JASA ARTICLE

Squeezed waveguides as a framework to study vowel-like acoustic resonances

A. Eliraki, F. Vixege, X. Pelorson, and A. Van Hirtum[a)]
Université Grenoble Alpes, Centre National de la Recherche Scientifique, Grenoble Institut National Polytechnique, Laboratoire des Écoulements Géophysiques et Industriels, Grenoble, France

ABSTRACT:
The relationship between vocal tract geometry and acoustic resonances is often examined using speaker-specific data, limiting physical studies to a narrow set of static, rigid configurations and frequently producing discontinuous area functions that are poorly suited for studying dynamic vocal-tract behavior. To address these constraints, this study introduces a theoretical squeezed waveguide vocal tract (SWVT) framework for analyzing the relationship between acoustic resonances (up to 4–5 kHz) and smooth area functions that characterize vowel-like vocal tract configurations with one or two constrictions. The relevance of seven SWVT parameters—constriction positions, degrees, extents, and waveguide length—is established through theoretical analysis and experimental validation. Experiments are performed using both rigid, static waveguides and a deformable, molded waveguide. The molded waveguide facilitates experimental investigation of constriction degrees up to full closure and provides a basis for future studies on the (aero-)acoustic–geometry relationship in dynamic SWVT configurations.
 https://doi.org/10.1121/10.0042185

(Received 13 May 2025; revised 27 November 2025; accepted 14 December 2025; published online 13 January 2026)

[Editor: Zhaoyan Zhang] Pages: 272–284

I. INTRODUCTION

The study of acoustic resonances in the vocal tract (VT) during vowel production, referred to as formants, has been a central focus in spoken language research.[1–7] Driven by continuing advances in medical imaging and machine learning, a growing body of literature documents benchmark datasets of VT area functions and corresponding formants in order to describe how articulatory movements dynamically shape VT geometry and thereby influence vowel acoustics.[8–12] However, VT configurations are known to vary widely due to inter- and intra-speaker factors (e.g., age, gender, posture, body size[1,13–16]) and contextual influences (e.g., prosody, duration, environment[5,16–19]). To generalize beyond limited speaker-specific data,[20] efforts have been made to formalize the relationship between VT geometry and vowel production through the development of parameterized VT models, grounded in biomechanical, articulatory and data reduction principles.[15,21–30] Nevertheless the necessary geometric detail of the VT remains debated.[31] As a result, physical studies in vowel sound production mostly rely on a narrow set of static VT geometries, which are frequently characterized by discontinuous area functions.[6,12,31–33] This reliance not only restricts the generalizability of such studies, but also hinders their extension to dynamic VT configurations and aero-acoustic conditions.

To address these limitations, we introduce a novel parameterized framework for exploring vowel-like acoustics using deformable squeezed waveguides, i.e., cylindrical tubes compressed by up to two sets of parallel bars (pinchers). This framework builds on preliminary findings[34] showing that continuous area functions can be obtained analytically as a function of a limited set of physically meaningful parameters. Such a formulation would enable systematic parametric investigations, thereby in the long run facilitating physical aeroacoustic studies of VT dynamics, provided that acoustic vowel features can be accurately reproduced.

Therefore, in this work, the resonance space of squeezed waveguide VT-like (SWVT) geometries is compared to formant spaces[1,2,4,12,14,35] from American English and French vowels (Fig. 1 of the supplementary material). Resonances are computed using a simplified one-dimensional acoustic model, which has proven effective for predicting vowel formants up to approximately 4–5 kHz.[6,21,22,32,33,36–38] This frequency range covers the first few formants, which define the formant spaces critical for vowel perception and, consequently, for vowel classification.[32,36]

In the following, the SWVT framework is outlined in Sec. II. The theoretical resonance space of the seven-parameter SWVT framework is compared to vowel formant data in Sec. III. Selected SWVT area functions are considered in Sec. IV. Theoretical results of the SWVT framework are experimentally validated in Sec. VI for the vowels /i/, /a/ and /u/, using both rigid and deformable waveguides, in accordance with the experimental approach detailed in Sec. V. The influence of constriction features is further explored in Sec. VII. Conclusions are formulated in Sec. VIII.

II. INTRODUCING THE SWVT FRAMEWORK

The squeezed waveguide seven-parameter VT-like framework, or SWVT framework, is proposed to investigate

[a)]Email: annemie.vanhirtum@univ-grenoble-alpes.fr

https://doi.org/10.1121/10.0042185

the relationship between SWVT geometries (Sec. II A) and its acoustic resonances (Sec. II B).

A. SWVT framework: The waveguide

The squeezed VT-like waveguide, the SWVT geometry, is illustrated schematically in Fig. 1. A circular, uniform waveguide with variable length L_0 and wall thickness $d_0 = 3$ mm is aligned along the x axis, with its cross section lying in the yz-plane. The internal diameter is $D_0 = 25$ mm, yielding a cross-sectional area $A_0 = 490\,\mathrm{mm}^2$, consistent with values reported for unconstricted or open VT regions during adult vowel production.[6,12,32] The cross section is initially circular, with both minor (b_0 along the z-axis) and major (a_0 along the y axis) axes equal to $D_0/2$. Deformation is introduced via one or two sets of opposing pincher bars (indexed by $i \in 1,2$). Each pincher consists of two symmetrical, rounded bars (radius $r_p = 3.2$ mm) featuring a central flat segment of length $l_c \geq 0$ mm. The total pincher length along the x axis is $2r_p + l_c \geq 6.4$ mm, and the bars are aligned parallel to the y axis. Pinching induces a local constriction along the x axis, reducing the cross-sectional area $\hat{A}(x) \leq A_0$ over a length of up to $4 \times D_0$, symmetrically centered at x_c.[34,39]

Following Van Hirtum *et al.*,[39] the effect of a single pincher with $l_c = 0$ mm for a waveguide of length L_0 is described by two parameters: the constriction position x_c and the constriction degree $0 \leq \mathcal{P} < 100\%$, where $\mathcal{P} = 1 - \beta$ and $\beta = b(x_c)/b_0$. The deformed cross-sections are approximated as rounded stadium rings, with a reduced minor axis $\hat{b} < b_0$ and increased major axis $\hat{a} > a_0$. The waveguide geometry, including the area function $\hat{A}(x)$ and axes $\hat{a}(x)$ and $\hat{b}(x)$, can be expressed analytically as functions of the two-parameter set $(x_c, \mathcal{P})$. Eliraki[34] extended and validated this model using mechanical principles, demonstrating that it can be applied to configurations with two pinchers (each with $l_{c,1} = l_{c,2} = 0$ mm). In this case, modeled quantities such as $\hat{A}(x)$ are defined by a four-parameter set $(x_{c,1}, \mathcal{P}_1, x_{c,2}, \mathcal{P}_2)$, with pincher indices $i \in 1,2$, as shown in Fig. 1. Two pinchers can produce up to two constrictions, enabling for two- to five-tube configurations depending on constriction positions ($x_{c,1}$, $x_{c,2}$) and degrees ($\mathcal{P}_1$, $\mathcal{P}_2$). Unlike traditional concatenated tube representations,[6,21,22,32,36] the resulting area functions more closely resemble those of human speakers and exhibit smoothness, which facilitates subsequent studies with airflow, e.g., those involving fluid-structure interaction at the glottal source.

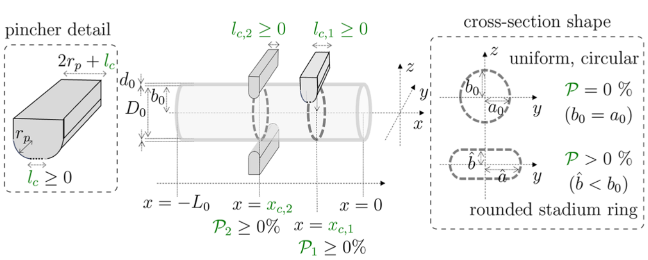


FIG. 1. The squeezed waveguide VT-like (SWVT) framework. A cylindrical tube of length L_0 and diameter D_0 is squeezed between one or two sets (indexed by $i \in 1,2$) of rounded parallel pincher bars ($r_p = 3.2$ mm), imposing constrictions defined by the parameter set $(l_c, x_c, \mathcal{P})_i$ (in green), where $l_{c,i}$ is the pincher extent, $x_{c,i}$ the constriction position, and $\mathcal{P}_i$ the constriction degree. The deformation results in stadium-shaped cross-sections with minor ($\hat{b} \leq b_0$) and major ($\hat{a} \geq a_0$) axes, indicated for both undeformed ($\mathcal{P} = 0\%$) and deformed ($0 < \mathcal{P} \leq 100$ in %) cases.

Considering the constriction length as a parameter is motivated, as during vowel production[12] (Fig. 1 of the supplementary material), the constricted region may extend from a few millimeters (e.g., /u/) up to several dozen millimeters (e.g., /o/). This variability led to the explicit inclusion of constriction length as a parameter in traditional tube models.[6,21,22,32,36] For each pincher, the minimum area is confined to a waveguide segment of fixed length, determined by its width $2r_p + l_c \geq 6.4$ mm, or more precisely, by the flat segment length l_c. When $l_c = 0$ mm, the minimum area is restricted to a fixed short length of $2r_p = 6.4$ mm. The constricted segment length can be increased either explicitly, by setting $l_c > 0$ mm to create a uniform constriction of length l_c with a single pincher, or indirectly, by positioning two pinchers close enough that the region between $x_{c,1}$ and $x_{c,2}$ remains constricted, albeit non-uniformly. It is assumed that the analytic geometric waveguide model, validated for one and two pinchers with $l_c = 0$ mm,[34,39] also holds for pinchers with $l_c > 0$ mm, when the minimum area is extended over the length l_c. Thus, the modeled SWVT waveguide geometry—and derived functions such as $\hat{A}(x)$, $\hat{b}(x)$, and $\hat{a}(x)$—is defined by a six-parameter set $(l_{c,1}, x_{c,1}, \mathcal{P}_1, l_{c,2}, x_{c,2}, \mathcal{P}_2)$, where $i \in 1,2$ indexes the pinchers (see Fig. 1). Including the waveguide length L_0 yields a general seven-parameter framework: $(L_0, l_{c,1}, x_{c,1}, \mathcal{P}_1, l_{c,2}, x_{c,2}, \mathcal{P}_2)$, which represents up to two area reduction portions in a waveguide with a fixed maximal area $A_0 = 490\,\mathrm{mm}^2$, as $b_0 = D_0/2$ is held constant.

B. SWVT framework: Theoretical acoustic resonances

The acoustic pressure within the waveguide is computed using a theoretical one dimensional acoustic model based on spatial transmission line theory.[37–39] The waveguide geometry is represented by its one-dimensional area function $\hat{A}(x)$, discretized with a spatial step of 0.3 mm ($\Delta x = 0.3$ mm). This step size was selected to ensure the theoretical model's output is independent of the discretization and to assure an approximately smooth geometry variation (see Sec. II A). A flanged infinite baffle exit condition is implemented in order to approximate the presence of the face at the VT outlet ($x = 0$). The one-dimensional theory neglects higher order acoustic mode propagation within the waveguide. At room temperature (21 °C), the cut-on frequency of these modes for a uniform ($\mathcal{P} = 0\%$) waveguide portion is approximately 8.0 kHz and decreases with increasing $\mathcal{P}$ due to the increase in the major axis $a(x)$ along the constricted section (see Fig. 1).[39–41] For waveguides pertinent to the VT (length $L_0 \approx 180$ cm, area $A_0 \approx 490\,\mathrm{mm}^2$), the first three acoustic resonance frequencies typically occur below 4–5 kHz.[34,39] Therefore, the framework's resonance

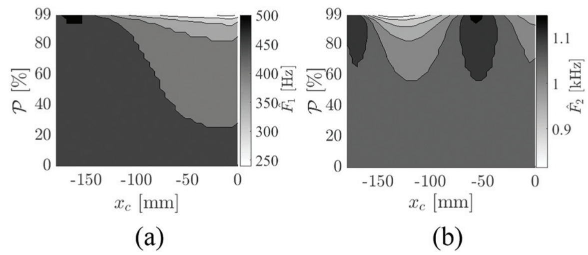


FIG. 2. Theoretical resonance frequencies $\hat{F}_{i\in\{1,2\}}$ for a single pincher as function of position x_c ($-L_0$ to 0) and constriction degree $\mathcal{P}$ (0%–99%) for $L_0 = 180\,\text{mm}$ and $l_c = 0\,\text{mm}$. (a) $\hat{F}_1(x_c, \mathcal{P})$, (b) $\hat{F}_2(x_c, \mathcal{P})$.

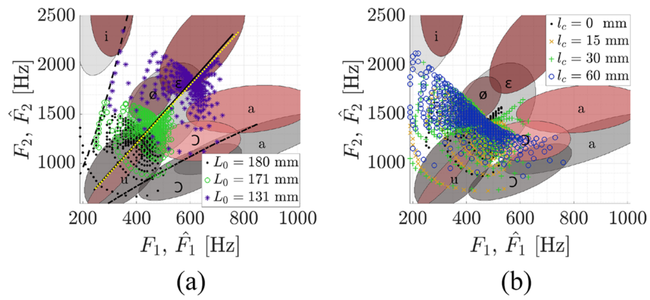


FIG. 3. Theoretical SWVT resonances ($\hat{F}_1$, $\hat{F}_2$) vs formants (F_1, F_2) for $-L_0 \leq x_c \leq 0\,\text{mm}$ and $0\% \leq \mathcal{P} \leq 99\%$: (a) SWVT subset ($L_0, l_c = 0\,\text{mm}, \text{x}_\text{c}, \mathcal{P}$) for $L_0 \in \{131, 171, 180\}$ mm. Relationships characterizing $(\hat{F}_1, \hat{F}_2)$ statistics for $L_0 \in [100\ 340]$ mm, $\hat{F}_{2stat}(\hat{F}_{1stat})$ with *stat* $\in \{\text{min}, \text{max}, \text{mean}\}$ are indicated (black lines). Reference curve for a uniform waveguide (dotted yellow line); (b) SWVT subset (L_0 = 180 mm, l_c, x_c, $\mathcal{P}$) for $l_c \in \{0, 15, 30, 60\}$ mm. (a) (L_0, $l_c = 0$ mm, x_c, $\mathcal{P}$). (b) (L_0 = 180 mm, l_c, x_c, $\mathcal{P}$).

space defined by the first two theoretical resonances $\hat{F}_i$ with $i \in \{1, 2\}$ can be investigated. Resonance frequencies $\hat{F}_i$ are estimated as peak frequencies in the theoretical transfer function $H(f)$ (in dB) between the amplitude of the acoustic volume flow rate U at the inlet (U_{-L_0} at $x = -L_0$) and outlet (U_0 at $x = 0$) as

$$\forall f, \quad H(f) = 20 \log_{10} \frac{|U_0(f)|}{|U_{-L_0}(f)|}. \tag{1}$$

Frequencies from 90 Hz to 6 kHz are considered in 10 Hz increments to cover the frequency range of interest.

For a waveguide with fixed length L_0 and a single pincher of length l_c, theoretical resonances can be expressed as functions of two SWVT parameters, $\hat{F}_i(x_c, \mathcal{P})$, assuming a single constriction (i.e., $\mathcal{P}_2 = 0\%$ so that pincher index $i = 1$ can be omitted for clarity). As an example to illustrate the influence of constriction parameters, theoretical resonance frequencies $\hat{F}_{i\in\{1,2\}}(x_c, \mathcal{P})$ for $L_0 = 180\,\text{mm}$ (reference VT length) and $l_c = 0$ mm are visualised in Fig. 2. Note that theoretical resonances cannot be visualized in this manner when more than two SWVT parameters are varied. In the following, constriction position(s) $x_{c,i}$ are varied from the inlet ($x = -L_0$) to the outlet ($x = 0$) in 5 mm steps, while constriction degree(s) $\mathcal{P}_i$ range from 0% (fully open) to 99% (near closure), using 5% increments for $0\% \leq \mathcal{P} \leq 80\%$ and 1% increments for $80\% < \mathcal{P} \leq 99\%$. The following sections compare the SWVT framework with literature data, focusing on vowel formants (Sec. III) or area functions $A(x)$ (Sec. IV).

III. SWVT vs LITERATURE: RESONANCE SPACES

The vowel space defined by the first two formants (F_1, F_2) of American English (gray ellipsoids) and French (red ellipsoids) vowels, spanning $200 \leq F_1 \leq 1000\,\text{Hz}$ and $500 \leq F_2 \leq 2500\,\text{Hz}$ (Fig. 1 of the supplementary material), is compared to the theoretical resonance space $(\hat{F}_1, \hat{F}_2)$ generated from various SWVT parameter subsets symbols.

A. One constriction: SWVT subset $(L_0, l_c, x_c, \mathcal{P})$

The theoretical resonance space $(\hat{F}_1, \hat{F}_2)$ for the SWVT subset $(L_0, l_c, x_c, \mathcal{P})$, considering one constriction (omitting pincher index $i = 1$ for clarity), is compared to the vowel formant space (F_1, F_2) in Fig. 3. The effects of constriction parameters $(x_c, \mathcal{P})$ are analyzed using reference values of $L_0 = 180\,\text{mm}$ and $l_c = 0\,\text{mm}$ (see Fig. 2), as well as for varying L_0 with $l_c = 0\,\text{mm}$ [Fig. 3(a)] and varying l_c with $L_0 = 180\,\text{mm}$ [Fig. 3(b)].

1. Reference: $(L_0 = 180\ \text{mm}, l_c = 0\ \text{mm}, x_c, \mathcal{P})$

Theoretical SWVT resonances $\hat{F}_1(x_c, \mathcal{P})$ and $\hat{F}_2(x_c, \mathcal{P})$ are plotted in Fig. 2. Theoretical $\hat{F}_1$ varies between 250 and 500 Hz (overall increase in 100%) and $\hat{F}_2$ between 800 and 1500 Hz (overall increase in 77%). Consequently, vowels in Fig. 3 with formants outside these ranges cannot be reproduced with the reference SWVT subset. Smallest and greatest resonance frequencies $\hat{F}_{i\in 1,2}$ in Fig. 2 are all observed in the parameter region associated with large constriction degrees $\mathcal{P}$.

In this high $\mathcal{P}$-range, $\hat{F}_1$ [Fig. 2(a)] decreases with constriction position x_c which is a tendency not accounted for in the thumb rule[32] (i.e., the empirical rule associated with the vowel trapezoid, stating that F_1 primarily increases with greater constriction degree, while F_2 increases with a forward shift of the constriction from the pharyngeal region toward the lips), where x_c is solely associated with F_2. Variation of $\hat{F}_1$ occurs for rather large $\mathcal{P} > 90\%$. 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 in $\hat{F}_1$ for decreasing constriction degree $\mathcal{P}$ (close to open following the thumb rule[32]) is retrieved for x_c away from the inlet up to the waveguide's outlet. However, for the x_c-range near the inlet, the tendency is reversed as $\hat{F}_1$ decreases for decreasing $\mathcal{P}$, which thus contradicts the thumb rule for formant F_1.

In the high $\mathcal{P}$-range, $\hat{F}_2$ [Fig. 2(b)] exhibits two successive maxima and minima for increasing x_c. Nevertheless, when not considering the extrema near the inlet and outlet, $\hat{F}_2$ increases with x_c (back to front). This trend is consistent with the rule of thumb[32] for formant F_2.

This increase or decrease in $\hat{F}_{1,2}$ 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 VT geometries are associated with the same $(\hat{F}_1, \hat{F}_2)$.

2. Influence of L_0, subset $(L_0, l_c = 0\ \text{mm}, x_c, \mathcal{P})$

Theoretical SWVT resonances $(\hat{F}_1, \hat{F}_2)$ for waveguide lengths $L_0 \in \{131, 171, 180\}$ mm and $l_c = 0$ mm are plotted in Fig. 3(a). These values correspond to typical VT lengths ranging from those of a child to an adult producing different vowels,[32] e.g., for a single male adult speaker[12] from 161 mm for /ae/ to 196 mm for /u/. Assuming the reference VT length of $L_0 = 180$ mm, the theoretical resonance space is confined to the range $250 \leq F_1 \leq 500$ Hz and $0.8 \leq F_2 \leq 1.5$ kHz (see also Fig. 2). This resonance space, defined by $(\hat{F}_1, \hat{F}_2)$, forms a tilted, butterfly-shaped isosceles trapezium. As L_0 increases, the resonance space shifts toward lower frequencies and narrows, leading to a decrease in both $\hat{F}_1$ and $\hat{F}_2$, as well as in their difference $\hat{F}_2 - \hat{F}_1$.

These tendencies are also observed in the behavior of descriptive statistics (in Hz) as a function of waveguide length L_0 (in mm); specifically, the minima $\min(\hat{F}_{1,2})(L_0)$, maxima $\max(\hat{F}_{1,2})(L_0)$, and mean $\bar{\hat{F}}_{1,2}(L_0)$. Statistics are computed across 12 different waveguide lengths L_0 (in mm) within the range $100 \leq L_0 \leq 340$ mm, covering typical human VT lengths. The resulting trends are accurately approximated (coefficient of determination $R^2 \geq 0.99$) by continuous quadratic functions, as shown in Fig. 4. Notably, these curves exhibit the greatest variation for L_0 values within the range typical of human articulation (161–196 mm). This indicates that SWVT resonance characteristics are most sensitive to small changes in L_0 within this range, thereby promoting greater resonance variation in the SWVT space $(\hat{F}_1, \hat{F}_2)$.

The relationships between the statistical quantities $\hat{F}_{2stat}$ and $\hat{F}_{1stat}$ [Fig. 3(a)], where $stat \in \{\min, \max, \text{mean}\}$, for $L_0 \in [100\ 340]$ mm, is accurately approximated as

$$\begin{cases} \hat{F}_{2\min}(\hat{F}_{1\min}) = -0.46 \cdot \hat{F}_{1\min}^2 + 1.98 \cdot \hat{F}_{1\min} + 48, \\ \hat{F}_{2\max}(\hat{F}_{1\max}) = 0.016 \cdot \hat{F}_{1\max}^2 + 0.28 \cdot \hat{F}_{1\max} + 260, \\ \hat{F}_{2\text{mean}}(\hat{F}_{1\text{mean}}) = 0.096 \cdot \hat{F}_{1\text{mean}}^2 + 2.99 \cdot \hat{F}_{1\text{mean}} + 0.3. \end{cases} \tag{2}$$

The theoretical resonance space $(\hat{F}_1, \hat{F}_2)$ is bounded by the curves $\hat{F}_{2stat}(\hat{F}_{1stat})$ representing the extrema ($stat \in \{\min, \max\}$), while $\hat{F}_{2\text{mean}}(\hat{F}_{1\text{mean}})$ illustrates the average trend. Overall, this confinement indicates that the theoretical resonance space associated with the SWVT subset $(L_0, l_c = 0\,\text{mm}, x_c, \mathcal{P})$ overlaps only with the central region of the vowel formant space $(\hat{F}_1, \hat{F}_2)$, and supports the observed trend that $\hat{F}_1$, $\hat{F}_2$ and their difference $\hat{F}_2 - \hat{F}_1$ all decrease as L_0 increases. Consequently, varying L_0 allows to shift along the direction defined by the mean statistical trend, consistent with Eq. (2), which shows that $\hat{F}_{2\text{mean}} \approx 3\hat{F}_{1\text{mean}}$ and that $\hat{F}_{1\text{mean}}$ decreases with increasing L_0. The alignment between the mean trend (black full line) and the reference curve [yellow dotted line in Fig. 3(a)], which represents the first two resonances of a uniform tube[40,41] of length L_0 and diameter D_0, supports the conclusion that the mean is governed solely by L_0.

3. Influence of l_c, subset $(L_0 = 180\ \text{mm}, l_c, x_c, \mathcal{P})$

Theoretical SWVT resonances $(\hat{F}_1, \hat{F}_2)$ for pincher length $l_c \in \{0, 15, 30, 60\}$ mm and $L_0 = 180$ mm are shown in Fig. 3(b). Increasing l_c expands the theoretical SWVT resonance space $(\hat{F}_1, \hat{F}_2)$, increasing the overlap with the vowel formant space (F_1, F_2) observed for $l_c = 0$ mm. The overall position of $(\hat{F}_1, \hat{F}_2)$ remains largely unchanged, confirming the statement in Sec. III A 2) that the position is primarily determined by the fixed length $L_0 = 180$ mm. The expansion is most pronounced in the direction perpendicular to the statistical mean trend $\hat{F}_{2\text{mean}}(\hat{F}_{1\text{mean}})$ [see Fig. 3(a)], such that the theoretical resonance space remains its tilted, butterfly-shaped isosceles trapezium form, with its width increasing as l_c increases. Thus, overall, l_c influences the extrema statistics $\hat{F}_{2stat}(\hat{F}_{1stat})$, where $stat \in \{\min, \max\}$. Compared to $l_c = 0$ mm, this can result in higher values of either $\hat{F}_1$ or $\hat{F}_2$, but not both simultaneously. As a result, the overlap with the vowel formant space (F_1, F_2) remains confined to a subregion of that space.

B. Two constrictions: SWVT subset $(L_0 = 180\ \text{mm}, l_{c,1}, x_{c,1}, \mathcal{P}_1, l_{c,2}, x_{c,2}, \mathcal{P}_2)$

This section examines how the theoretical resonance space observed for a single constriction (Sec. III A) expands in the case of two constrictions (pincher index $i = \{1, 2\}$ in Fig. 1). The theoretical space $(\hat{F}_1, \hat{F}_2)$ for the SWVT subset

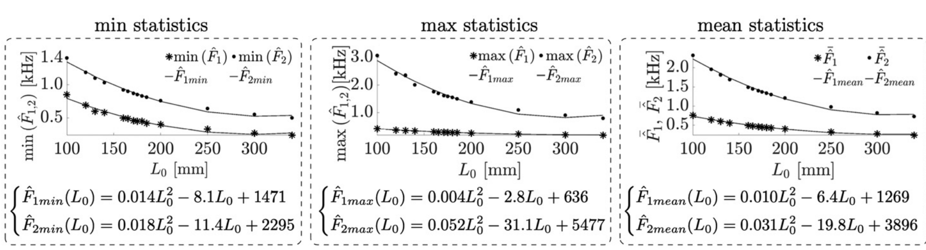


FIG. 4. Descriptive statistics of theoretical SWVT resonance space $(\hat{F}_1, \hat{F}_2)$ for subset $(L_0,\ l_c = 0\ \text{mm},\ x_c,\ \mathcal{P})$ as a function of waveguide length L_0 (in mm) with $100 \leq L_0 \leq 340$ mm. Fitted quadratic approximations of $\hat{F}_{1stat}$ and $\hat{F}_{2stat}$ (in Hz) with $stat \in \{\min, \max, \text{mean}\}$.

https://doi.org/10.1121/10.0042185

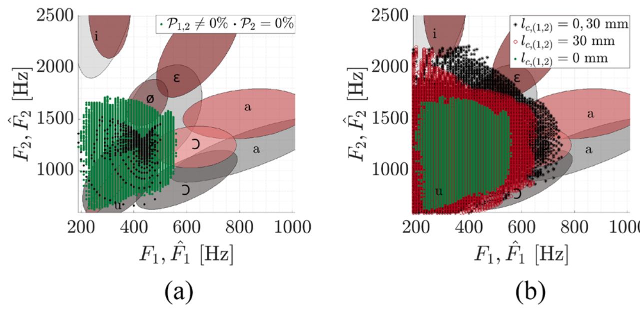


FIG. 5. Theoretical SWVT resonances $(\hat{F}_1, \hat{F}_2)$ vs formants (F_1, F_2) for $-L_0 \leq x_{c,1}, x_{c,2} \leq 0$ mm and $0\% \leq \mathcal{P}_1, \mathcal{P}_2 \leq 99\%$ ($L_0 = 180$ mm): (a) SWVT subsets for $l_{c,1} = l_{c,2} = 0$ mm, one constriction ($\mathcal{P}_2 = 0\%$) vs two constrictions ($\mathcal{P}_{1,2} \neq 0\%$); (b) SWVT subset for $l_{c,1}, l_{c,2} \in \{0, 30\}$ mm for two constrictions. (a) $l_{c,1} = l_{c,2} = 0$ mm. (b) $l_{c,1}, l_{c,2} \in \{0, 30\}$ mm.

$(L_0 = 180\,\text{mm}, l_{c,1}, x_{c,1}, \mathcal{P}_1, l_{c,2}, x_{c,2}, \mathcal{P}_2)$ is compared to the vowel formant spaces (F_1, F_2).

Figure 5(a) shows the theoretical SWVT resonance space $(\hat{F}_1, \hat{F}_2)$ for $l_{c,1} = l_{c,2} = 0$ mm ($\mathcal{P}_{1,2} \neq 0\%$), alongside the resonance space for a single constriction ($\mathcal{P}_2 = 0\%$). The overall position of the resonance space, whether for one or two constrictions, remains unchanged. This confirms the previous section's (Sec. III A) finding that the position is primarily determined by the fixed length $L_0 = 180$ mm. Nevertheless, the size of the resonance space is expanded in all directions, resulting in a significant improvement in its overlap with the vowel formant space (F_1, F_2), although it remains localized based on L_0.

Figure 5(b) illustrates the effect of increasing the pincher extent $l_{c,1}$, $l_{c,2}$, or both from 0 to 30 mm. This leads to a further expansion of the resonance space in all directions, improving its overlap with the vowel formant space (F_1, F_2), even though the waveguide length L_0 remains fixed at 180 mm. Notably, the theoretical SWVT resonance space $(\hat{F}_1, \hat{F}_2)$ either reaches the boundary of the French vowel resonance space or overlaps at least partially with American English vowels in the (F_1, F_2) space. This is clearly observed for vowels such as /i/, /a/, and /ε/ in Fig. 5(b).

IV. SWVT vs LITERATURE: AREA FUNCTIONS

As shown in Sec. III, the SWVT framework can approximate the formants (F_1, F_2) of American English and French vowels using $L_0 = 180$ mm. The analysis below identifies the subsets of SWVT parameters $(L_0 = 180\,\text{mm}, l_{c,1}, x_{c,1}, \mathcal{P}_1, l_{c,2}, x_{c,2}, \mathcal{P}_2)$ that achieve this approximation, along with the corresponding SWVT area functions $\hat{A}(x)$.

The most direct method for selecting SWVT parameter subsets and corresponding area functions $\hat{A}(x)$ for a given vowel is to ensure that the resulting theoretical resonances $(\hat{F}_1, \hat{F}_2)$ lie within the target vowel's space (F_1, F_2). For French vowels,[2] this method is referred to as the **Calliope criterion**, and for American English vowels,[1] the **Peterson criterion**. The ellipsoids representing vowel spaces exhibit substantial variation due to intra- and inter-speaker differences, as noted in the introduction (Sec. I). As a result, SWVT parameter selection is not unique. To address this, for American English vowels, an alternative criterion is proposed based on area functions $A(x)$ and formant frequencies derived from a single adult male speaker, as reported by Story.[12] This approach is referred to as $\mathbf{MRI_0}$ (for area functions) and **Story** (for formants), and is detailed in Sec. IV A. For French vowels, however, only the Calliope criterion is used.

A. SWVT-tailored area functions and resonance invariants

Compared to the Peterson criterion, which uses a range of vowel formant frequencies, the **Story criterion**, provides fixed, well-defined formant values.[12] However, these values were obtained under specific, and possible variable, conditions (MRI environment, temperature). On the other hand, the corresponding VT area functions $A(x)$ ($\mathbf{MRI_0}$) vary in length L from 161 to 196 mm. In this section, these MRI-based VT area functions $A(x)$ ($\mathrm{MRI_0}$) are adapted to fit within the SWVT domain. This is done by reshaping them to satisfy the condition $\hat{A}(x) \leq A_0$ over the interval $-L_0 \leq x \leq 0$, where the reference length is $L_0 = 180$ mm and the maximum cross section area is $A_0 = 490\,\text{mm}^2$ (see Fig. 1). This approach allows for the derivation of a reference set of theoretical resonance invariants, offering an additional criterion for selecting SWVT parameters.

1. SWVT tailored area functions: MRI_I, MRI_{II}

The first type of SWVT-tailored area functions, denoted $\mathbf{MRI_I}$ (supplementary material Fig. 2 for /i/, /a/, /u/, and /ε/), is obtained by rescaling the longitudinal dimension with a factor $\alpha_x = L_0/L$, ensuring that the adjusted length equals the reference length L_0; either stretching ($\alpha_x > 1$) or compressing ($\alpha_x > 1$) the x-dimension. Additionally, any cross-sectional area exceeding A_0 is capped at A_0. This thresholding preserves the original cross section where $A(x) \leq A_0$, while reducing values that exceed A_0.

The second type, denoted $\mathbf{MRI_{II}}$ (supplementary material Fig. 2 for /i/, /a/, /u/, and /ε/), also rescales the x-dimension by α_x, but further adjusts the cross-sectional areas by a factor $\alpha_A = A_0/\max(A(x))$. As a result, the $\mathrm{MRI_{II}}$ functions conform to the full SWVT domain, with both the length equal to L_0 and the maximum cross-sectional area equal to A_0. Unlike $\mathrm{MRI_I}$, this method modifies the cross section along the entire x-range; either enlarging it ($\alpha_A > 1$) or reducing it ($\alpha_A < 1$), depending on the original area function $\mathrm{MRI_0}$.

2. Theoretical resonance invariants of MRI_I, MRI_{II}

Section III outlined that theoretical resonances $(\hat{F}_1, \hat{F}_2)$ depend on the degree of area reduction (impact $\mathcal{P}_{1,2}$), the length of the constricted portion (impacts $l_{c,1}$ and $l_{c,2}$), and the waveguide length (impact L_0). The SWVT-tailoring of $\mathrm{MRI_0}$, as described in Sec. IV A 1, involves both area scaling—which influences the degree of constriction—and longitudinal scaling—which affects both the waveguide length and the length of the constricted region(s). As a

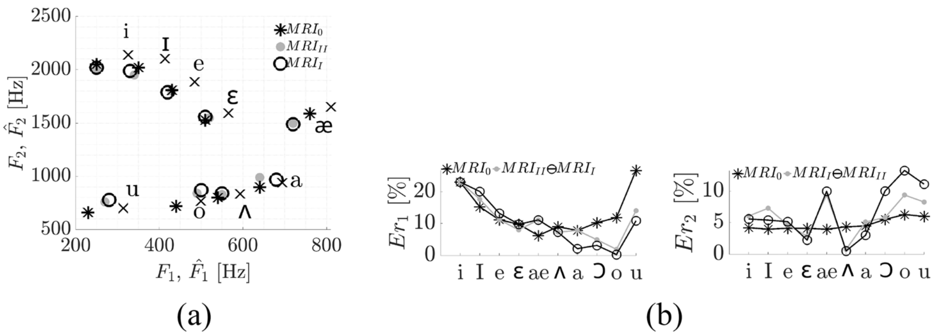


FIG. 6. For $\mathrm{MRI_0}$, $\mathrm{MRI_I}$, and $\mathrm{MRI_{II}}$: (a) Theoretical resonance frequencies $(\hat{F}_1, \hat{F}_2)$ and formants (F_1, F_2) ($\times$, Story) (Ref. 12). (b) Relative accuracy Er_1 and Er_2 with respect to formants (F_1, F_2). (a) $(\hat{F}_1, \hat{F}_2)$, (b) Er_1, Er_2.

result, the theoretical resonances for $\mathrm{MRI_0}$, $\mathrm{MRI_I}$, and $\mathrm{MRI_{II}}$ may differ.

The theoretical resonance space $(\hat{F}_1, \hat{F}_2)$ for area functions $\mathrm{MRI_0}$, $\mathrm{MRI_I}$, and $\mathrm{MRI_{II}}$, along with reference formant values (F_1, F_2) (Story[12]) for 10 vowels, is shown in Fig. 6(a). While all four data points for each vowel lie within the same general region, noticeable differences are present. For MRI-based geometries, these discrepancies may be linked to the magnitude of the scaling factors (α_x, α_A), as seen for vowels like /a/ and /u/. As mentioned previously, discrepancies between theoretical resonances obtained for $\mathrm{MRI_0}$ and formants can be partly attributed to different measurement conditions.

Observed differences (Fig. 6(a)) between theoretical resonances $\hat{F}_i$ for $\mathrm{MRI_0}$, $\mathrm{MRI_I}$, and $\mathrm{MRI_{II}}$ and formants F_i for each vowel are quantified considering the mean relative absolute difference (Er, relative difference in short) with respect to N formants F_i with index i in the set $\{i_1, \ldots, i_n\}$ obtained as

$$Er_{i_1 \ldots i_n} = \frac{1}{N} \sum_{i \in \{i_1, \ldots, i_n\}} \frac{|\hat{F}_i - F_i|}{F_i}. \tag{3}$$

For each vowel, relative differences Er_1 and Er_2 are shown in Fig. 6(b). The difference Er_2 remains below 15%, while Er_1 can reach up to 25%. Overall, discrepancies between vowel formants and theoretical resonance frequencies are all of the same order of magnitude, including those from the SWVT-tailored area functions $\mathrm{MRI_I}$ and $\mathrm{MRI_{II}}$. Therefore, these theoretical resonances $\hat{F}_i$, defined under well-specified conditions, serve as **resonance invariants** and can be effectively used for SWVT parameter selection for a given American English vowel. In Sec. IV A 3, **$\mathrm{MRI_{II}}$ resonance invariants** are used to select SWVT parameter for a single pincher imposing one constriction.

3. SWVT parameter selection with respect to resonance invariants from $\mathrm{MRI_{II}}$

Relative differences Er_{12} between the theoretical resonances of the SWVT reference subset ($L_0 = 180\,\mathrm{mm}$, $l_c = 0\,\mathrm{mm}$, $\mathrm{x_c}$, $\mathcal{P}$) and the vowel formants (F_1, F_2) (i.e., the Story criterion[12]) for /a/ are shown in Fig. 7 (and in supplementary material Fig. 3 for /i/, /a/, and /u/). The SWVT parameter set minimizing Er_{12} is highlighted in the zoomed-in range $15\% \leq Er_{12} \leq 30\%$ and the selected SWVT parameters $(x_c, \mathcal{P})$ are indicated. In general, selected configurations exhibit severe constrictions degrees $\mathcal{P} \geq 98\%$ and constriction positions shift from front to back in agreement with vowel area functions $A(x)$ associated with $\mathrm{MRI_0}$.

Using vowel formants (F_1, F_2) (i.e., the **Story** criterion) to compute the relative differences is considered a poorly defined selection criterion due to inter- and intra-speaker variability. Therefore, relative differences Er_{12} with respect to theoretical resonance frequencies $\hat{F}_i$ from $\mathrm{MRI_0}$ (the **$\mathrm{MRI_0}$** criterion) and resonance invariants from $\mathrm{MRI_{II}}$ (the **$\mathrm{MRI_{II}}$** criterion) (see Sec. IV A 2) are assessed as well. The comparison of minimum relative differences obtained for these three selection criteria on the SWVT subset ($L_0 = 180\,\mathrm{mm}, \mathrm{l_c}, \mathrm{x_c}, \mathcal{P}$) for constriction lengths $l_c \leq 60\,\mathrm{mm}$ is plotted for /a/ in Fig. 7(b) (and in supplementary material Fig. 4 for /i/, /a/, and /u/). All three selection criteria exhibit consistent trends with respect to the choice of l_c, confirming that the resonance invariants from $\mathrm{MRI_{II}}$ are suitable for SWVT parameter selection. The overall minimum of the relative differences $Er_{12}(l_c)$ occurs at a shorter $l_c \approx 20\,\mathrm{mm}$ for /u/, compared to $l_c \geq 50\,\mathrm{mm}$ for /i/ and /a/. This is consistent with reported area functions $\mathrm{MRI_0}$.[12]

B. SWVT area function selection for vowels /i/, /a/, and /u/

SWVT area functions $\hat{A}(x)$ are selected by minimizing the relative differences Er_{12} between theoretical SWVT

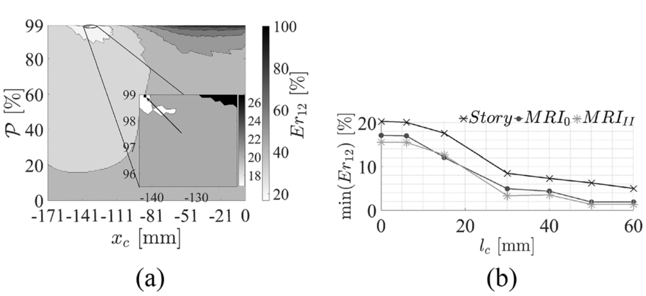


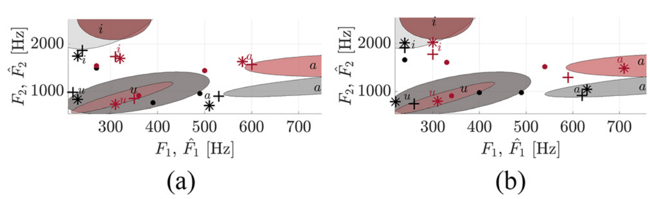
FIG. 7. For /a/: (a) Relative differences Er_{12} for theoretical resonances of SWVT parameter subset ($L_0 = 180$ mm, $l_c = 0$ mm, x_c, $\mathcal{P}$) with respect to formants (F_1, F_2) (Story) (Ref. 12) with $\min(Er_{12}) = 17\%$ for $x_c = -142$ mm and $\mathcal{P} = 99\%$ (arrow in zoom); (b) minima of relative differences Er_{12} as a function of l_c with respect to three criteria. (a) Er_{12}, (b) $\min(Er_{12})$.

resonance frequencies $(\hat{F}_1, \hat{F}_2)$ from the SWVT parameter subset and reference frequencies for each vowel. The $\mathbf{MRI_{II}}$ criterion is applied for American English vowels, so that reference vowel frequencies are the resonance invariants from $\mathrm{MRI_{II}}$ (Sec. IV A 2). Results obtained from selected SWVT parameters for **American English vowels** are plotted in **black**, in accordance with the gray-shaded ellipsoids in the vowel space. For French vowels, the **Calliope criterion** is applied. Results obtained from selected SWVT parameters for **French vowels** are plotted in **red**, in accordance with the red-shaded ellipsoids in the vowel space (F_1, F_2). To lighten notations, the used criterion and formant space coloring is omitted from figure captions and legends in this section.

Selected configurations for subsets of the SWVT parameter space corresponding to configurations with one (Sec. IV B 1) or two constrictions (Sec. IV B 2) are considered. Results are presented for the American English and French vowels /i/ (front constriction in $\mathrm{MRI_0}$), /a/ (back constriction in $\mathrm{MRI_0}$) and /u/ (center and front constriction in $\mathrm{MRI_0}$). Following Sec. IV, $L_0 = 180\,\mathrm{mm}$ is held constant and the length is not further considered.

1. One constriction: SWVT selections in $(L_0 = 180\ \mathrm{mm}, l_c, x_c, \mathcal{P})$

Theoretical resonances $(\hat{F}_1, \hat{F}_2)$ associated with selected SWVT area functions $\hat{A}(x)$ (Fig. 5 of the supplementary material) for the parameter subset $(L_0 = 180\,\mathrm{mm}, \mathrm{l_c} = 0\,\mathrm{mm}, \mathrm{x_c}, \mathcal{P})$ are shown (•) in Fig. 8(a). Overall, the tendencies for American English and French vowels are similar. For the vowel /a/, the constriction is located near the back (the glottis) at $-165 < x_c < -144\,\mathrm{mm}$, with $\mathcal{P} = 96\%$. For /i/, the constriction is near the front (the lips) at $-57 < x_c < -51\,\mathrm{mm}$, with $96 < \mathcal{P} < 99\%$. For /u/, the constriction lies near the center of the waveguide at $-124 < x_c < -113\,\mathrm{mm}$, with $\mathcal{P} = 99\%$. These observations are consistent with literature. However, selected SWVT area functions $\hat{A}(x)$ exhibit constrictions slightly upstream (i.e., further back) compared to those observed in VT area functions $A(x)$ ($\mathrm{MRI_0}$). As in Sec. III A, the theoretical resonances $(\hat{F}_1, \hat{F}_2)$ do not reach the target vowel spaces for /i/ and /a/. However, for /u/, the resonances fall within the expected vowel space for both American English and French, supporting the relevance of the proposed selection criteria.



(a) (b)

FIG. 8. Theoretical resonances $(\hat{F}_1, \hat{F}_2)$ of selected SWVT area functions $\hat{A}(x)$ $(L_0 = 180\,\mathrm{mm})$ for American English (black) and French (red) vowels /i/, /a/, and /u/ from SWVT subsets: (a) $(l_c, x_c, \mathcal{P})$ with $l_c = 0\,\mathrm{mm}$ (•), $l_c = 30\,\mathrm{mm}$ (+), and $l_c = 60\,\mathrm{mm}$ (∗); (b) $(l_{c,1}, x_{c,1}, \mathcal{P}_1, l_{c,2}, x_{c,2}, \mathcal{P}_2)$ with $l_{c,1} = l_{c,2} = 0\,\mathrm{mm}$ (•), $l_{c,1} = l_{c,2} = 30\,\mathrm{mm}$ (*), and $l_{c,1}, l_{c,2} \in \{0, 30\}\,\mathrm{mm}$ with $l_{c,1} \neq l_{c,2}$ (+). (a) $(l_c, x_c, \mathcal{P})$, (b) $(l_{c,1}, x_{c,1}, \mathcal{P}_1, l_{c,2}, x_{c,2}, \mathcal{P}_2)$.

Theoretical resonances $(\hat{F}_1, \hat{F}_2)$ associated with selected SWVT area functions $\hat{A}(x)$ (Fig. 6 of the supplementary material) for the parameter subset $(L_0 = 180\,\mathrm{mm}, \mathrm{l_c} \in \{30, 60\}\,\mathrm{mm}, x_c, \mathcal{P})$ are shown (+, ∗) in Fig. 8(a). Thus, the influence of $l_c > 0$ mm is assessed. Similar to the case of $l_c = 0\,\mathrm{mm}$, the same general tendencies are observed for both American English and French vowels. For the vowel /a/, the constriction is located near the back (the glottis) between $155 < x_c < 130\,\mathrm{mm}$ with $96 < \mathcal{P} < 99\%$. For /i/, the constriction is positioned near the front (the lips) at $52 < x_c < 58\,\mathrm{mm}$ with $93 < \mathcal{P} < 96\%$. For /u/, the constriction lies toward the center of the waveguide, at $140 < x_c < 125\,\mathrm{mm}$ with $\mathcal{P} = 99\%$. As with $l_c = 0\,\mathrm{mm}$, all SWVT area functions $\hat{A}$ show constrictions located slightly upstream (i.e., toward the back) relative to those observed in the VT area functions $A(x)$ ($\mathrm{MRI_0}$). As found in Sec. III A, increasing the constriction extent ($l_c \in \{30, 60\}\,\mathrm{mm}$), enables theoretical resonances $(\hat{F}_1, \hat{F}_2)$ to reach the target vowel spaces for /u/ as well as for /i/ and /a/. The resonances obtained for $l_c = 30\,\mathrm{mm}$ or $l_c = 60\,\mathrm{mm}$ are closely adjacent.

2. Two constrictions: SWVT selections in $(L_0 = 180\ \mathrm{mm}, l_{c,1}, x_{c,1}, \mathcal{P}_1, l_{c,2}, x_{c,2}, \mathcal{P}_2)$

Theoretical resonances $(\hat{F}_1, \hat{F}_2)$ associated with selected SWVT area functions $\hat{A}(x)$ (Fig. 7 of the supplementary material) for the parameter subset $(L_0 = 180\,\mathrm{mm}, \mathrm{l_{c,1}} = 0\,\mathrm{mm}, \mathrm{x_{c,1}}, \mathcal{P}_1, \mathrm{l_{c,2}} = 0\,\mathrm{mm}, \mathrm{x_{c,2}}, \mathcal{P}_2)$ are shown (•) in Fig. 8(b). As in Sec. III B, the theoretical resonances $(\hat{F}_1, \hat{F}_2)$ do not reach the target vowel spaces for /i/, /a/, though they do fall within the correct region for /u/ in both American English and French. The influence of adding a second constriction varies by vowel:

- For /a/, where the VT area function $A(x)$ ($\mathrm{MRI_0}$) typically shows a single constricted portion, the SWVT area function $\hat{A}(x)$ forms two closely spaced constrictions, effectively increasing the extent of the constricted region (though not uniformly).
- For /u/ and /i/, which show up to two constrictions in the VT area function $A(x)$ (front and back for /i/; center and front for /u/), SWVT area functions $\hat{A}(x)$ either display two closely spaced constrictions—broadening the constricted region—or two more widely spaced constrictions, which replicate the presence of a second constriction.

As with the single-constriction SWVT configurations (see Sec. IV B 1), all SWVT area functions $\hat{A}(x)$ show constrictions located slightly upstream (i.e., further back) compared to those observed in the VT area functions $A(x)$ ($\mathrm{MRI_0}$). Importantly, adding a second constriction does not significantly improve the accuracy of the results compared to single-constriction models, particularly when $l_c > 0\,\mathrm{mm}$.

Theoretical resonances $(\hat{F}_1, \hat{F}_2)$ associated with selected SWVT area functions $\hat{A}(x)$ (Fig. 8 of the supplementary material) for the parameter subset $(L_0 = 180\,\mathrm{mm}, \mathrm{l_{c,1}}, \mathrm{x_{c,1}}, \mathcal{P}_1, \mathrm{l_{c,2}}, \mathrm{x_{c,2}}, \mathcal{P}_2)$ for either $l_{c,1} = l_{c,2} = 30\,\mathrm{mm}$ or $l_{c,1}, l_{c,2} \in \{0, 30\}\,\mathrm{mm}$ with $l_{c,1} \neq l_{c,2}$ are shown (∗, +) in Fig. 8(b). Increasing the constriction length allows the theoretical resonances $(\hat{F}_1, \hat{F}_2)$ to reach the target vowel spaces, which confirms findings in Sec. III B. As for $l_{c,1} = l_{c,2} = 0\,\text{mm}$, the second constriction is either adjacent to the first one in order to extent the constricted region or widely spaced in order to reproduce a second constriction location.

V. EXPERIMENTAL METHODS

A. Experimental setup and data analysis

The setup in Fig. 9 is designed to measure acoustic pressure at various positions x_p ($x_p \in \{-5, -37, -60\}$ mm) along the centerline of the waveguide under test (rigid or deformable).[34] The waveguide is positioned within an insulation chamber (Décibel France, working dimensions $118 \times 65 \times 95\,\text{cm}$, volume $0.73\,\text{m}^3$). A compression chamber (Monacor KU-916T) is connected to its inlet at $x = -L_0$ via an adapter featuring a centrally positioned communication hole (2 mm in diameter and 6 mm in length) through which a 30-s linear frequency sweep (frequency range from 0.1 up to 6 kHz) is repeatedly emitted into the waveguide. A rigid Plexiglas rectangular baffle (365×352 mm in the yz-plane, with a thickness of 5 mm) is attached at the waveguide exit located at $x = 0$. This baffle approximates the infinite baffle assumption used in the theoretical acoustic model (see Sec. II B) and, unless stated differently, imposes a circular exit condition to the waveguide.

Data generation (sweep signal V_s in V) and acquisition of acoustic pressure (V_p in V), and wall displacement ($\Delta a \geq 0$ in mm) occurs as outlined in Eliraki,[34] so that the acoustic pressure amplitude ($|P_{x_p}(f)|$) as a function of frequency f at each probe position x_p can be obtained.[34,39] Measured acoustic waveguide resonances, denoted $\bar{f}_i$, are identified at frequencies corresponding to peaks in the acoustic pressure amplitude (in dB). The temperature during experiments (21 to 28 °C) can differ from 21 °C accounted for in the theoretical acoustic model (Sec. II B). Subsequent differences between measured and theoretical resonance frequencies can reach up to 2%.

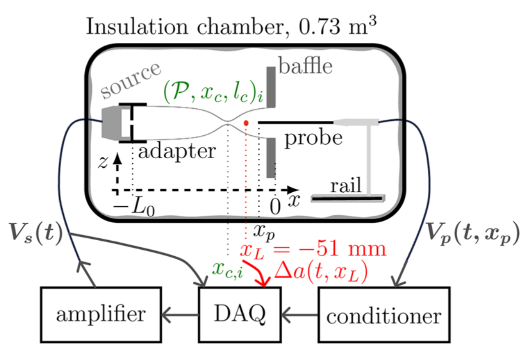


FIG. 9. Setup (xz-plane) for measuring the acoustic pressure [$V_p(t, x_p)$, V] along the centerline of a waveguide [$(\mathcal{P}, x_c, l_c)_i$, pincher index $i \in \{1, 2\}$] at probe position x_p, following a source excitation signal $V_s(t)$ (V) transmitted via an adapter. The wall displacement along the major axis Δa (mm) is measured using a laser transceiver at position x_L.

B. Experimental waveguides

1. Rigid (R) and molded (M) waveguides

Rigid (R) waveguides are 3D fabricated using 3D printing (RAISE3D Pro3, accuracy ±0.01 mm), ensuring precise reproduction of the intended waveguide geometry. This includes rigid waveguides based on VT area functions $A(x)$ (MRI$_0$), SWVT-tailored VT area function geometries (MRI$_\text{I}$, MRI$_\text{II}$, see Sec. IV A 1) and selected SWVT area functions $\hat{A}(x)$ for the American English vowels /i/, /a/, /u/ (see Sec. IV B).

Deformable molded (M) uniform cylindrical waveguides ($L_0 = 180\,\text{mm}$, $D_0 = 25\,\text{mm}$, $d_0 = 3\,\text{mm}$), matching the SWVT dimensions shown in Fig. 1, are molded using a custom-designed mold and elastomer silicone mixture [Dragonskin (Smooth-On Inc.)].[34] SWVT parameters can be applied to these deformable waveguides using a custom-built, controllable squeezing apparatus,[34,42] which is installed within the insulation chamber (Fig. 9). The pinchers, designed to impose $l_c \in \{0, 30, 60\}$ mm, are 3D printed (RAISE3D Pro3, accuracy ±0.01 mm) to match the required SWVT parameters. Selected SWVT area functions $\hat{A}(x)$ for the American English vowels /i/, /a/, /u/ (see Sec. IV B) can be replicated using the deformable molded waveguide.

In contrast to the theoretical exploration of the SWVT parameter space presented in Secs. III and IV, the experimental setup imposes physical constraints near the inlet (at $x = -L_0$ mm) and outlet (at $x = 0$ mm) due to the connection of the waveguide with the source (at $x = -L_0$ mm) and with the baffle (at $x = 0$ mm). For a waveguide with $L_0 = 180$ mm, this reduced the theoretical x_c-range from $-180 \leq x_c \leq 0$ mm to the experimental x_c-range:

- $-155 + l_c/2 \leq x_c \leq -15 - l_c/2$ mm for molded (M) deformable waveguides,
- $-180 \leq x_c \leq -15 - l_c/2$ mm for rigid (R) waveguides.

Constriction degrees up to full closure can be applied for deformable (M) waveguides. However, for rigid (R) waveguides, fabrication is not feasible for constriction degrees exceeding $\mathcal{P} > 96\%$. This constrains the range of experimentally achievable SWVT configurations and thus restricts the experimentally assessable resonance space $(\hat{F}_1, \hat{F}_2)$ (Fig. 9 of the supplementary material). In the case that selected SWVT parameter subsets for a given American English vowel, based on the MRI$_\text{II}$-criterion (see Sec. IV B), are experimentally not feasible, SWVT parameters are re-selected using the formant space-based **Peterson**-criterion (Table I of the supplementary material).

2. Major axis validation for molded waveguides

The analytical model used to represent the SWVT waveguide geometry (see Sec. II A) was previously validated for deformable molded (M) waveguides with one or two pinchers, using $l_{c,1} = l_{c,2} = 0\,\text{mm}$.[34,39] In this section, the validation is extended to cases where $l_c \geq 0$ mm, by comparing SWVT theoretical values and experimental measurements of wall displacement Δa as a function of

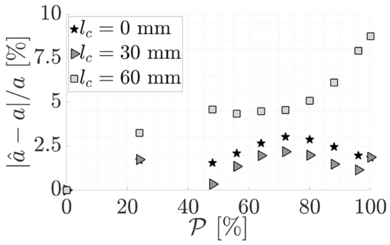


FIG. 10. Mean relative error (in %) as a function of $\mathcal{P}$ (in %) between the measured major axis $a \approx \Delta a + a_0$ and theoretical major axis $\hat{a}$ for $l_c \in \{0, 30, 60\}$ mm.

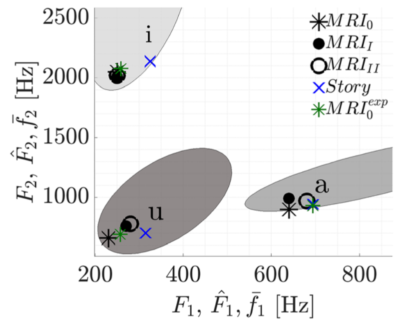


FIG. 11. Experimental resonance frequencies $(\bar{f}_1, \bar{f}_2)$ of rigid 3D printed MRI_0 waveguides ($\mathrm{MRI}_0^{\mathrm{exp}}$); theoretical resonances $(\hat{F}_1, \hat{F}_2)$ for MRI_0, MRI_I, and $\mathrm{MRI}_{\mathrm{II}}$; formant spaces (F_1, F_2) for the American English vowels /i/, /a/, /u/, along with reference formants (×, Story).

constriction degree $\mathcal{P}$ and normalized position $(x_c - x_L - l_c/2)/a_0$ (Fig. 10 of the supplementary material). The mean relative error between the theoretical major axis $\hat{a}$ and the measured values $a \approx a_0 + \Delta a$ (see Fig. 1) as a function of $\mathcal{P}$ for $l_c \in \{0, 30, 60\}$ mm is shown in Fig. 10. For $l_c = \{0, 30\}$ mm, the experimental and theoretical results show good agreement across all values of $\mathcal{P}$, with relative difference remaining below 5% and following similar trends. For $l_c = 60$ mm, good agreement is observed for $\mathcal{P} \leq 70\%$, with relative differences around 5%. However, for $\mathcal{P} > 70\%$, the discrepancy increases, likely due to physical constraints discussed in Sec. V B 1. Nevertheless, since all mean differences remain below 10%, the theoretical geometrical waveguide model is considered to accurately represent the deformation of molded (M) waveguides using the experimental setup described in Sec. V for $l_c \geq 0$ mm.

VI. SWVT VALIDATION: /I/, /A/, AND /U/

Experimental resonances $\bar{f}_i$ of rigid (R) and molded (M) waveguides detailed in Sec. V B 1 are compared to theoretical resonance frequencies $\hat{F}_i$ in order to validate the SWVT framework. As VT area functions $A(x)$ (MRI_0) are available for American English only, the experimental validation focuses on the American English vowels /i/, /a/, and /u/.

A. Rigid MRI_0 waveguides

Experimental resonance frequencies $(\bar{f}_1, \bar{f}_2)$ of rigid 3D printed MRI_0 waveguides ($\mathrm{MRI}_0^{\mathrm{exp}}$) are plotted in Fig. 11. For all three vowels, experimental values validate theoretical values $\hat{F}_i$ for MRI_0 as discrepancies are small. This provides an additional validation of the theoretical acoustic model used in the SWVT framework for the shown frequency range shown (up to 2.5 kHz). In Sec. VI B, rigid (R) and molded (M) waveguides obtained with selected SWVT parameters are considered.

B. Rigid and deformable SWVT waveguides

As outlined in Sec. V B 1, rigid (R) and deformable molded (M) waveguides shaped with selected SWVT parameters are experimentally evaluated for each of the American English vowels /i/, /a/, and /u/. Multiple waveguides are assessed per vowel (Table I of the supplementary material), depending on the specific SWVT parameter subset used for selection (Sec. IV B) and the experimentally accessible parameter range (see Sec. IV B 1). Waveguides are labeled as /$\mathrm{a}^*_{(s,)l_c}$/, /$\mathrm{i}^*_{(s,)l_c}$/, /$\mathrm{u}^*_{(s,)l_c}$/ with superscript $* \in \{\mathrm{M, R}\}$ indicating the wall material and, where appropriate, s indexing selections from similar SWVT parameter subsets.

For deformable molded waveguides (M), the first resonance frequency $\bar{f}_{1,M}$ is plotted. This frequency is defined to reflect the effects of vibro-acoustic coupling with a wall vibration frequency (≈ 200 Hz) specific to the experimental setup, which only occurs in molded waveguides.[34] This coupling leads to amplitude modulation that shifts the first resonance to a lower frequency $\bar{f}_{1,l}$ (around 210 Hz) and generates an additional resonance $\bar{f}_{1,h}$ (around 640 Hz). For molded waveguides, this modulation is accounted for by defining the first resonance as $\bar{f}_{1,M} = |\bar{f}_{1,h} - \bar{f}_{1,l}|$. Thus, $(\bar{f}_{1,M}, \bar{f}_2)$ are reported for molded deformable waveguides, while $(\bar{f}_1, \bar{f}_2)$ are reported for rigid waveguides.

1. One constriction SWVT configurations

Experimentally measured resonance frequencies $(\bar{f}_{1(,M)}, \bar{f}_2)$ for one constriction configurations are presented in Fig. 12(a) (full and zoomed view in Fig. 11 of the supplementary material), based on the SWVT parameter subset ($L_0 = 180$ mm, $\mathrm{l_c} \in \{0, 30, 60\}$ mm, $x_c, \mathcal{P}$). Overall, the measured resonances for rigid waveguides (+) closely match or even overlap with the theoretical SWVT resonances (□). Since the SWVT parameters were selected to approximate the target vowel formants, this agreement validates the SWVT framework for one constriction configurations with $L_0 = 180$ mm and $l_c \leq 60$ mm.

Discrepancies are observed between the measured resonances for deformable waveguides (•) and both the theoretical SWVT resonances (□) and the ones measured for rigid waveguides (+). While these discrepancies remain limited, or even negligible, for the American English vowels /a/ (e.g., a_0^M in proximity to a_0^R) and /u/ (e.g., $u_{60,0}^M$ near $u_{60,0}^R$), they become pronounced for /i/. For instance, comparing resonances for $i_{0,30,69}^M$ and for $i_{0,30,60}^R$, this discrepancy is entirely due to a shift in the first resonance frequency, i.e., between $\bar{f}_1$ and $\bar{f}_{1,M}$. Since the first resonance frequency

https://doi.org/10.1121/10.0042185

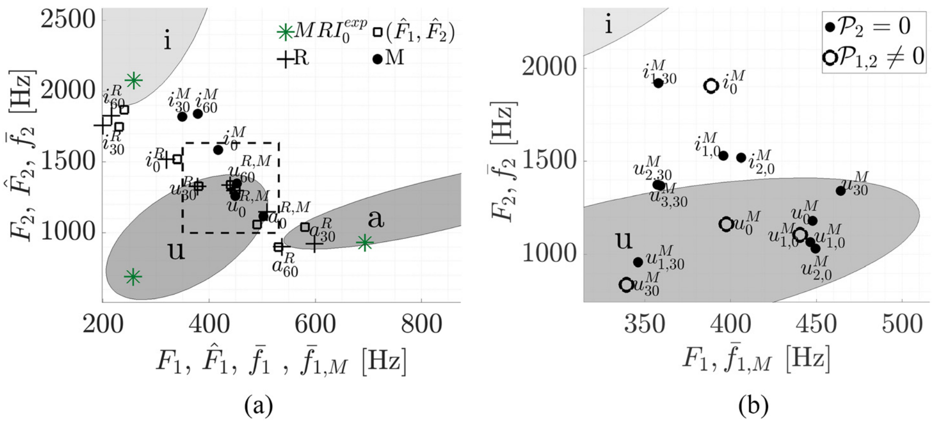


FIG. 12. Theoretical $(\hat{F}_1, \hat{F}_2)$ and measured resonance frequencies $(\bar{f}_{1,M}, \bar{f}_2)$ for deformable (M) SWVT waveguides and $(\bar{f}_1, \bar{f}_2)$ for rigid (R) waveguides (SWVT, $\mathrm{MRI}_0^{\mathrm{exp}}$) with length $L_0 = 180$ mm; formant spaces (F_1, F_2): (a) one constriction ($\mathcal{P} = 0\%$) cases selected from SWVT subset (L_0, $l_c \in \{0, 30, 60\}$ mm, x_c, $\mathcal{P}$); (b) one ($\mathcal{P}_2 = 0\%$) and two ($\mathcal{P}_{1,2} \neq 0\%$) constriction cases, selected from SWVT subsets (L_0, $l_c \in \{0, 30\}$ mm, x_c, $\mathcal{P}$) and (L_0, $l_{c,1} \in \{0, 30\}$ mm, $x_{c,1}$, $\mathcal{P}_1$, $l_{c,2} \in \{0, 30\}$ mm, $x_{c,2}$, $\mathcal{P}_2$) with $l_{c,1} = l_{c,2}$. (a) $\mathcal{P} = 0\%$, (b) $\mathcal{P} = 0\%$, $\mathcal{P}_{1,2} \neq 0\%$).

associated with selected SWVT parameters for /i/ is relatively low compared to /a/ or /u/, it is hypothesized that vibro-acoustic coupling prevents the first resonance frequency $\bar{f}_{1,M}$ to reach low $\hat{F}_1 \lesssim 350$ Hz. Nevertheless, the measured results support the validity of the SWVT framework for configurations characterized by $\hat{F}_1 \gtrsim 350$ Hz.

2. One and two constriction SWVT configurations

Experimentally measured resonance frequencies $(\bar{f}_{1,M}, \bar{f}_2)$ for both one ($\mathcal{P}_2 = 0\%$) and two ($\mathcal{P}_{1,2} \neq 0\%$) constriction configurations are presented in Fig. 12(b). These are based on selections from the SWVT parameter subset ($L_0 = 180$ mm, $\mathrm{l_c} \in \{0, 30\}$ mm, x_c, $\mathcal{P}$) for one constriction, and (L_0=180mm,$\mathrm{l}_{c,1}\in\{0,30\}$mm,$x_{c,1}$,$\mathcal{P}_1$,$l_{c,2}\in\{0,30\}$mm, $x_{c,2}$,$\mathcal{P}_2$) with $l_{c,1}=l_{c,2}$ for two constrictions.

All measurements are obtained using deformable molded waveguides, with $\bar{f}_{1,M} \gtrsim 350$ Hz, consistent with the findings for one constriction configurations described in Sec. VI B 1. Overall, $350 \lesssim \bar{f}_{1,M} \lesssim 480$ Hz is observed. Notably, when comparing, for example, i_0^M (°, $\mathcal{P}_{1,2} \neq 0$) and $i_{1(2),0}^M$ (•, $\mathcal{P}_2 = 0$) or u_{30}^M (°, $\mathcal{P}_{1,2} \neq 0$) and $i_{1,30}^M$ (•, $\mathcal{P}_2 = 0$), it is evident that introducing two constrictions extends the experimentally accessible $\bar{f}_2$-range to both higher (as observed for i^M) and lower (as observed for u^M) frequencies. This confirms the theoretical trend described in Sec. III B.

VII. SWVT VALIDATION: CONSTRICTION FEATURES

A. Influence of pincher extent, l_c: One constriction

Both the theoretical SWVT resonances (Sec. III) and experimental results (Sec. VI B) support the conclusion that the constriction length l_c contributes to achieving the target vowel formant space. Therefore, the influence of the constriction extent $l_c \leq 60$ mm is experimentally investigated using a molded waveguide with a fixed length $L_0 = 180$ mm, for one constriction configurations. The constriction position is fixed at $x_c = -90$ mm, centered within the waveguide, to minimize the potential impact of physical constraints at the inlet or outlet, as outlined in Sec. V B 1. For each assessed l_c, the constriction degree $\mathcal{P}$ is varied from 0% (no constriction) to 99% (near full closure). The measured values of $\bar{f}_{1,M}(\mathcal{P})$ and $\bar{f}_2(\mathcal{P})$ for this SWVT parameter subset ($L_0 = 180$ mm, $l_c \in \{0, 30, 60\}$ mm, $x_c = -90$ mm, $\mathcal{P}$) are plotted in Fig. 13. The general trend observed for both $\bar{f}_{1,M}(\mathcal{P})$ and $\bar{f}_2(\mathcal{P})$ is not altered by variations in l_c.

The first resonance frequency $\bar{f}_{1,M}(\mathcal{P})$ gradually increases (despite that fluctuations are observed), reaching a maximum around $\mathcal{P} \approx 80\%$, after which it declines rapidly with further increases in $\mathcal{P}$. This peak is more pronounced for $l_c = 60$ mm, while the subsequent decline is steeper for $l_c = 30$ mm. The frequency range $435 \lesssim \bar{f}_{1,M}(\mathcal{P}) \lesssim 465$ Hz

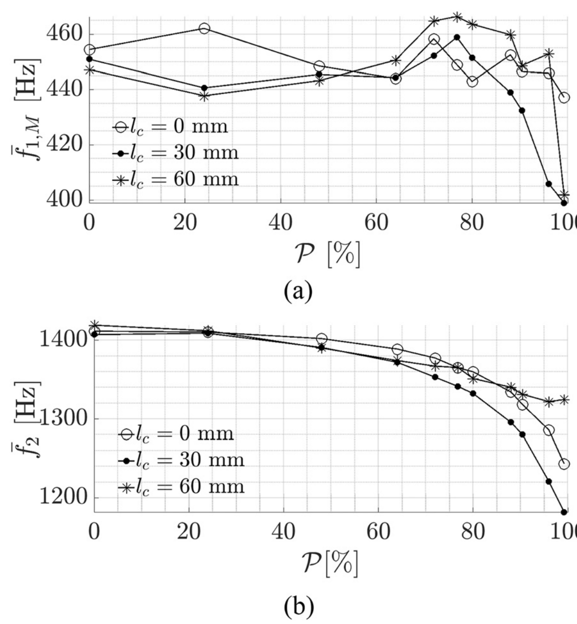


FIG. 13. Resonance frequencies of a molded waveguide ($L_0 = 180$ mm) with one constriction at $x_c = -90$ mm for $l_c \in \{0, 30, 60\}$ mm and $0 \leq \mathcal{P} \leq 99\%$: (a) $\bar{f}_{1,M}(\mathcal{P})$, (b) $\bar{f}_2(\mathcal{P})$. (a) $\bar{f}_{1,M}(\mathcal{P}; l_c)$, (b) $\bar{f}_2(\mathcal{P}; l_c)$.

observed for $l_c = 0$ mm extends by ≈ 70 Hz to $400 \lessapprox \bar{f}_{1,M}(\mathcal{P}) \lessapprox 470$ Hz for $l_c > 0$ mm. For the second resonance frequency $\bar{f}_2(\mathcal{P})$, a general decrease is observed with increasing $\mathcal{P}$, with the most pronounced decline occurring at $l_c = 30$ mm. The range $1250 \lessapprox \bar{f}_2(\mathcal{P}) \lessapprox 1400$ Hz for $l_c = 0$ mm also expands by ≈ 70 Hz to $1180 \lessapprox \bar{f}_2(\mathcal{P}) \lessapprox 1400$ Hz for $l_c > 0$ mm. These results confirm that increasing the constriction length broadens the frequency ranges associated with both the first and second resonance frequencies.

B. Influence of constriction spacing, S_c: Two constrictions

Both the theoretical SWVT resonances (Sec. III) and experimental results (Sec. VI B) support the conclusion that, in two constriction configurations, the spacing between constrictions contributes to achieving the target vowel formant space. Therefore, the influence of constriction spacing, defined as $S_c = |x_{c,2}| - |x_{c,1}|$ (in mm), is experimentally assessed. Concretely, resonance frequencies are measured using a molded waveguide ($L_0 = 180$ mm, $l_{c,1} = l_{c,2} = 0$ mm) with $\mathcal{P}_2$ fixed either at 77% or 90%, while $\mathcal{P}_1 \in \{77, 85, 90, 93, 96\}$% is varied for each spacing value in $S_c \in \{30, 60, 120\}$ mm. Large constriction degrees are assessed, as they characterize the selected SWVT parameters used to represent vowels [and typical vowels area reduction degrees $\min(A(x))/A_0 \approx 93\% - 97\%$ (Fig. 1 in the supplementary material)]. The constriction spacing is symmetrical centered around $x = -56.23$ mm. The measured values of $\bar{f}_{1,M}(\mathcal{P})$ and $\bar{f}_2(\mathcal{P})$ are presented in Fig. 14.

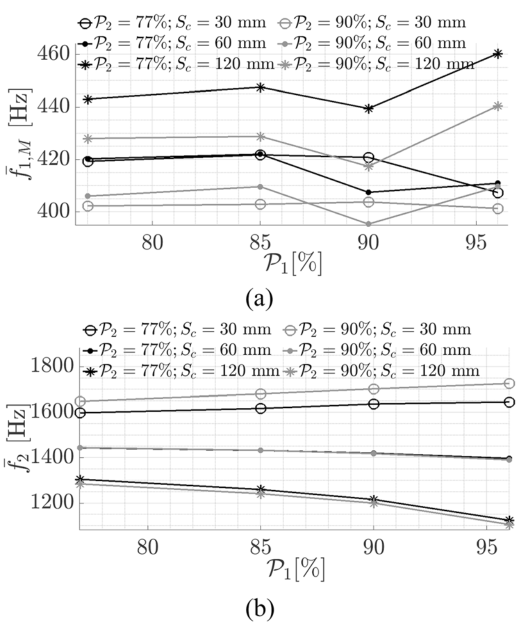


FIG. 14. Resonance frequencies of a molded waveguide ($L_0 = 180$ mm, $l_{c,1} = l_{c,2} = 0$ mm) for two constrictions spaced symmetrically around $x = -56.23$ mm. $\mathcal{P}_2$ is fixed at 77% or 90%, while $\mathcal{P}_1 \in \{77, 85, 90, 93, 96\}$ (in %) is varied for each constriction spacing value $S_c \in \{30, 60, 120\}$ mm: (a) $\bar{f}_{1,M}(\mathcal{P}_1; S_c, \mathcal{P}_2)$, (b) $\bar{f}_2(\mathcal{P}_1; S_c, \mathcal{P}_2)$. (a) $\bar{f}_{1,M}(\mathcal{P}_1; S_c, \mathcal{P}_2)$, (b) $\bar{f}_2(\mathcal{P}_1; S_c, \mathcal{P}_2)$.

The impact of S_c on the first resonance frequency $\bar{f}_{1,M}(\mathcal{P}_1)$ becomes more noticeable as S_c increases. However, for $S_c \leq 120$ mm, the overall effect on the frequency range of $\bar{f}_{1,M}$ remains limited, with an extension of no more than $\lessapprox 50$ Hz, regardless of $\mathcal{P}_2$. It is also observed that varying $\mathcal{P}_1$ can either decrease $\bar{f}_{1,M}$ (e.g., when increasing $\mathcal{P}_1$ from 85% to 90%) or increase it (e.g., when increasing $\mathcal{P}_1$ from 90% to 95%). Thus, compared to the influence of l_c on $\bar{f}_{1,M}$—which was found (Sec. VII A) to extend the frequency range by up to $\lessapprox 70$ Hz—the influence of S_c is similar or slightly smaller magnitude. Therefore, findings for $\bar{f}_{1,M}$ support the statement in Sec. IV B 2, that adding a second constriction does not significantly improve the accuracy of the results compared to single-constriction models, particularly when $l_c > 0$ mm.

For closely spaced constrictions ($S_c \in \{30, 60\}$ mm), the second resonance frequency $\bar{f}_2$ remains fairly constant, varying by no more than approximately 70 Hz, regardless of $\mathcal{P}_2$ and $\mathcal{P}_1$. In contrast, for widely spaced constrictions ($S_c = 120$ mm) $\bar{f}_2$ decreases by approximately 200 Hz with increasing $\mathcal{P}_1$, independent of $\mathcal{P}_2$. Furthermore, increasing the spacing from 30 to 120 mm reduces $\bar{f}_2$ by around 300 Hz for $\mathcal{P}_1 = 77\%$, and by up to 600 Hz for $\mathcal{P}_1 = 96\%$, regardless of $\mathcal{P}_2$. These results demonstrate that, at high constriction degrees, the spacing between constrictions can have a substantial impact on the observed second resonance frequency $\bar{f}_2$, thereby supporting the conclusion in Sec. VI B 2 and illustrated in Fig. 12 (comparing i_0^M where $\mathcal{P}_{1,2} \neq 0\%$, with $i_{1(2),0}^M$, where $\mathcal{P}_2 = 0\%$).

C. Measured resonance frequencies: Influence of l_c and S_c

Figure 15 (zoomed views in Fig. 12 of the supplementary material) presents an overview of the measured resonance frequencies ($\bar{f}_{1,M}$,$\bar{f}_2$), illustrating the influence of constriction length l_c (one constriction configurations, black) and constriction spacing S_c (two constriction configurations,

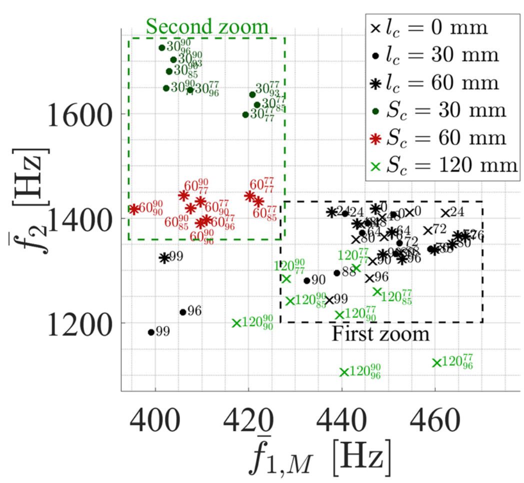


FIG. 15. Measured resonance frequencies ($\bar{f}_{1,M}$,$\bar{f}_2$): Influence of pincher extent length l_c for one constriction configurations (black symbol and $\mathcal{P}$) and constriction spacing S_c for two constriction configurations with $l_{c,1} = l_{c,2} = 0$ mm (non-black symbol and $S_{c\mathcal{P}_2}^{\mathcal{P}_1}$).

non-black), as discussed in Secs. VII B and VII C, respectively. Overall, the color clusters indicate that these parameters allow exploration of distinct regions within the resonance space $(\overline{f}_{1,M}, \overline{f}_2)$. This demonstrates the relevance and utility of these parameters within the SWVT framework.

The measured first resonance $\overline{f}_{1,M}$ varies from 390 up to 470 Hz, while the second resonance $\overline{f}_2$ spans from 1100 up to 1800 Hz. This frequency range allows to reach up to the ellipsoids associated with American English vowels. However, the region where $\overline{f}_{1,M} \gtrapprox 420$ Hz and $\overline{f}_2 \gtrapprox 1500$ Hz coincide is not explored by the experimentally assessed parameters. This omission is partly expected, given that the experiments were conducted using a waveguide of fixed length $L0 = 180$ mm. As discussed in relation to Fig. 3(a) (Sec. III A 2), varying L_0 enables exploration of the vowel formant space along the direction corresponding to the mean trend $\hat{F}_{2mean}(\hat{F}_{1mean})$, which aligns with the unexplored region observed in Fig. 15.

VIII. CONCLUSION

Overall, the theoretical SWVT resonance space $(\hat{F}_1, \hat{F}_2)$ for one constriction, defined by the SWVT subset $(L_0, l_c, x_c, \mathcal{P})$, forms a tilted, butterfly-shaped isosceles trapezium. Its position and overlap with the formant space (F_1, F_2) are influenced by the waveguide length L_0, which shifts the mean position of $(\hat{F}_1, \hat{F}_2)$, and by the minimum constriction length l_c, which affects the spacing of the extrema and thus the width of the SWVT resonance space. The same tendencies were observed considering two constrictions. It was demonstrated that when increasing the extend length of at least one of the two pinchers, the SWVT framework approximates the formants (F_1, F_2) of American English vowels using $L_0 = 180$ mm.

This motivates the investigation of selected SWVT parameter subsets for these vowels by introducing a well-defined objective selection criterion ($\mathrm{MRI}_{\mathrm{II}}$) based on the theoretical resonance frequencies, or resonance invariants, derived from the VT area functions reported by Story, adapted to the geometrical domain of a cylindrical uniform waveguide with a fixed length of $L_0 = 180$ mm. The general features of the SWVT area functions corresponding to the selected parameters under the $\mathrm{MRI}_{\mathrm{II}}$-criterion, such as severe constriction degree and constriction location, are consistent with those observed in the VT area functions. However, to fully cover the vowel formants space, the SWVT parameter subset must either include constriction lengths $l_c \geq 0$ mm within one constriction configurations, be extended to include two constriction configurations, or account for both (i.e., two constriction configurations with $l_{c,1(2)} \geq 0$ mm). This underscores the relevance of the parameters in the SWVT framework for examining the relationship between acoustic resonance frequencies and area functions. For future studies it is of interest to compare selected SWVT configurations with other geometric VT models of vowels.

Experimental validation was assessed using rigid and deformable waveguides for selected SWVT parameter sets representing American English vowels /i/, /a/, and /u/. Overall, the experimental results align with theoretical resonance frequencies, thereby validating the selected SWVT parameter sets for both waveguide types. Notably, validation on deformable waveguides is particularly valuable, as it enables the assessment of constriction degrees up to 99%, a range not easily achievable with 3D printed rigid waveguides and offers the potential for dynamic variation of SWVT parameters for (aero-)acoustic studies.

Beyond the immediate perspectives discussed above, this work opens several promising directions for future research. The SWVT framework provides a systematic tool for exploring the relationship between vocal tract geometry and acoustic resonances, and it can be readily extended to additional vowels, consonants, or dynamic articulatory gestures. Its simple yet flexible parametric representation offers a robust foundation for researchers developing computational or physical models of the vocal tract. Furthermore, integration of the SWVT approach with inverse-modeling or machine-learning techniques may enable prediction of articulatory configurations from acoustic data, with potential impact on speech synthesis, articulatory phonetics, and clinical speech-therapy applications. The methodological concept demonstrated here—controlling resonance frequencies through one or two localized deformations of an otherwise cylindrical waveguide—also suggests applications beyond speech science. Similar principles could inform the design of ducts, pipes, and acoustic enclosures in automotive or aeronautical engineering, where targeted resonance control is essential for noise reduction, sound-quality enhancement, or rapid prototyping. Overall, the SWVT framework not only advances theoretical and experimental studies of speech production but also provides a versatile tool with broader relevance to engineering acoustics.[43–45]

SUPPLEMENTARY MATERIAL

See the supplementary material for a comprehensive overview of used literature on vowel formants and area functions, additional theoretical results for the SWVT framework and details of its experimental validation.

ACKNOWLEDGEMENTS

Work partly supported by the Full3DTalkingHead project (No. ANR-20-CE23-0008-03).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

DATA AVAILABILITY

Data available on request from the authors.

[1]G. Peterson and H. Barney, "Control methods used in a study of the vowels," J. Acoust. Soc. Am. **24**, 175–184 (1952).
[2]L. Calliope and G. Fant, *La Parole et Son Traitement Automatique* (*Speech and its Automatic Processing*) (Masson, Paris, France, 1989).
[3]D. Kewley-Port and C. Watson, "Formant-frequency discrimination for isolated English vowels," J. Acoust. Soc. Am. **95**, 485–496 (1994).
[4]J. Hillenbrand, L. Getty, M. Clark, and K. Wheeler, "Acoustic characteristics of American English vowels," J. Acoust. Soc. Am. **97**, 3099–3111 (1995).
[5]C. Gendrot and M. Adda-Decker, "Impact of duration on F1/F2 formant values of oral vowels: An automatic analysis of large broadcast news corpora in French and German," in *Interspeech 2005* (2005), pp. 2453–2456.
[6]G. Fant, *Acoustic Theory of Speech Production: With Calculations Based on X-Ray Studies of Russian Articulations* (Mouton, Berlin, Germany, 1971).
[7]G. Fant, "The relations between area functions and the acoustic signal," Phonetica **37**, 55–86 (1980).
[8]M. Labrunie, P. Badin, D. Voit, A. Joseph, J. Frahm, L. Lamalle, C. Vilain, and L.-J. Boë, "Automatic segmentation of speech articulators from real-time midsagittal MRI based on supervised learning," Speech Commun. **99**, 27–46 (2018).
[9]I. Douros, Y. Xie, C. Dourou, K. Isaieva, P.-A. Vuissoz, J. Felblinger, and Y. Laprie, "3D dynamic spatiotemporal atlas of the vocal tract during consonant–vowel production from 2D real time MRI," J. Imag. **8**, 227 (2022).
[10]K. Isaieva, F. Odille, Y. Laprie, G. Drouot, J. Felblinger, and P.-A. Vuissoz, "Super-resolved dynamic 3D reconstruction of the vocal tract during natural speech," J. Imag. **9**, 233 (2023).
[11]V. Ribeiro, K. Isaieva, J. Leclere, J. Felblinger, P.-A. Vuissoz, and Y. Laprie, "Automatic segmentation of vocal tract articulators in real-time magnetic resonance imaging," Comput. Methods Programs Biomed. **243**, 107907 (2024).
[12]B. Story, "Comparison of magnetic resonance imaging-based vocal tract area functions obtained from the same speaker in 1994 and 2002," J. Acoust. Soc. Am. **123**, 327–335 (2008).
[13]P. Lieberman, "The evolution of human speech: Its anatomical and neural bases," Curr. Anthropol. **48**, 39–66 (2007).
[14]J. Eichhorn, R. Kent, D. Austin, and H. Vorperian, "Effects of aging on vocal fundamental frequency and vowel formants in men and women," J. Voice **32**, 644.e1–644.e9 (2018).
[15]B. Story, H. Vorperian, K. Bunton, and R. Durtschi, "An age-dependent vocal tract model for males and females based on anatomic measurements," J. Acoust. Soc. Am. **143**, 3079–3102 (2018).
[16]M. DiNino, "Age and masking effects on acoustic cues for vowel categorization," JASA Express Lett. **4**, 060001 (2024).
[17]B. Story, I. Titze, and E. Hoffman, "The relationship of vocal tract shape to three voice qualities," J. Acoust. Soc. Am. **109**, 1651–1667 (2001).
[18]K. Tom, I. Titze, E. Hoffman, and B. Story, "Three-dimensional vocal tract imaging and formant structure: Varying vocal register, pitch, and loudness," J. Acoust. Soc. Am. **109**, 742–747 (2001).
[19]J. Tanner, "Prosodic and durational influences on the formant dynamics of Japanese vowels," JASA Express Lett. **3**, 085202 (2023).
[20]B. De Boer and W. Tecumseh Fitch, "Computer models of vocal tract evolution: An overview and critique," Adapt Behav. **18**, 36–47 (2010).
[21]T. Chiba and M. Kajiyama, *The Vowel, Its Nature and Structure* (Phonetic Society of Japan, Tokyo, Japan, 1958).
[22]H. Dunn, "The calculation of vowel resonances, and an electrical vocal tract," J. Acoust. Soc. Am. **22**, 740–753 (1950).
[23]E. L. Saltzman and K. G. Munhall, "A dynamical approach to gestural patterning in speech production," Ecol. Psychol. **1**, 333–382 (1989).
[24]S. Maeda, "Compensatory articulation during speech: Evidence from the analysis and synthesis of vocal-tract shapes using an articulatory model," in *Speech Production and Speech Modelling*, pp. 131–149.
[25]C. Browman and L. Goldstein, "Articulatory phonology: An overview," Phonetica **49**, 155–180 (1992).
[26]R. Wilhelms-Tricarico, "A biomechanical and physiologically-based vocal tract model and its control," J. Phon. **24**, 23–38 (1996).
[27]B. Story and I. Titze, "Parameterization of vocal tract area functions by empirical orthogonal modes," J. Phon. **26**, 223–260 (1998).
[28]P. Mokhtari, T. Kitamura, H. Takemoto, and K. Honda, "Principal components of vocal-tract area functions and inversion of vowels by linear regression of cepstrum coefficients," J. Phon. **35**, 20–39 (2007).
[29]J. Gaines, K. Kim, B. Parrell, V. Ramanarayanan, S. Nagarajan, and J. Houde, "Discrete constriction locations describe a comprehensive range of vocal tract shapes in the Maeda model," JASA Express Lett. **1**, 124402 (2021).
[30]J. Lucero, "Parametrization of the vocal tract area function using a subset selection approach," J. Acoust. Soc. Am. **149**, 4036–4038 (2021).
[31]M. Arnela, S. Dabbaghchian, R. Blandin, O. Guasch, O. Engwall, A. Van Hirtum, and X. Pelorson, "Influence of vocal tract geometry simplifications on the numerical simulation of vowel sounds," J. Acoust. Soc. Am. **140**, 1707–1718 (2016).
[32]K. Stevens, *Acoustic Phonetics* (MIT Press, Cambridge, MA, 2000).
[33]R. Blandin, M. Arnela, R. Laboissière, X. Pelorson, O. Guasch, A. Van Hirtum, and X. Laval, "Effects of higher order propagation modes in vocal tract like geometries," J. Acoust. Soc. Am. **137**, 832–843 (2015).
[34]A. Eliraki, "Physical study of the influence of waveguide properties on vocal tract acoustics," Ph.D. thesis, University Grenoble Alpes, Saint-Martin-d'Hères, France (2025), pp. 24–51, submitted.
[35]G. Yeung, S. Lulich, J. Guo, M. Sommers, and A. Alwan, "Subglottal resonances of American English speaking children," J. Acoust. Soc. Am. **144**, 3437–3449 (2018).
[36]J. Flanagan, *Speech Analysis Synthesis and Perception* (Springer, Berlin, Germany, 1965).
[37]J. Kelly and C. Lochbaum, "Speech synthesis," in *Proceedings of the 4th International Congress on Acoustics* (1962), pp. 1–4.
[38]M. Sondhi and J. Schroeter, "A hybrid time-frequency domain articulatory speech synthesizer," IEEE Trans. Acoust. **35**, 955–967 (1987).
[39]A. Van Hirtum, R. Blandin, and X. Pelorson, "Validation of an analytical compressed elastic tube model for acoustic wave propagation," J. Appl. Phys. **118**, 224905 (2015).
[40]P. Morse and K. Ingard, *Theoretical Acoustics* (Princeton University Press, Princeton, NJ, 1986).
[41]A. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications* (Springer, Berlin, Germany, 2019).
[42]A. Van Hirtum, R. Blandin, and X. Pelorson, "A setup to study aeroacoustics for finite length ducts with time-varying shape," Appl. Acoust. **105**, 83–92 (2016).
[43]B. S. Atal, J. Chang, M. Mathews, and J. Tukey, "Inversion of articulatory-to-acoustic transformation in the vocal tract by a computer-sorting technique," J. Acoust. Soc. Am. **63**, 1535–1555 (1978).
[44]S. Ouni and Y. Laprie, "Modeling the articulatory space using a hypercube codebook for acoustic-to-articulatory inversion," J. Acoust. Soc. Am. **118**, 444–460 (2005).
[45]S. Panchapagesan and A. Alwan, "A study of acoustic-to-articulatory inversion of speech by analysis-by-synthesis using chain matrices and the maeda articulatory model," J. Acoust. Soc. Am. **129**, 2144–2162 (2011).