RESONANCE FREQUENCIES IN VOCAL TRACT LIKE WAVE-GUIDES

Al Oualid Eliraki, Xavier Pelorson, Annemie van Hirtum

29th International Congress
on Sound and Vibration
The annual congress of
the International Institute
of Acoustics and Vibration (IIAV)

9–13 July 2023

ICSV29

Annual Congress of the International Institute of Acoustics and Vibration (IIAV)

RESONANCE FREQUENCIES IN VOCAL TRACT LIKE WAVEGUIDES

Al Oualid ELIRAKI, Xavier Pelorson and Annemie Van Hirtum
LEGI, UMR CNRS 5519, Grenoble Alpes University, France
e-mail: al.oualid.eliraki@univ-grenoble-alpes.fr, annemie.vanhirtum@univ-grenoble-alpes.fr

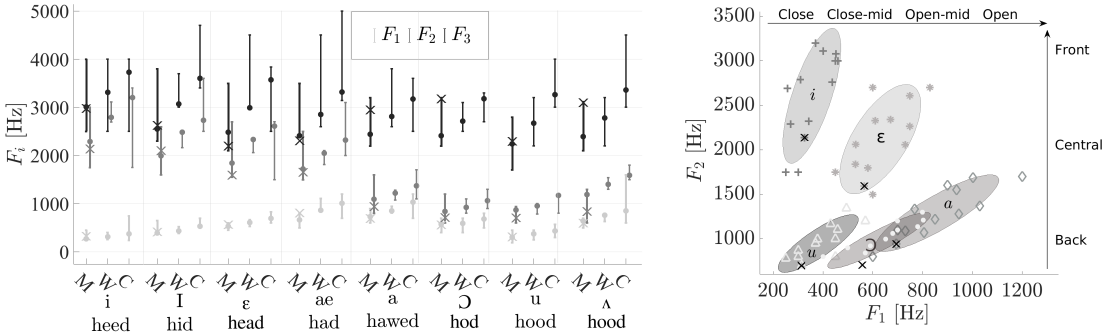
Abstract

Human voice production involves neuromuscular control of the vocal tract articulators resulting in complex time-varying vocal tract geometries associated with different speech sounds. In order to study and measure the acoustics of a dynamic vocal tract, a simplified two-parameter vocal tract model is investigated, which is suitable to be exploited in physical studies using a dynamic mechanical vocal tract replica allowing to impose both model parameters which are the position and degree of constriction. Acoustic resonance frequencies up to 5 kHz are modelled for the full parameter space. Results are compared to formants of human speakers for the simplified two-parameter vocal tract model including selected vocal tract geometries derived from MRI data. It is found that the agreement for the first three resonance frequencies for 10 vowels is within 10% since the mean relative error varies from 0.9% up to 9.1%. Overall, the simplified two-parameter vocal tract model allows to reproduce resonances up to 5 kHz. Advantages and limitations of the model are discussed.

Keywords: Acoustics; Waveguide dynamics; Vowels.

1. Introduction

The human vocal tract extends from the glottis at its back to the lips at its front. Sound propagation inside the vocal tract is an important area of study in speech acoustics. In this paper, we approach the problem from the point of view of vowels, and investigate the relationship between acoustic and geometrical waveguide features. Indeed, the vocal tract is a complex waveguide which is shaped dynamically through articulation. As such, even for a single sustained vowel, cross-section areas throughout the vocal tract can vary from a few up to several hundred mm^2 [6]. Acoustic resonance frequencies or formants issued by the vocal tract waveguide vary considerably depending on the vocal tract geometry. The first formant frequencies (F_1 up to F_3), which are crucial to distinguish between vowels, vary in a frequency band up to about 5 kHz and are reported to differ considerably between speakers. The inter-speaker variability observed for adult males, adult women and children is illustrated in Fig. 1. Mean formant frequencies from Table II of the reference paper by Peterson and Barney [5] for a total of 76 speakers (33 men, 18 women and 15 children) are plotted as well as the inter-speaker range obtained considering the maximum and minimum formants for 43 adult males, 53 adult women and 43 children from [1, 2, 3, 4, 5, 6, 8]. The range for each single formant shows the variability associated with each vowel, which can amount to the same order of magnitude as the formant frequency. The size of the (F_1, F_2) space associated with each vowel is then indicated in Fig. 1(b) with an ellipse whose axes and orientation are obtained following a principal component analysis of the co-variance matrix of 13 (M,



(a) frequency ranges for F_1, F_2, F_3

(b) formant space (F_1, F_2)

Figure 1: Vowel formants F_i observed for American English speakers (43 adult men (M), 53 adult woman (W), 43 children (C)) [1, 2, 3, 4, 5, 6, 8], values (male adult) of Table II in [6] ($\times$) are indicated as a reference: a) frequency range (bars) and mean values ($\bullet$, Table II [5]), b) formant space from 13 (F_1, F_2) pairs (M, W, C) for each vowel with influence of constriction degree (F_1) and location (F_2).

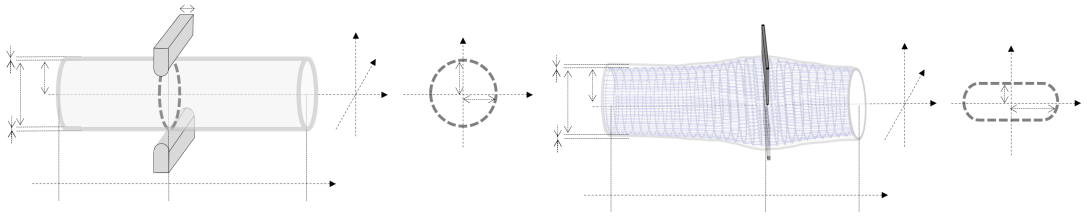
W and C) formant pairs (F_1, F_2) [1, 2, 3, 4, 5, 6, 8]. Phonetic textbooks generally relate an increase of F_1 to an articulatory release of a vocal tract area constriction (close towards open) and an increase of F_2 to an articulatory forward shift of a vocal tract constriction (back towards front) [1]. Following this reasoning, the vowel space spanned by formant frequencies (F_1, F_2) plotted in Fig. 1(b) reflects both the degree and the position of a main vocal tract constriction. Using this reasoning a simplified geometrical waveguide model capable to impose these geometrical constriction features should be able to reconstruct the formant space for vowels up to some degree. Therefore, in this work, it is aimed to investigate the relationship beween formants and vocal tract geometry using a simplified two-parameter geometrical deformable waveguide model proposed in [7]. It is sought to what extent this model in combination with a one-dimensional (1D) acoustic model is capable to explain the extent of the vowels space for formants up to 5 kHz. The model outcome in terms of acoustic resonance frequencies is evaluated against formants given in literature for sustained vowels uttered by human speakers. The ability of simplified waveguides to reproduce vowel formants is discussed.

2. Methodology

2.1 Simplified vocal tract model and 1D acoustic model

2.1.1 Simplified two-parameter vocal tract model

A simplified two-parameter vocal tract model is based on the analytical compressed elastic cylindrical waveguide model proposed and validated in [7]. The undeformed uniform cylindrical waveguide of length L_0 and wall thickness d is positioned along the x-direction and its uniform cross-section along the yz-plane is characterised by the internal diameter D with circular cross-section area $A_0 = \pi D^2/4$. The waveguide is deformed by squeezing it locally between a pincher consisting of two rounded symmetrical bars (diameter 6.4 mm) positioned parallel to the y-direction as depicted in Fig. 2(a). Pinching squeezes the waveguide locally so that a constriction is induced along the longitudinal x-direction reducing the cross-sectional area along the deformed portion of the waveguide. The geometrical effect of squeezing can be accurately described as a function of two geometrical parameters, namely constriction position x_c and constriction degree $\mathcal{P}(\beta) = 1 - \beta$ with $\beta = b_{x_c}/b_0$, $b_0 = D/2$ and $b_{x_c} \leq b_0$ denoting the minor axis of the cross-section at $x = x_c$ corresponding to its half-length along the z-axis. Increasing the squeezing



(a) without constriction, $\mathcal{P} = 0$ % (b) with constriction, $\mathcal{P} = 88$ %

Figure 2: Simplified two-parameter $(x_c, \mathcal{P})$ geometrical vocal tract model as a cylindrical waveguide of length L_0 squeezed with a single pincher (two parallel bars) at constriction position x_c imposing constriction degree $\mathcal{P} = 1 - b_{x_c}/b_0$: a) $\mathcal{P} = 0$ %, b) $\mathcal{P} = 88$ % and stadium-shaped cross-sections.

effort increases thus $\mathcal{P}$ from $\mathcal{P}(\beta = 1) = 0$ ($b_{x_c} = b_0$) for the undeformed waveguide to $\mathcal{P}(\beta < 1) > 0$ ($b_{x_c} < b_0$) for the deformed waveguide until at most $\mathcal{P}(\beta = 0) = 1$ ($b_{x_c} = 0$) for complete closure. The deformation is quasi-symmetrically around the constriction position x_c over a length of $8 \times b_0$ (or $4 \times D$) and deformed cross-sections can be accurately modelled analytically considering as a stadium ring with rounded edges as illustrated in Fig. 2(b). The cross-sectional area $A(x)$ and major axis $a(x)$ as a function of longitudinal x-position is then obtained for imposed parameter set $(\mathcal{P}, x_c)$ as

$$A(x; \mathcal{P}, x_c) = \pi\beta^2(x; \mathcal{P}, x_c)b_0^2 + 2A_0\beta(x; \mathcal{P}, x_c)\big(1 - \beta(x; \mathcal{P}, x_c)\big), \tag{2.1.1}$$

$$a(x; \mathcal{P}, x_c) = \beta(x; \mathcal{P}, x_c)b_0 + \frac{\pi}{2}b_0\big(1 - \beta(x; \mathcal{P}, x_c)\big) \tag{2.1.2}$$

with $\beta(x; \mathcal{P}, x_c) = b(x; \mathcal{P}, x_c)/b_0$ obtained from minor axis $b(x; \mathcal{P}, x_c)$ as

$$b(x; \mathcal{P}, x_c) = b_0 \left(1 - \mathcal{P}\left(\frac{(x - x_c)^2}{\alpha_{\mathcal{P}}^2} + 1\right)^{-1}\right) \tag{2.1.3}$$

with $\alpha_{\mathcal{P}}$ denoting half peak width at half amplitude of the deformed longitudinal outer edge approximated as a quadratic function of the constriction degree $\alpha_{\mathcal{P}}(\mathcal{P}) = 48\mathcal{P}^2 - 70\mathcal{P} + 39$. As illustrated in Fig. 1(b), the cross-section shape at each x-position along the longitudinal waveguide axis can be modelled from the minor $b(x)$ and major $a(x)$ axes. The outlined simplified and analytical two-parameter geometrical vocal tract waveguide model allows thus naturally a quantitative and systematic assessment of the influence of the constriction position x_c and constriction degree $\mathcal{P}$ on the acoustic waveguide resonances. The model is thus particularly suitable to investigate to what extent constriction parameters $(x_c, \mathcal{P})$ can reproduce the formants illustrated in Fig. 1. In the following, geometrical waveguide characteristics $D = 25$ mm, $d = 3$ mm and $L_0 = 171$ mm are held constant. The waveguide outlet at the lips ($x = 0$) is set in an rectangular flat baffle (37 cm $\times$ 35 cm and thickness 5 mm) in order to represent the face.

2.1.2 Acoustic model

The lowest cut-on frequency f_c associated with the first higher order acoustic propagation mode depends on the largest dimension of the waveguide cross-sections as well as on the cross-section shape [1]. The largest cross-section dimension corresponds to the half-width of the stadium ring at the constriction, *i.e.* major axis $a(x = x_c)$ denoted a_c in short (Eq. (2.1.1)). For very small (near 0) and very large (near 1) constriction degrees $\mathcal{P}$, the stadium ring cross-section shape can be approximated as circular and rectangular respectively [7]. For these shapes, cut-on frequencies f_c as a function of major axis a_c are

given analytically as $f_c = \alpha_\circ c/a_c$ (circular, small Pd) and $f_c = \alpha_\square c/a_c$ (rectangular, large $\mathcal{P}$) with with $c = 340$ m/s denoting the sound speed in air at room temperature (21 °C) and $\alpha_\circ = 0.92/\pi$ and $\alpha_\square = 1/4$. Accounting for the waveguide dimensions, this implies that the lowest cut-on frequency decreases from $f_c = 7.9$ kHz ($\mathcal{P} = 0$ %, circular) to $f_c = 4.3$ kHz (rectangular, $\mathcal{P} = 98$ %). For vowels (Fig. 1(a)) the first three formants $F_{i\in\{1,2,3\}}$ are smaller than f_c, thus a plane wave acoustic model approach, neglecting higher order mode propagation, can be used to evaluate to which extent acoustic resonances frequencies of the simplified two-parameter vocal tract model can explain vowel formants and their variability.

For plane waves, the acoustic field inside the waveguide (Fig. 2) at each longitudinal position x and time t is given as $p(x,t) = P(x)e^{j\omega t}$ with radian wave frequency $\omega = 2\pi f$ and complex-valued pressure amplitude $P(x) = P_+e^{-jkx} + P_-e^{jkx}$ obtained as the sum of an acoustic wave with amplitude P_+ traveling in the positive x-direction and an acoustic wave with amplitude P_- traveling in the negative x-direction where k denotes the wavenumber. For a plane wave propagating in the positive direction, the reflection coefficient R_x at position x is the ratio of the reflected and incident wave $R_x = P_-/P_+e^{2jkx}$ from which the impedance Z_x at position x is obtained as $Z_x = Z_c(1+R_x)/(1-R_x)$ with characteristic impedance $Z_c = \rho c/A(x)$ and air density $\rho = 1.2$ kg/m^3. At the open waveguide outlet ($x = 0$), the assumption of a plane wave front underlying R_x no longer holds. The reflection coefficient at the outlet R_0 is then obtained from the radiation impedance as $R_0 = -|R_0|e^{-2jk\delta}$ with δ the real part of the end correction defining the effective acoustic waveguide length. The reflection coefficient Z_0 is obtained from R_0 following Z_x. The real part of the radiation impedance represents the radiation loss. In the case of a flanged outlet condition, corresponding to the rectangular baffle at $x = 0$, the modulus of the reflection coefficient R_0 and end correction δ are approximated for $kb_0 < 3.5$ or wave frequencies $f < 15$ kHz as

$$|R_0| = \frac{1 + 0.323(kb_0) + 0.077(kb_0)^2}{1 + 0.323(kb_0) + (1 - 0.077)(kb_0)^2} \tag{2.1.4}$$

$$\frac{\delta}{b_0} = 0.8216\left(1 + \frac{0.77(kb_0)^2}{1 + 0.77(kb_0)}\right)^{-1}. \tag{2.1.5}$$

Using a transmission line principle for impedance matching, the impedance between neighbouring points x_i and x_{i+1} with $\Delta x = x_{i+1} - x_i$ along the positive x-directions yields

$$\frac{Z_i}{Z_c} = \frac{j\tan k\Delta x + Z_{i+1}/Z_c}{1 + j\dfrac{Z_{i+1}}{Z_c}\tan k\Delta x}. \tag{2.1.6}$$

As outlet radiation impedance Z_0 is known, the impedance at each x position can be determined when the one-dimensional waveguides area function $A(x)$ is known. The impedance at each x-position along the waveguide yields $Z(x) = P(x)/U(x)$ with acoustic volume flow rate $u(x,t) = U(x)e^{j\omega t}$. Applying the transmission line principle results then in the transfer matrix

$$\begin{bmatrix} P(x_i) \\ U(x_i) \end{bmatrix} = \begin{bmatrix} \cos k\Delta x & jZ_c \sin k\Delta x \\ jZ_c^{-1}\sin k\Delta x & \cos k\Delta x \end{bmatrix} \times \begin{bmatrix} P(x_{i+1}) \\ U(x_{i+1}) \end{bmatrix}. \tag{2.1.7}$$

Thus, from an imposed $U(x = -L)$ at the waveguide's inlet, $P(x = -L)$ is found and the acoustic field inside the waveguide can be determined from Eq. 2.1.7. For each frequency, transfer function H

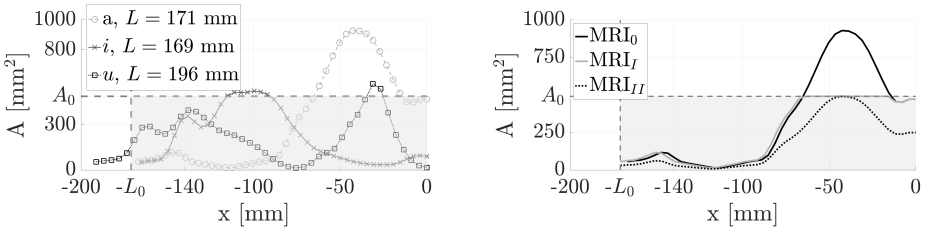
$$H = 20\log_{10}\left|\frac{U_0}{U_{-L}}\right|. \tag{2.1.8}$$

between inlet ($x = -L$) and outlet ($x = 0$) allows to determine resonance peak frequencies $\hat{F}$. The model outcome is quantitatively evaluated by considering the mean relative absolute error with respect to N formants F_i reported in [6] for a male adult with formant index i in the set $\{i_1, \ldots, i_n\}$ as

$$Er_{i_1 \ldots i_n} = \frac{1}{N} \sum_{i \in \{i_1, \ldots, i_n\}} \frac{|\hat{F}_i - F_i|}{F_i}. \tag{2.1.9}$$

2.2 MRI-based vocal tract waveguides

Realistic vocal tract 1D cross-sections area functions $A(x)$ obtained from MRI imaging and associated formants ($\times$ in Fig. 1) for 10 different sustained vowel utterances (6 up to 10 seconds) from Table II in [6] are used as a reference. Original area functions $A(x)$, as plotted in Fig. 3(a) and denoted MRI_0, are all characterised by at least one major constriction. The influence of such a constriction can be investigated with the simplified two-parameter vocal tract model presented in Section 2.1.1. Nevertheless, due to the constant overall dimensions of the last $(-L_0, A_0)$, cross-section areas $A(x)$ can be larger than the maximum cross section area A_0 of the simplified vocal tract model. The same way the length of the vocal tract L can be either shorter or longer than L_0. To relate vocal tract geometries MRI_0 to the simplified two-parameter vocal tract model two geometrical transformations are assessed forcing the MRI area function to fit within the unconstricted ($\mathcal{P} = 0$ %) simplified two-parameter vocal tract domain. A first type of reshaped vocal tract geometries labelled MRI_I is obtained by rescaling the longitudinal dimension with a factor $\alpha_x = L_0/L$ so that the rescaled length matches L_0. In addition, cross-section areas greater than A_0 are set to A_0. A second type of reshaped geometries labelled MRI_{II} is obtained by scaling the x-dimension again with α_x and cross-section areas with factor $\alpha_A = A_0/\max(A(x))$. Rescaled MRI_{II} vocal tract lengths matches thus L_0 and the maximum cross-section area matches A_0. All three (MRI_0, MRI_I, MRI_{II}) MRI-related vocal tract area functions $A(x)$ are illustrated in Fig. 3(b) for vowel textbfa. Scaling with α_L (MRI_I and MRI_{II}) alters the constriction position x_c, which will either shift towards the lip (x_c decreases) when $\alpha_x < 1$ ($L > L_0$) or towards the glottis (x_c increases) when $\alpha_x < 1$ ($L < L_0$). On the other hand scaling with α_A (MRI_{II}) affects constriction degree $\mathcal{P}$, which will either decrease for $\alpha_A > 1$ ($A_0 > \max(A(x))$) or increase for $\alpha_A < 1$ ($A_0 < \max(A(x))$). Thus, transformations might affect formants, but favour use of the simplified two-parameter vocal tract model in order to investigate observed differences between geometries and between MRI_I and MRI_{II} in particular.



(a) MRI_0 for **a**, **i**, **u**

(b) vowel **a**: MRI_0, MRI_I, MRI_{II}

Figure 3: Area functions $A(x)$ from MRI imaging [6] and simplified two-parameter vocal tract domain $(-L_0, A_0)$ for $\mathcal{P} = 0$ % (gray shaded): a) original MRI_0 , b) MRI_0 and reshaped MRI_I and MRI_{II}.

3. Results

3.1 Resonance space for MRI-based vocal tract waveguides

Model results obtained when applying the 1D acoustic model approach outlined in Section 2.1.2 to waveguides obtained using vocal tract area functions MRI_0, MRI_I and MRI_{II} (Section 2.2) are plotted in Fig. 4 and scaling factors (α_x, α_A) used to obtain MRI_I and MRI_{II} from the original area function MRI_0 from [6] are indicated. Modelled transfer functions $H(f)$ for vowels **i, ɛ, a, u** are plotted in Fig. 4(a)-4(d) and resonance peaks are indicated. Depending on the vowel, transfer functions $H(f)$ obtained for MRI_0, MRI_I and MRI_{II} are a near match (**i** or **ɛ**) or exhibit significant differences (**a** or **u**). The modelled formant space $(\hat{F}_1, \hat{F}_2)$ for MRI_0, MRI_I and MRI_{II} as well as literature values (F_1, F_2) associated with MRI_0 for 10 vowels are plotted in Fig. 4(e). Although that for each vowel all four data points remain in the same region differences are apparent. For MRI-based geometries, the discrepancy might be related to the magnitude of scaling factors (α_x, α_A) (see *e.g.* **a** or **u**). For MRI_0 and literature values reported in [6] the observed discrepancy can be partly explained by a mismatch in tempeature and thus sound speed. The discrepancy is further quantified considering Er_1, Er_2, Er_3 for each vowel in Fig. 4(f)- 4(h). Relative error values Er_2 and Er_3 are less than 10 % whereas Er_1 can increase up to 23 %. The overall mean relative error Er_{123} varies from 0.9 % up to 9.1 %. Overall, the discrepancy is thus acceptable for all geometries including MRI_I and MRI_{II} which fit within the largest $A(x)$ domain occupied by the simplified two-parameter vocal tract model. This leads to the statement that formants $\hat{F}_{i\in 1,2,3}$ can be modelled within this domain. In how far imposing a single constriction with parameters $(x_c, \mathcal{P})$ allows to explain observed formants and associates tendencies is sought in the next section.

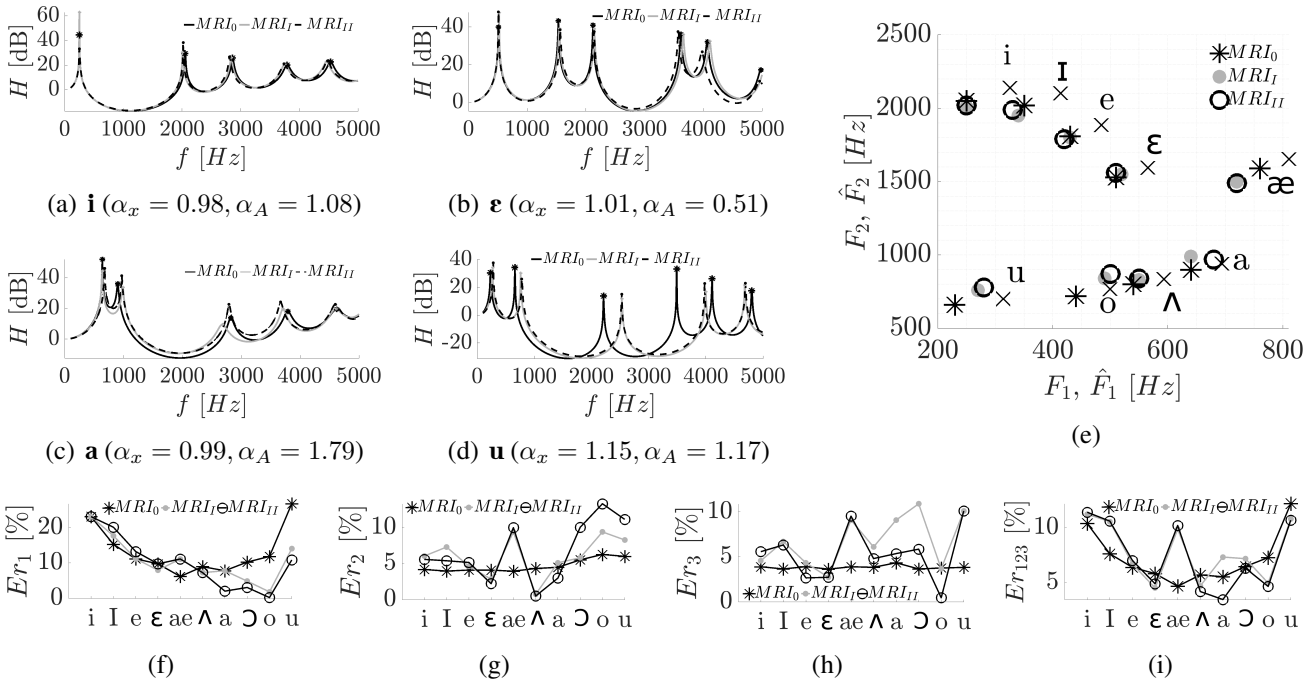
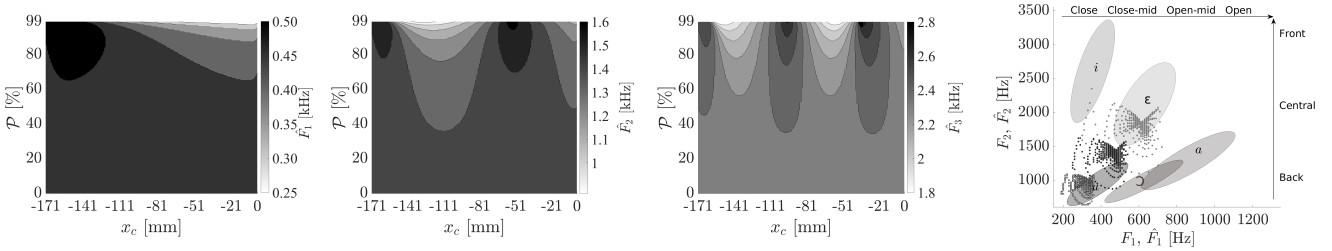


Figure 4: Model outcome for vocal tract area functions MRI_0, MRI_I and MRI_{II}: a-d) transfer functions $H(f)$ with resonance peaks (•), e) modelled $(\hat{F}_1, \hat{F}_2)$ (✳, ●, ○) and (F_1, F_2) (×) from literature [6], f-h) relative accuracy Er_1, Er_2 and Er_3 and i) mean relative accuracy Er_{123}.

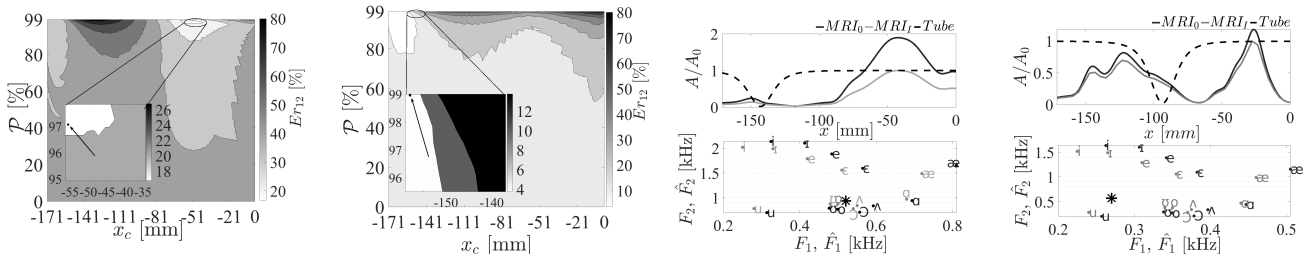
3.2 Resonance space for simplified two-parameter vocal tract model

Modelled resonance frequencies $\hat{F}_1(x_c, \mathcal{P})$, $\hat{F}_2(x_c, \mathcal{P})$ and $\hat{F}_3(x_c, \mathcal{P})$ for the simplified vocal tract model described in Section 2.1.1 are shown in Fig. 5 as a function of its two-parameter space $(x_c, \mathcal{P})$ with $-L_0 \leq x_c \leq 0$ and $0\% \leq \mathcal{P} < 100\%$. Modelled $\hat{F}_1$ vary between 250 Hz and 500 Hz (overall increase of 100 %), modelled $\hat{F}_2$ between 900 Hz and 1600 Hz (overall increase of 77 %) and modelled $\hat{F}_3$ between 1800 Hz and 2800 Hz (overall increase of 55 %). Consequently, vowels with formants outside these ranges can not be reproduced with the simplified two-parameter vocal tract model using the chosen conditions. This is illustrated in Fig. 5(d) showing the limited resonance space $(\hat{F}_1, \hat{F}_2)$, associated with the complete parameter space $(x_c, \mathcal{P})$, compared to vowel formant space. The influence of varying waveguide length $L = L_0$ on the resonance space is shown for a shorter ($L = 130$ mm) and longer ($L = 250$ mm) waveguide. Smallest and greatest resonance frequencies $\hat{F}_{i\in1,2,3}$ are all observed in the parameter region associated with large constriction degrees $\mathcal{P}$. In this high $\mathcal{P}$-range, $\hat{F}_1$ (Fig. 5(a)) decreases with constriction position x_c which is a tendency not illustrated in Fig. 1(b) where x_c is solely associated with F_2. Variation of $\hat{F}_1$ occurs for rather extreme $\mathcal{P} > 90$ %. Consequently, extreme x_c near the outlet and inlet as well as extreme constriction degrees $\mathcal{P} > 90$ % are needed to exploit the full range of $\hat{F}_1$. The increase of $\hat{F}_1$ for decreasing constriction degree $\mathcal{P}$ (close to open) pointed out in Fig. 1(b) is retrieved for x_c away from the inlet (17 %) up to the waveguide's outlet (83 %). However, for the x_c-range near the outlet (17%), the tendency is reversed as $\hat{F}_1$ decreases for decreasing $\mathcal{P}$, which thus contradicts the thumb rule in Fig. 1(b). In the high $\mathcal{P}$-range, $\hat{F}_2$ (Fig. 5(b)) exhibits two successive maxima and minima for increasing x_c. Nevertheless, when not considering the extrema (10 %) near the inlet and outlet, $\hat{F}_2$ increases with x_c (back to front). This tendency follows thus the thumb rule indicated in Fig. 1(b). Overall, the tendency is again most prominent for large $\mathcal{P}$ although that for some x_c the finding holds for $\mathcal{P} > 40$ %. In the high $\mathcal{P}$-range, $\hat{F}_3$ (Fig. 5(c)) exhibits three successive maxima and minima for increasing x_c. Consequently, alternating tendencies for increasing x_c occur as near the outlet (14 %), $\hat{F}_3$ decreases, next $\hat{F}_3$ increases towards the center (36 %), followed again by a decrease of $\hat{F}_3$ (26 %) where-after $\hat{F}_3$ increases towards the outlet (24%). This increase or decrease of $\hat{F}_{i\in1,2,3}$ with x_c for high enough $\mathcal{P}$ is not necessarily linear, but can be concentrated in a compact parameter region. For constriction degrees $\mathcal{P}$ lower than the high $\mathcal{P}$-range, the variation of resonance frequencies with either x_c or $\mathcal{P}$ is limited, suggesting that a multitude of vocal tract geometries are associated with a same set $(\hat{F}_1, \hat{F}_2, \hat{F}_3)$. From the viewpoint of dynamic resonance frequency variation by two-parameter articulation, this might be a favourable feature. The relative error for the full two-parameter space $Er_{12}(x_c, \mathcal{P})$ of the simplified vocal tract model is quantified for each vowel as illustrated



(a) $\hat{F}_1(x_c, \mathcal{P})$ (b) $\hat{F}_2(x_c, \mathcal{P})$ (c) $\hat{F}_3(x_c, \mathcal{P})$ (d) $(\hat{F}_1, \hat{F}_2)$, influence of L

Figure 5: Modelled resonance frequencies $\hat{F}_{i\in1,2,3}$ for the full space $(x_c, \mathcal{P})$ of the simplified two-parameter vocal tract model: a) $\hat{F}_1$, b) $\hat{F}_2$, c) $\hat{F}_3$, d) $(\hat{F}_1, \hat{F}_2)$ on formant space (F_1, F_2) of Fig. 1(b) for waveguide lengths $L = 250$ mm (○), $L = L_0$ (●, 171 mm) and $L = 130$ mm (●).



(a) **i**, 17 %: (-52 mm, 98 %) (b) **ɛ**, 4 %: (-159 mm, 99 %) (c) **a**, 17 %: (-142 mm, 99 %) (d) **u**, 21 %: (-93.5 mm, 99 %)

Figure 6: a,b) Error $Er_{12}(x_c, \mathcal{P})$ and simplified vocal tract selection for $\min(Er_{12})$ (arrow in zoom), c,d) $A(x)/A_0$ (top) and $(\hat{F}_1, \hat{F}_2)$ (bottom) for selection (Tube, $*$), MRI_0 ($\bullet$), MRI_I ($\bullet$), vowels (F_1, F_2) [6].

in Fig. 6(a)-6(b)). The parameter set minimising the error (zoom near $\min Er_{12} < 28$ %) is attributed to each vowel resulting in a simplified vocal tract selection so that modelled resonances $\hat{F}_{i\in1,2,3}$ can be compared to those obtained for MRI_0 , MRI_I and formants $F_{i\in1,2,3}$ [6] as illustrated in Fig. 6(c)-6(d). In general, x_c is shifted towards the inlet suggesting that the constriction onset primes.

4. Conclusion and perspectives

It is investigated to what extent a simplified two-parameter vocal tract model accounting for a single vocal tract constriction within a parameter space limited by its maximum area and length is able to reproduce vowel formants up to 5 kHz. MRI-based vocal tract geometries allowed to show that the parameter space is sufficient to approximate formants. Although that simplified geometries allow to retrieve thumb rules concerning the influence of constriction position and degree on resonance frequencies, resonance spaces overlap only with a subset of the formant spaces retrieved for vowels.

Acknowledgements

Funding by Fully3DTalkingHead (ANR-20-CE23-0008-03) project.

REFERENCES

1. M. J. Crocker. *Handbook of acoustics*. John Wiley & Sons, 1998.
2. M. C. J. Hillenbrand, L.A. Getty and K. Wheeler. Acoustic characteristics of American English vowels. *J. Acoust. Soc. Am.*, 97(5):3099–3111, 1995.
3. D. A. J.T. Eichhorn, R.D. Kent and H. Vorperian. Effects of aging on vocal fundamental frequency and vowel formants in men and women. *J Voice*, 32(5):644–e1, 2018.
4. D. Kewley-Port and C. S. Watson. Formant-frequency discrimination for isolated english vowels. *J. Acoust. Soc. Am.*, 95(1):485–496, 1994.
5. G. Peterson and H. Barney. Control methods used in a study of the vowels. *J. Acoust. Soc. Am.*, 24:175–184, 1952.
6. B. H. 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(1):327–335, 2008.
7. A. Van Hirtum, R. Blandin, and X. Pelorson. Validation of an analytical compressed elastic tube model for acoustic wave propagation. *J. Appl. Phys.*, 118(22):224905, 2015.
8. G. Yeung, S. M. Lulich, J. Guo, M. S. Sommers, and A. Alwan. Subglottal resonances of American English speaking children. *J. Acoust. Soc. Am.*, 144(6):3437–3449, 2018.