting the SPL in the upper frequency interval from the SPL in the lower frequency interval. The results show that the level difference is lower for nasality, because the SPL is lower in the lower frequency region and higher in the higher frequency region for the nasalized vowel [a:]. Consequently, it means that the frequency region of higher formants is relatively more important for loudness in the nasalized vowel [a:] than in the non-nasalized one.

C. Results from *in vivo* experiments

The sound spectra for the nasalized and non-nasalized vowels [a:] and [i:], obtained *in vivo* from the female subject, are shown in Fig. 13. The red lines help to identify the formant frequencies, which are presented in Tables II and III.

In correspondence with the experiments on the models, the first nasal resonance at 522 Hz appeared below the first formant frequency 701 Hz in the *in vivo* measured vowel [a:]. While the first and second formant frequencies f_{R1} and f_{R2} at ca. 723 and 1172 Hz were changed by nasality by only –3.0 and –3.6%, respectively, the third formant frequency f_{R3} increased from 2705 Hz by ca. 9% and f_{R5} decreased from 4330 Hz by ca. 4%. These frequency shifts were similar to those observed in the models and showed the tendency to form a singer's/speaker's formant cluster in the frequency range 3–4 kHz (see Table III and Fig. 13).

When comparing the nasal and non-nasal spectra for the vowel [i:], additional nasal formants were to some extent identifiable in Fig. 13 just below and above the first formant frequency. These two measured nasal formants are likely related to the nasal resonances N_1 and N_2 computed for the vowel [i:] with zero damping (recall Fig. 10). FE modelling revealed that in the vowel [i:] the nasality effect is maximal for the first resonance frequency f_{R1} which increased by ca. 47%. The effect on the higher resonance frequencies was rather small.

In vivo measurements revealed that the first formant frequency f_{F1} of the vowel [i:] increased by 13.5%, while its second formant frequency f_{F2} decreased by 6% and the third formant frequency f_{F3} decreased by 7.5% due to nasalization. The largest nasalization-caused shifts in the *in vivo* measurements were found for the formant frequencies f_{F4} and f_{F5} . Here f_{F4} decreased by ca. 26% and f_{F5} by ca. 20%,

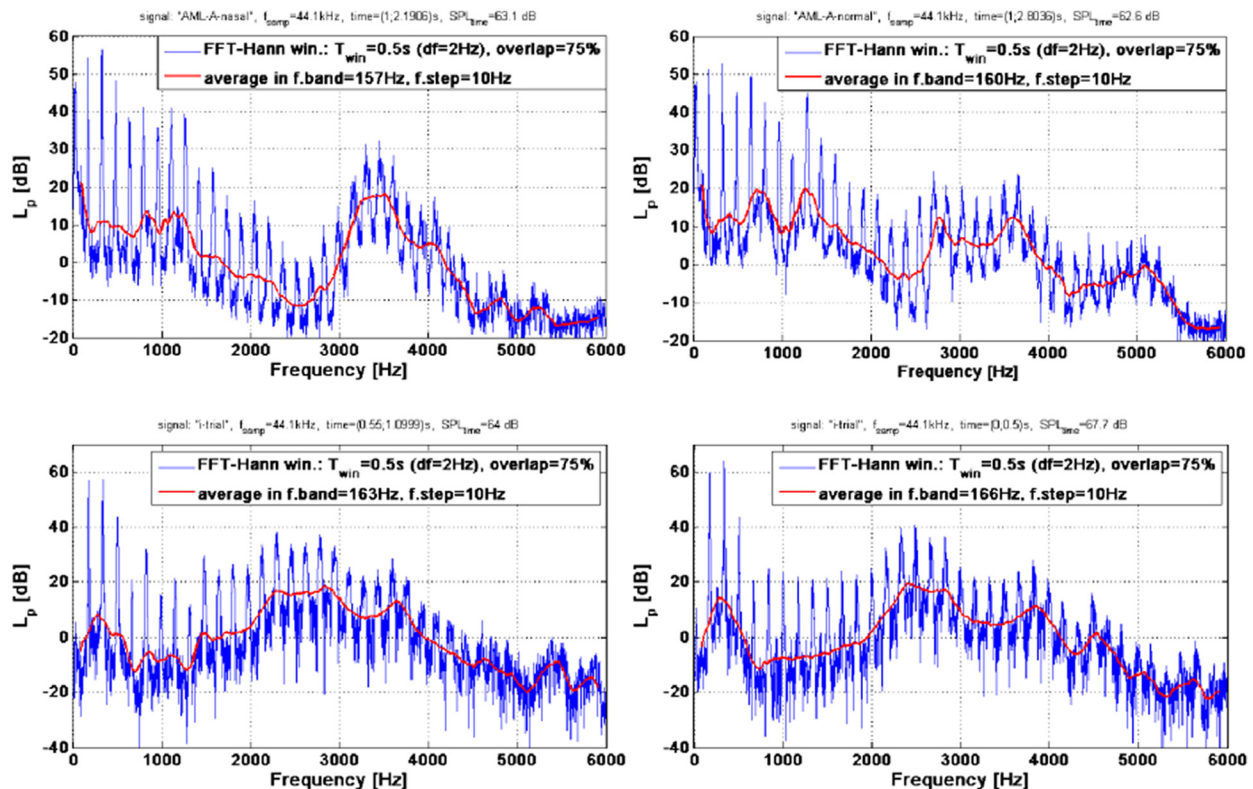


FIG. 13. (Color online) Spectra of the vowels [a:] (upper panel) and [i:] (lower panel) measured *in vivo* in the female subject; nasalized (left) and non-nasalized (right) phonation. The thick red lines show the formants in the spectra after removing the fundamental frequencies ($f_0 = 157$ or 160 Hz for the vowel [a:] and $f_0 = 163$ or 166 Hz for the vowel [i:]) and their higher harmonics [for the method, see Radolf *et al.* (2016) and Horáček *et al.* (2017)].

TABLE V. SPLs measured in the female subject in front of the lips for nasalized and non-nasalized phonation of vowel [a:]. The results are shown for two adjacent frequency intervals.

Vowel [a:] phonation-female	SPL ₁ [dB] at 0.13–2.5 kHz	SPL ₂ [dB] at 2.5–6 kHz	SPL ₁ -SPL ₂ [dB]
nasalized	62.54	45.57	16.97
non-nasalized	61.65	41.0	20.65
difference	0.89	4.57	-3.68

and approached the resonance frequency f_{F3} emphasizing the formant cluster in the frequency range 2–4 kHz in the nasalized vowel [i:] (see Table III and Fig. 13).

Tables V and VI present the SPL levels computed from the sound spectra measured in the female subject for [a:] and [i:] in two consecutive frequency intervals. The frequency intervals are divided at the formant-free zone so that the second interval contains the formants above 1.5 kHz for [a:] and 2 kHz for [i:] (recall Fig. 13). Similarly to the measurement on the physical model of vowel [a:] (see Table IV), the calculated level difference between the first and second interval decreased by ca. 4 dB, making the energy above 2 kHz more prominent for nasal phonation of both vowels.

D. Comparison of the computed and experimental results

Figure 14 compares graphically the vocal tract resonance and formant frequencies for the two vowels obtained from the FEM, physical model and *in vivo* measurements. Compared to FEM, all the resonance frequencies measured on the physical model of the vowel [a:] are shifted down. Both the non-nasalized and nasalized phonations show maximal differences between the FEM and physical model

TABLE VI. SPLs measured identically as in Table V but this time for the vowel [i:].

Vowel [i:] phonation-female	SPL ₁ [dB] at 0.13–2 kHz	SPL ₂ [dB] at 2–6 kHz	SPL ₁ -SPL ₂ [dB]
nasalized	63.74	53.07	10.67
non-nasalized	67.63	53.18	14.45
difference	-3.89	-0.11	-3.78

measurements around 16%–24% for the first two resonance frequencies f_{R1} and f_{R2} . Comparing the FEM results with the *in vivo* measurements of vowel [a:], the maximal difference is -21% for the first resonance frequency f_{R1} in the non-nasalized phonation (see Fig. 14 and Table II). For the non-nasalized vowel [i:], the maximal differences between the FEM and *in vivo* measurements are -22% in f_{R1} and -15% in f_{R3} . For the nasalized [i:], they differ by 23% in f_{R4} (see Table III and Fig. 14).

Qualitatively, the computed resonance frequencies are in a good correspondence with the *in vivo* measurements for both the vowels [a:] and [i:]. The oro-nasal double resonances N_1 and N_2 appeared, in both the nasalized vowels, in the proximity of the first VT resonance R_1 .

The FE model, however, did not capture the sound bone conduction that exists in humans, nor the sound propagation through the hard walls of the VT in the physical model. This could be the reason for some of the observed differences between the FEM, physical model and *in vivo* results.

In order to find out about the influence of vowel nasalization on the formant damping, the half power 3 dB bandwidths of the resonances R_1 and R_2 were investigated. For the non-nasalized vowel [a:] the bandwidths were found to be ca. 55 and 47 Hz, respectively, in the damped FE model

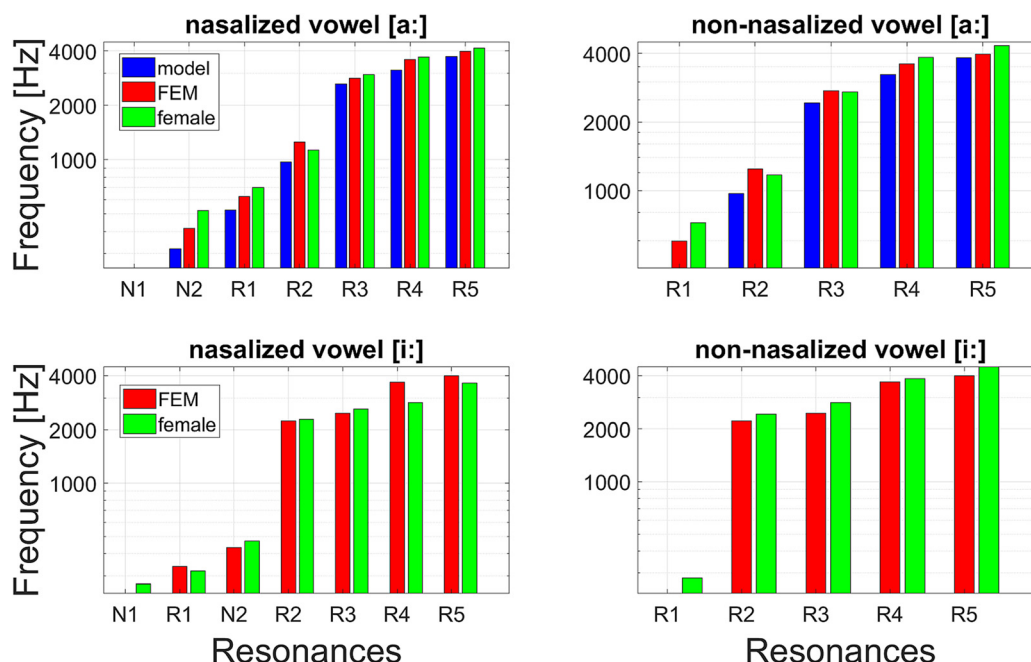


FIG. 14. (Color online) Comparison of results from FE computer modelling, physical model and *in vivo* measurements for the vowels [a:] (upper panel) and [i:] (lower panel) with nasality (left) and without nasality (right).

(Fig. 9) and ca. 32 and 95 Hz in the physical model (Fig. 12). The bandwidth of the resonance R_1 for the non-nasalized vowel [i:] was found to be ca. 64 Hz in the damped FE model (Fig. 10). All these values correspond to the formant bandwidths published by Fant (1956), House and Stevens (1958), or Jovičič (1998). Fant notes that bandwidths 40–250 Hz are probable limits for the first three formants. House and Stevens (1958) measured the bandwidths of first three vocal tract resonances between 53 and 66 Hz for vowel [a:] and between 66 and 76 Hz for vowel [i:]. Jovičič (1998) found the mean formant bandwidths for voiced vowels to be 46 Hz for [i:] and 63 Hz for [a:]. This indicates that the damping properties in the FE model were not overestimated, and that they were comparable to the physical model.

We should note that in the vowel [a:], nasality increased the bandwidth of the first resonance R_1 from 55 to 75 Hz in the FE model, and from 32 to 42 Hz in the physical model. The bandwidth of the first resonance R_1 for the vowel [i:] was not apparently influenced by the nasality.

IV. DISCUSSION

This study explored the effects of nasalization by using CT-based computational modelling, experiments on CT-based physical models, and acoustic recordings from a female subject, from whom the vocal tract CT was originally obtained. The results confirm the previously known findings but also bring new ones.

- (1) The results from FE modelling revealed the four lowest nasal resonances for both vowels [a:] and [i:] to be in the vicinity of the first formant frequencies. The first double peak N_1 at frequencies of ca. 245 and 280 Hz corresponded to oro-nasal resonances of the two maxillary sinuses. The second double peak N_2 at the frequencies of ca. 425 and 472 Hz corresponded to oro-nasal resonances of the two sphenoid sinuses (Tables II and III). These observed resonance frequencies correspond well to the Helmholtz resonance frequencies computed according to Eq. (1), which gives the resonances at 232 and 271 Hz for the left and right maxillary sinuses, respectively, and similarly the resonances at frequencies 419 and 459 Hz for the left and right sphenoid sinuses, respectively.
- (2) When the acoustic energy losses of the VT walls were incorporated in the FE model via the boundary admittance coefficient, none of the acoustic resonances of the NT clearly appeared at the lips for the nasalized vowels [a:] and [i:]. This was in contradiction with the physical model and *in vivo* results in which the nasal resonances remained in the vicinity of the first formant frequency R_1 in the nasalized vowels. Such discrepancy could be explained by partial transmission of sound from the nose cavities to the oral cavity through wall or bone conduction.
- (3) The changes of the computed and measured resonance frequencies due to nasality were qualitatively in

agreement. However, all the FEM computed resonances R_1 – R_5 systematically occurred at higher frequencies (up to 23%–24%) than those observed in the physical model, see Table II. This was true for the nasalized as well as for the non-nasalized vowel [a:]. The differences were likely caused also by the interaction of vibrating walls of the acoustic cavities with the acoustic resonances that were not included in the FE model.

- (4) The low vs high-frequency level differences in the measured spectra for vowel [a:] (see Tables IV and V) showed that the acoustic energy in the region of formants F_3 – F_5 can become more dominant in the sound spectrum due to nasality. This agrees with the findings of Sundberg *et al.* (2007) that nasalization of the vowel [a:] decreases the level of the first formant thus perceptually enhancing the prominence of the harmonics around the singer's formant region near 3 kHz. Similar, but less prominent effect of nasality was found also for the vowel [i:] measured in the female *in vivo* (see Table VI). Here the SPL below 2 kHz decreased while SPL above 2 kHz remained similar. Similar effect was observed for both vowels also in the damped FE models (see Figs. 6 and 10 and Tables II and III). The peak levels of the first resonance frequency R_1 substantially decreased by the nasality while the peak levels of the higher resonances remained similar. For vowel [a:] the resonance frequency R_3 rose towards the resonances R_4 and R_5 thus increasing the acoustic energy in the singer's formant frequency region.
- (5) Many local inner nasal resonances were detected in the nostrils through the FE modelling without acoustic damping of the VT walls. Their acoustic output was very small at the lips, however, and it disappeared for the damped FE model. These resonances thus do not appear to affect markedly the acoustic output out of the mouth but some of these resonances may be sensed as vibrations in the nose, face and head and potentially offer an important tool for optimizing the voice quality.
- (6) The numerical simulations performed in the frequency range 0–6 kHz showed that the shapes of the acoustic inner-nose modes of vibration and the associated resonance frequencies are nearly identical for both the nasalized vowels [a:] and [i:].
- (7) The first two double peak resonances of the left and right acoustic nasal cavity of maxillary and sphenoid sinuses were in the vicinity of the first acoustic resonance R_1 of the VT of the studied vowels. They affected the first formant frequency of the vowels and its peak level. It has been well known that the NT acts as a side branch to the VT and that nasalization leads to appearance of antiresonances in the voice spectrum. In our undamped FE models, the major antiresonances were found in the neighbourhood of the first resonance R_1 for both studied nasalized vowels (recall Figs. 6 and 10). These antiresonances pair up with two resonance peaks that appear as additional double peaks in the computed acoustic pressure response spectra in the VT.

(8) These antiresonances were found between the two resonances N_1 of the left and right maxillary sinuses, between the two resonances N_2 of the sphenoid sinuses and the between two resonances N_3 of the frontal sinuses. At the mouth output, the antiresonances completely disappeared for the damped FE models. Clearly identifiable and a more typical antiresonance was found only below the first resonance R_1 in the measurement on a physical model for nasalized vowel [a:] [recall Fig. 12(c)].

Some limitations of this study are related to geometrical, physical and physiological differences between the FE and the experimental model as well as between the models and humans. In our study, we attempted to eliminate these as much as possible by using data from the same female subject for the FE modelling, physical modelling and *in vivo* investigations.

Our computational model included all seven main human nasal cavities, and the total volume of the acoustic cavities in the NT was larger than the volume of the VT for both vowels. All nasal cavities were accessible to sound excitation. However, the ostia of the paranasal cavities are small and may also be blocked in reality, as known from clinical ear-nose-throat practices.

Resonance frequencies were evaluated in the physical model by using sweep excitation by an earphone. This does not correspond to the reality in humans, but it corresponds well to the computational model excitation. In the computations, the results were obtained from the FE model, where the input airflow pulse signal was applied at the vocal folds location, and the acoustic output pressure was simulated at the lips. Consequently, the boundary conditions given by the geometry of the laryngeal cavities at the inlet to the VT are different in the model, excited by the earphone, and in the human excited by the vibrating vocal folds. Moreover, in the case of the vocal folds the glottis is periodically closing and opening to the subglottic spaces thus constantly changing the boundary conditions.

Some inaccuracy in modelling can also be caused by the complex boundary conditions considered at the lips. The mouth opening and lip protrusion will typically result in differences in the formant frequencies. Recently Vampola *et al.* (2018) used an FE model of the nasalized vowel [a:] with lip protrusion and found that it can shift the frequencies of the first and second VT resonance F_1 and F_2 by ca. 50 and 140 Hz, i.e., by ca. 7%–11%.

In humans, the changes in the velum position are naturally associated with changes in the VT, since when the velum drops it starts filling the pharyngeal space. Second, the palatal and glottal movements are linked via palatal and laryngeal muscles, see, e.g., Kogo (1987). Such complex simultaneous geometrical changes of the acoustic spaces in pharynx, oral and nasal cavities were not considered in any models used in this study.

The laboratory experiments were carried out at the room air temperature of about 22°–25 °C. In humans the

temperature would be around 36 °C, which results in ca. 3% increase of the formant frequencies due to the higher speed of sound in warmer air. Similarly, the effects of yielding walls of human VT were considered in the models only in a simplified form by modifying the damping properties of the VT and NT walls in the FE models. Furthermore, due to very high computation cost, the combined radiation of the sound from the mouth and nostrils out to the open space was not included in our FE models. The mentioned limitations should be addressed in future studies.

V. CONCLUSION

This study introduces a highly detailed CT-based model of the vocal tract coupled with a nasal tract including the maxillary, sphenoid, ethmoid and frontal sinuses. Our experiments confirm the initial findings by Sundberg *et al.* (2007) and show that vowel nasalization leads to decrease of the radiated acoustic energy in the frequency region covering the first and second VT formant, but maintains or even increases the radiated acoustic energy in the region of the third and higher VT formants around 3–5 kHz. These frequencies belong to the region of the highest sensitivity of the human ear and to the region of the singer's/speaker's formant cluster the enhancement of which has been found to be perceived as helpful for good or “resonant” voice quality (Sundberg, 1995; Bele, 2006). Our results therefore indicate that the nasal cavities may also play a beneficial role in producing the “resonant voice” by adding more resonances to the formant cluster around 3–5 kHz. Furthermore, the many resonances in the nasal cavities may potentially lead to vibratory sensations that can help the singers or professional speakers to regulate their voice quality even if these resonances as such would not appear in the acoustic radiated spectra (Yiu *et al.*, 2012; Chen *et al.*, 2014).

Some discrepancies observed among the FE modelling, the physical model and *in vivo* investigations indicate that a more exact modelling of the influence of nasality on voice should include also the effects of the bone conduction of sound in the human head. This was demonstrated by measuring the vibration response of the VT hard walls of the 3D-printed model. Acoustic resonances of the VT were reflected also in the vibration of the VT walls. Therefore, the bone conduction in the human head can also enhance the sensation of the acoustic resonances in different nasal cavities.

ACKNOWLEDGMENTS

The study was supported by the Czech Science Foundation, Grant No. 19-04477S “Modelling and measurements of fluid-structure-acoustic interactions in biomechanics of human voice production.”

Amela, M., Blandin, R., Dabbaghchian, S., Guasch, O., Alías, F., Pelorson, X., Van Hirtum, A., and Engwall O. (2016a). “Influence of lips on the production of vowels based on finite element simulations and experiments,” *J. Acoust. Soc. Am.* **139**(5), 2852–2859.

Amela, M., Dabbaghchian, S., Blandin, R., Guasch, O., Engwall, O., Van Hirtum, A., and Pelorson X. (2016b). “Influence of vocal tract geometry

- simplifications on the numerical simulation of vowel sounds," *J. Acoust. Soc. Am.* **140**(3), 1707–1718.
- Arneta, M., Guasch, O., and Alías F. (2013). "Effects of head geometry simplifications on acoustic radiation of vowel sounds based on time-domain finite-element simulations," *J. Acoust. Soc. Am.* **134**(4), 2946–2954.
- Austin, S. F. (1997). "Movement of the velum during speech and singing in classically trained singers," *J. Voice* **11**, 212–221.
- Bauer, J. (2002). "Die Bedeutung des supraglottischen Raumes für die sängerische Klangformung unter besonderer Berücksichtigung der Funktion des Velum palatinum" ("The importance of the supraglottic space for the singing sound formation with special consideration of the function of the velum palatinum"), Inaugural dissertation, Medizin, Johannes Gutenberg-Universität Mainz, Mainz, Germany.
- Bele, I. V. (2006). "The speaker's formant," *J. Voice* **20**, 555–578.
- Birch, P., Gümöes, B., Prytz S., Karle A., Stavard H., and Sundberg J. (2002a). "Effects of a velopharyngeal opening on the sound transfer characteristics of the vowel [a]," *Speech Music Hear.* **43**, 9–15.
- Birch, P., Gümöes, B., Prytz S., Karle A., Stavard H., and Sundberg J. (2002b). "Effects of a velopharyngeal opening on the sound transfer characteristics of the vowel [a]," in *Speech, Music and Hearing*, KTH Stockholm, Sweden TMH-QPSR. Vol. 43, pp. 9–15.
- Chen, F. C., Ma, E. P. M., and You, E. M. L. (2014). "Facial bone vibration in resonant voice production," *J. Voice* **28**, 596–602.
- Dang, J., and Honda, K. (1996). "Acoustic characteristics of the human paranasal sinuses derived from transmission characteristic measurement and morphological observation," *J. Acoust. Soc. Am.* **100**, 3374–3383.
- Dang, J., and Honda K. (1997). "Acoustic characteristics of the piriform fossa in models and humans," *J. Acoust. Soc. Am.* **101**, 456–465.
- Dedouch, K., Horáček, J., and Švec, J. (2000). "Frequency modal analysis of supraglottal vocal tract," in *Proceedings of the 7th International Conference on Recent Advances in Structural Dynamics*, ISVR, edited by N.S. Ferguson, University of Southampton, 24–27 July, Vol. II, pp. 863–874.
- Delvaux, B., and Howard, D. M. (2014). "A new method to explore the spectral impact of the piriform fossae on the singing voice: Benchmarking using MRI-based 3D-printed vocal tracts," *PLoS One* **9**, e102680.
- Fant, G. (1960). *Acoustic Theory of Speech Production*, 2nd ed. (Mouton, S'Gravenage).
- Fant, G. M. (1956). "On the predictability of formant levels and spectrum envelopes from formant frequencies," in *For Roman Jakobson*, edited by M. Halle, H. Lunt, and M. MacLean (Mouton and Company, Gravenhage, the Netherlands), pp. 109–120.
- Fujita, S., and Honda, K. (2005). "An experimental study of acoustic characteristics of hypopharyngeal cavities using vocal tract solid models," *Acoust. Sci. Technol.* **26**, 353–357.
- Gully, A. J., Daffern, H., and Murphy, D. T. (2018). "Diphthong synthesis using the dynamic 3D digital waveguide mesh," *IEEE/ACM Trans. Audio Speech Lang. Process.* **26**(2), 243–255.
- Hattori, S., Yamamoto, K., and Fujimura, O. (1958). "Nasalization of vowels in relation to nasals," *J. Acoust. Soc. Am.* **30**, 267–274.
- Havel, M., Becker, S., Schuster, M., Johnson, T., Maier, A., and Sundberg, J. (2017). "Effects of functional endoscopic sinus surgery on the acoustics of the sinonasal tract," *Rhinology* **55**, 81–89.
- Havel, M., Kornes, T., Weitzberg, E., Lundberg J. O., and Sundberg, J. (2016). "Eliminating paranasal sinus resonance and its effects on acoustic properties of the nasal tract," *Logoped. Phoniater. Vocol.* **41**(1), 33–40.
- Horáček, J., Radolf, V., and Laukkanen, A.-M. (2017). "Low frequency mechanical resonance of the vocal tract in vocal exercises that apply tubes," *Biomed. Sign. Process. Control* **37**, 39–49.
- House, A. S., and Stevens K. N. (1956). "Analog studies of the nasalization of vowels," *J. Speech Hear Disord* **21**, 218–232.
- House, A. S., and Stevens, K. N. (1958). "Estimation of formant bandwidths from measurements of transient response of the vocal tract," *J. Speech Hear. Res.* **1**(4), 309–315.
- Huffman M. K. (1990). "The role of F1 amplitude in producing nasal percepts," *J. Acoust. Soc. Am.* **88**, S54.
- Jovičić, S. T. (1998). "Formant feature difference between whispered and voiced sustained vowels," *Acta Acust. Acust.* **84**, 739–743.
- Kogo, M., Nishio, J., Matsuya, T., Hamamura, Y., and Miyazaki T. (1987). "Coordination of the levator veli palatini and intrinsic laryngeal muscles: An evoked electromyographic study in the dog," *Cleft Palate J.* **24**(2), 119–125.
- Kozaki-Yamaguchi, Y., Suzuki, N., Fujita, Y., Yoshimasu, H., Akagi, M., and Amagasa, T. (2005). "Perception of hypernasality and its physical correlates," *Oral Sci. Int.* **2**(1), 21–35.
- Motoki K. "Three-dimensional acoustic field in vocal-tract," *Acoust. Sci. Technol.* **23**, 207–212 (2002).
- Nieuwenhof, B. v. d., and Coyette, J.-P. (2001). "Treatment of frequency-dependent admittance boundary conditions in transient acoustic finite/infinite-element models," *J. Acoust. Soc. Am.* **110**(4), 1743–1751.
- Radolf, V., Horáček, J., Dlask, P., Otčenášek, Z., Geneid, A., and Laukkanen, A.-M. (2016). "Measurement and mathematical simulation of acoustic characteristics of an artificially lengthened vocal tract," *J. Sound Vib.* **366**, 556–570.
- Richards, E. J., and Mead, D. J. (1960). *Noise and Acoustic Fatigue in Aeronautics* (Wiley, London).
- Sundberg, J. (1987). *The Science of the Singing Voice* (Northern Illinois University Press, DeKalb, IL).
- Sundberg, J. (1995). "The singer's formant revisited," KTH, Department of Speech, Music and Hearing, Quarterly Progress and Status Report, Vol. 36(2–3).
- Sundberg, J., Birch, P., Gümöes, B., Stavard, H., Prytz, S., and Karle, A. (2007). "Experimental findings on the nasal tract resonator in singing," *J. Voice* **21**(2), 127–137.
- Švec, J. G., Herbst, C., Havlík, R., Horacek, J., Krupa, P., Lejska, M., and Miller, D. G. (2008). "Singer's formant: Preliminary results of MRI and acoustic evaluations of singers," in *Proceedings Interaction and Feedbacks 2008*, edited by I. Zolotarev (Institute of Thermomechanics, Prague).
- Takemoto, H., Adachi, S., Mokhtari, P., and Kitamura T. (2013). "Acoustic interaction between the right and left piriform fossae in generating spectral dips," *J. Acoust. Soc. Am.* **134**(4), 2955–2964.
- Takemoto, H., Mokhtari, P., and Kitamura T. (2010). "Acoustic analysis of the vocal tract during vowel production by finite-difference time-domain method," *J. Acoust. Soc. Am.* **128**(6), 3724–3738.
- Tanner, K., Roy, N., Merrill, R. M., and Power D. (2005). "Velopharyngeal port status during classical singing," *J. Speech Lang. Hear. Res.* **48**, 1311–1324.
- Titze, I. R. (1994). *Principles of Voice Production* (Prentice Hall, Upper Saddle River, NJ).
- Titze, I. R. (2006). *The Myoelastic Aerodynamic Theory of Phonation* (National Centre for Voice and Speech, Iowa City, IA).
- Titze, I. R., Baken, R. J., Bozeman, K. W., Granqvist, S., Henrich, N., Herbst, C. T., and Wolfe, J. (2015). "Toward a consensus on symbolic notation of harmonics, resonances, and formants in vocalization," *J. Acoust. Soc. Am.* **137**(5), 3005–3007.
- Vampola, T., Horáček, J., Laukkanen, A. M., and Švec, J. G. (2015a). "Human vocal tract resonances and the corresponding mode shapes investigated by three-dimensional finite-element modelling based on CT measurement," *Logoped. Phoniater. Vocol.* **40**, 14–23.
- Vampola, T., Horáček, J., and Švec, J. G. (2008a). "FE modeling of human vocal tract acoustics. Part I: Production of Czech vowels," *Acta Acust. Acust.* **94**, 433–447.
- Vampola, T., Horáček, J., and Švec J. G. (2015b). "Modeling the influence of piriform sinuses and valleculae on the vocal tract resonances and anti-resonances," *Acta Acust. Acust.* **101**, 594–602.
- Vampola, T., Horáček, J., Vokral, J., and Cerny, L. (2008b). "FE modeling of human vocal tract acoustics. Part II: Influence of velopharyngeal insufficiency on phonation of vowels," *Acta Acust. Acust.* **94**, 448–460.
- Vampola, T., Štorkán, J., Horáček J., and Radolf, V. (2018). "Numerical investigation of acoustic characteristics of 3D human vocal tract model with nasal cavities," in *Proceedings of 24th International Conference Engineering Mechanics 2018*, Paper #304, Svratka, Czech Republic, May 14–17, pp. 893–896.
- Vennard W. (1964). "An experiment to evaluate the importance of nasal resonance in singing," *Folia Phoniater. (Basel)* **16**, 146–153.
- Yiu, E. M. L., Chen, F. C., Lo, G., and Pang, G. (2012). "Vibratory and perceptual measurement of resonant voice," *J. Voice* **26**, 675.e13–675.e19.