MADRID
inter.noise 2019
June 16 - 19
NOISE CONTROL FOR A BETTER ENVIRONMENT

Modal perturbation analysis instead of nonlinear radiation pressure to derive the area sensitivity function for resonance tuning in an axisymmetric duct with variable cross-section

Guasch, Oriol[1]
GTM - Grup de recerca en Tecnologies Mèdia, La Salle, Universitat Ramon Llull
C/ Quatre Camins 2, 08022 Barcelona, Catalonia, Spain.

Arnela, Marc[2]
GTM - Grup de recerca en Tecnologies Mèdia, La Salle, Universitat Ramon Llull
C/ Quatre Camins 2, 08022 Barcelona, Catalonia, Spain.

ABSTRACT

Axisymmetric ducts with variable cross-section are of importance in many acoustic problems ranging from horn theory to vocal tract acoustics. Webster's equation is commonly used to describe their performance in the plane wave propagation regime. In some problems, mostly related to voice generation, one is interested in modifying the area of the duct cross-sections to adjust the frequency of a resonance. For instance, one may want to increase its value, or to bring a group of resonances closer together, to emulate effects that occur in natural voice production. To that goal, an optimization iterative process can be followed in which the cross sections are subsequently changed, according to an area sensitivity function, until the resonances of the duct are placed at the target position. Traditionally, the area sensitivity functions have been derived from the non-linear radiation pressure inside the duct. In this work we demonstrate there is no need to resort to such non-linear phenomenon because the same result can be deduced from a first order modal perturbation analysis of the duct eigenfrequencies. After proving that, we present some simulations in the framework of expressive vowel production.

Keywords: Modal perturbation analysis, formant variation, numerical voice generation, duct acoustics, variable cross-sectional area
I-INCE Classification of Subject Number: 76, 34
http://i-ince.org/files/data/classification.pdf

1. INTRODUCTION

Axisymmetric ducts with variable cross-sectional area are encountered in horn theory, musical instruments, large pipe installations or speech production. Determining how duct resonances change when modifying duct parameters such as the cross-sectional areas, the tube length or the wall damping, may serve several purposes in all those areas of acoustics. In voice production, for instance, the vocal tract can be approximated by an axisymmetric tube whose shape can evolve to produce expressivity effects when pronouncing a simple sound like a vowel (see e.g., [1,2]). The resonances of the vocal tract are usually referred to as formants in the speech community and the first two of them

[1]oriol.guasch@salle.url.edu
[2]marc.arnela@salle.url.edu

allow one to distinguish one vowel from another. The dependence of the formant frequency locations on the vocal tract parameters has been therefore a classical topic of research since the pioneering work in [3].

A first connection was established by Schroeder [4], who showed that formant variations were proportional to the radiation pressure (see e.g., [5]) at a vocal tract constriction. Since that work, the radiation pressure has remained as the main physical mechanism (circuit analogies aside) upon which sensitivity functions have been built and justified to move the vocal tract formants to targeted values, following an iterative optimization process for area and/or length perturbations [6,7]. However, the radiation pressure is essentially a non-linear phenomenon and one needs to resort to the second order wave equation to get non-null values for it.

It is the main purpose of this work to show there is actually no need to resort to the non-linear pressure radiation to determine the area perturbation sensitivity functions for formant tuning. The same functions can be derived from a standard modal perturbation analysis (see e.g., [8,9]) of the discretized linear Webster equation, which describes planar wave propagation inside the duct (it is to be noted that similar results could be obtained for duct length perturbations, but they will be not reported herein).

Despite of huge progress in numerical voice production, which allows one to simulate the generation of static sounds in complex three-dimensional (3D) vocal tracts using finite element FEM approaches (see e.g., [10-13]), the 3D generation of dynamic sounds as diphthongs is still challenging and requires substantial computational effort [14,15]. A 3D optimization process for formant tuning would therefore prove very costly, and one could think in converting the 3D optimization problem into a 1D one (relying on Webster's equation), and, once solved, revert to 3D geometries. This work reports some initial, necessary results, towards that goal.

2. SENSITIVITY FUNCTIONS BASED ON RADIATION PRESSURE

2.1 Ehrenfest's theorem and the acoustic radiation pressure

Let us consider the case of a rigid axisymmetric duct with centerline in the x direction, variable cross-sectional area $A(x)$, and open boundary conditions at both ends. Assume that we can induce local changes $\delta A(x)$ to modify the duct resonances. As said in the Introduction, a typical situation would be that of an idealization of the human vocal tract, whose geometry can be distorted to move its formants and attain a certain expressivity effect when pronouncing a vowel sound.

According to Schroeder [4], the Ehrenfest theorem allows one to relate the frequency perturbation δf_n of a resonance f_n, with the time averaged variation $\delta\bar{E}_n$ of the modal mechanical energy, $\bar{E}_n$, when adiabatically changing the cross section from $A(x)$ to $A(x)+\delta A(x)$, namely,

$$\frac{\delta f_n}{f_n} = \frac{\delta \bar{E}_n}{\bar{E}_n}. \tag{1}$$

Here and in what follows, an overbar designates a time averaged quantity. The increase/decrease of mechanical energy $\delta\bar{E}_n$ in Equation 1 is associated to the work done by the radiation force on the duct walls. Recall that the Reynolds form of the momentum equation for an inviscid fluid reads

$$\partial_t(\rho \boldsymbol{u}) + \nabla \cdot \boldsymbol{M} = 0, \tag{2}$$

where ρ stands for the fluid density, $\boldsymbol{u}$ for its velocity and $\boldsymbol{M}$ for the momentum flux density tensor,

$$\boldsymbol{M} = (p - p_0)\boldsymbol{I} + \rho \boldsymbol{u} \otimes \boldsymbol{u}, \tag{3}$$

with p and p_0 respectively being the total and constant ambient pressures. For a stationary sound field, the time average of Equation 2 produces

$$\nabla \cdot \bar{\boldsymbol{M}} = 0. \tag{4}$$

The divergence-free tensor $\bar{\boldsymbol{M}}$ is called the radiation stress tensor [5]. The mean excess pressure $\overline{p - p_0}$ in $\bar{\boldsymbol{M}}$ will be herein identified with the acoustic radiation pressure (it should be noted that this term is often used to directly designate $\bar{\boldsymbol{M}}$, which is somewhat misleading). The radiation force exerted by a wave on an immersed object with surface Γ and normal vector $\boldsymbol{n}$ is given by

$$\boldsymbol{F}^R = \int_\Gamma \bar{\boldsymbol{M}} \cdot \boldsymbol{n} d\Gamma. \tag{5}$$

In the present situation, we consider plane wave propagation in the axis direction inside the duct, so the radiation stress tensor has entries

$$\bar{M}_{ij}(x) = \left(\overline{p(x) - p_0}\right)\delta_{ij} + \overline{\rho u_i(x) u_j(x)}, \tag{6}$$

with u_x being the sole non-null velocity component. The radiation stress tensor in Equation 6 can be transformed to cylindrical coordinates to find the radiation force exerted on the duct walls, in the radial direction. The radial force at coordinate *x* inside the duct is simply given by

$$F_r^R(x) = \overline{p(x) - p_0}\, A(x). \tag{7}$$

As mentioned before, the variation of the modal mechanical energy $\delta\bar{E}_n$ in Equation 1 will equal the work performed by the radiation force when modifying the cross-section area from $A(x)$ to $A(x) + \delta A(x)$, so that,

$$\delta\bar{E}_n = -\bar{W} = -\int_0^L \overline{p(x) - p_0}\, \delta A(x)\, dx. \tag{8}$$

The radiation pressure $\overline{p(x) - p_0}$ in Equation 8 vanishes for standard linear wave propagation. Therefore, the mean excess radiation pressure is essentially a non-linear phenomenon. Expanding all variables in series, i.e., $p(x) = p_0 + \epsilon p_1(x) + \epsilon^2 p_2(x) + \cdots$, $\rho(x) = \rho_0 + \epsilon\rho_1(x) + \epsilon^2\rho_2(x) + \cdots$ and $\boldsymbol{u}(x) = \epsilon\boldsymbol{u}_1(x) + \epsilon^2\boldsymbol{u}_2(x) + \cdots$ one can derive the acoustic wave equation to second order and check that the radiation pressure equals minus the first order time averaged Lagrangian density [5],

$$\overline{p(x)-p_0}=\overline{p_2(x)}=-\bar{\mathcal{L}}_n(x)=\bar{U}_n(x)-\bar{K}_n(x)$$
$$=\frac{1}{2\rho_0 c_0^2}\bar{p}_1^2(x)-\frac{1}{2}\rho_0\bar{u}_1^2(x), \tag{9}$$

with $\bar{U}^n$ and $\bar{K}^n$ respectively denoting the time averaged potential and kinetic energies of the n-th mode at coordinate x, and $\bar{p}_1$ and $\bar{u}_1$ here standing for the time averaged first order acoustic pressure and acoustic particle velocity. Taking into account that the pointwise mechanical energy of the n-th resonance is given by $\bar{E}_n(x)=\bar{K}_n(x)+\bar{U}_n(x)$, Equation 1 becomes,

$$\frac{\delta f_n}{f_n}=-\frac{\int_0^L[\bar{U}_n(x)-\bar{K}_n(x)]\,\delta A(x)\,dx}{\int_0^L[\bar{U}_n(x)+\bar{K}_n(x)]\,A(x)\,dx}=\frac{\int_0^L\bar{\mathcal{L}}_n(x)\,\delta A(x)\,dx}{\bar{E}_n}, \tag{10}$$

where $\bar{E}_n$ is the total acoustic mechanical energy in the duct.

2.2 Area sensitivity functions and iterative optimization process

As noticed by Story [6], Equation 10 can be exploited to derive an iterative process to move a given resonance f_n to a target value f_n^{tg} through subsequent area perturbations. From Equations 9 and 10 we can introduce the sensitivity function S_n [6,7],

$$S_n(x)=\frac{\bar{\mathcal{L}}_n(x)A(x)}{\bar{E}_n}, \tag{11}$$

and rewrite Equation 10 as

$$\frac{\delta f_n}{f_n}=\int_0^L S_n(x)\frac{\delta A(x)}{A(x)}\,dx. \tag{12}$$

If one discretizes the duct into a series of N_s concatenated cones or cylinders, Equation 12 becomes

$$\frac{\delta f_n}{f_n}=\sum_{s=1}^{N_s}S_n^s\frac{\delta A_s}{A_s}. \tag{13}$$

To move a set of N_r resonances to new target positions, Story [6] proposes the following iterative scheme to modify the cross-sectional areas and achieve the final goal,

$$A_{k+1}(s)=A_k(s)+\sum_{n=1}^{N_r}z_{n_k}S_{n_k}(s), \tag{14}$$

where $s=1\cdots N_s$ and $k=0\cdots N_{\text{iter}}$, N_{iter} being the number of iterations to be performed. The coefficients z_{n_k} are determined at each iteration by

$$z_{n_k} = \alpha \frac{f_n^{tg} - f_{n_k}}{f_{n_k}}, \tag{15}$$

α being a speed up factor, typically set to $\alpha = 10$. The iterations continue until $\sqrt{\sum \left(f_n^{tg} - f_{n_k} \right)^2} \leq \varepsilon$, with ε being a predetermined tolerance value.

The key of the above iteration procedure relies on the sensitivity function of Equation 11, which has been deduced from the non-linear acoustic radiation pressure in Equation 9. Traditionally, this has been the way to justify the effects of cross-sectional area changes on the frequency locations of resonances. However, it shall be proved in subsequent sections that the same sensitivity functions can be deduced without the need for second order non-linearity. A first order, standard modal perturbation analysis of the duct resonances suffices to that purpose.

3. SENSITIVITY FUNCTIONS FROM MODAL PERTURBATION ANALYSIS

3.1 Discrete solution to Webster's eigenvalue problem

The propagation of plane waves inside a duct with varying cross section $A(x)$ and length L can be described by means of the Webster equation. The resonances in such a duct for open boundary conditions at both ends can be obtained from the solution to

$$\begin{gathered} \frac{d}{dx}\left(A \frac{dp}{dx}\right) + A k_0^2 p = 0, \\ p(0) = p(L) = 0, \end{gathered} \tag{16}$$

where $k_0 = \omega / c_0$ is the wavenumber, ω the angular frequency and p now designates the linear acoustic pressure instead of p_1, to lighten the forthcoming notation. Let us solve the eigenproblem of Equation 16 using a numerical method. Given that it is a one-dimensional problem, a finite difference scheme would suffice, yet we will resort to the finite element method (FEM) to facilitate future connection with three-dimensional models as well as to benefit from the resulting energy balance relations.

The weak formulation of Equation 16 reads

$$\int_0^L A(x) \frac{dq(x)}{dx} \frac{dp(x)}{dx} dx - k_0^2 \int_0^L A(x) q(x) p(x)\, dx = 0, \tag{17}$$

with q standing for a test function that satisfies the homogeneous boundary conditions in Equation 16. If we next discretize Equation 17 partitioning the duct length $[0, L]$ into N_e finite elements $\Omega_a = (x_a, x_{a+1})$, with x_a standing for a node and $a = 0 \cdots N$, we can expand the test function and the acoustic pressure as $q(x) = \sum_{a=0}^{N} N^a(x) Q^a$ and $p(x) = \sum_{b=0}^{N} N^b(x) P^b$, where the $N^a(x)$ correspond to polynomial basis functions and Q^a and P^b designate nodal values (i.e. $Q^a \equiv q(x_a)$ and $P^b \equiv p(x_b)$). Substituting the expansions into Equation 17 provides the algebraic matrix system

$$(-\omega^2 \boldsymbol{M} + \boldsymbol{K}) \boldsymbol{P} = \boldsymbol{0}, \tag{18}$$

with $\boldsymbol{P}$ representing the vector of unknown nodal pressures. The entries of the mass and stiffness matrices $\boldsymbol{M}$ and $\boldsymbol{K}$ are given by

$$M_{ab} = \frac{1}{\rho_0 c_0^2}\int_0^L A(x)N^a(x)N^b(x)\,dx,$$
$$K_{ab} = \frac{1}{\rho_0}\int_0^L A(x)\frac{dN^a(x)}{dx}\frac{dN^b(x)}{dx}dx. \tag{19}$$

The eigenproblem in Equation 18 will be satisfied for real eigenvalues (resonances/formants) ω_n with corresponding eigenvectors $\boldsymbol{\Psi}_n$ such that

$$(-\omega_n^2\boldsymbol{M} + \boldsymbol{K})\boldsymbol{\Psi}_n = \boldsymbol{0}. \tag{20}$$

If we introduce the dynamic stiffness matrix $\boldsymbol{Z} \equiv -\omega^2\boldsymbol{M} + \boldsymbol{K}$, Equation 20 can be more succinctly written as

$$\boldsymbol{Z}_n\boldsymbol{\Psi}_n = \boldsymbol{0}. \tag{21}$$

3.2 Perturbation analysis of the resonances

Our next interest is to see how the resonances ω_n get modified under a pointwise area perturbation from $A(x)$ to $A(x)+\delta A(x)$. To that purpose let us first consider the variation under an arbitrary generic parameter, say g, and follow the classical approach in [9]. If we pre-multiply Equation 21 by $\boldsymbol{\Psi_n^{\intercal}}$ to get $\boldsymbol{\Psi_n^{\intercal}Z}_n\boldsymbol{\Psi}_n = 0$ and then differentiate with respect to g we get

$$\partial_g\boldsymbol{\Psi_n^{\intercal}Z}_n\boldsymbol{\Psi}_n + \boldsymbol{\Psi_n^{\intercal}}\partial_g\boldsymbol{Z}_n\boldsymbol{\Psi}_n + \boldsymbol{\Psi_n^{\intercal}Z}_n\partial_g\boldsymbol{\Psi}_n = 0. \tag{22}$$

Because of Equation 21 and its transposed, only the middle term in Equation 22 differs from zero. Besides, and from its definition, the explicit derivative of $\boldsymbol{Z}_n$ with respect to g is given by

$$\partial_g\boldsymbol{Z}_{\mathrm{n}} = -2\omega_n\partial_g\omega_n\boldsymbol{M} - \omega_n^2\partial_g\boldsymbol{M} + \partial_g\boldsymbol{K}. \tag{23}$$

Substituting Equation 23 into Equation 22, considering that $-2\omega_n\boldsymbol{M} = \omega_n^{-1}[\boldsymbol{Z}_n - (\omega_n^2\boldsymbol{M} + \boldsymbol{K})]$ and making use of Equation 21 again, we arrive at

$$\frac{\partial_g\omega_n}{\omega_n} = \frac{\boldsymbol{\Psi_n^{\intercal}}\left(-\omega_n^2\partial_g\boldsymbol{M} + \partial_g\boldsymbol{K}\right)\boldsymbol{\Psi}_n}{\boldsymbol{\Psi_n^{\intercal}}(\omega_n^2\boldsymbol{M} + \boldsymbol{K})\boldsymbol{\Psi}_n}, \tag{24}$$

which written in terms of perturbations $\delta h = \partial_g h\delta g$ (h representing a generic function), yields,

$$\frac{\delta\omega_n}{\omega_n} = \frac{\delta f_n}{f_n} = \frac{\boldsymbol{\Psi_n^{\intercal}}(-\omega_n^2\delta\boldsymbol{M} + \delta\boldsymbol{K})\boldsymbol{\Psi}_n}{\boldsymbol{\Psi_n^{\intercal}}(\omega_n^2\boldsymbol{M} + \boldsymbol{K})\boldsymbol{\Psi}_n}. \tag{25}$$

Given that the kinetic energy of the *n*-th mode is $K_n = -(1/2)\mathrm{i}\omega_n^{-1}\boldsymbol{\Psi_n^{\intercal}K}\mathrm{i}\omega_n^{-1}\boldsymbol{\Psi}_n$ and the potential one $U_n = (1/2)\boldsymbol{\Psi_n^{\intercal}M\Psi}_n$, Equation 25 can be expressed as

$$\frac{\delta f_n}{f_n} = \frac{\delta K_n - \delta U_n}{E_n} = \frac{\delta \mathcal{L}_n}{E_n}. \tag{26}$$

Therefore, we have basically recovered the same result of Equation 10 (we have herein identified $\delta \mathcal{L}_n \equiv \int_0^L \bar{\mathcal{L}}_n(x)\delta A(x)\, dx$) but without the need to resort to any non-linear phenomenon, such as the radiation pressure. Note that Equation 26 is a very general well-known result that applies to any non-damped mass-stiffness matrix system [8,9].

3.2 Recovering the area sensitivity functions

From the above line of reasoning we can recover the area sensitivity functions in Section 2.2 as follows. The change of a matrix $\boldsymbol{X}$ under a parameter perturbation can be approximated by a finite difference, so that $\delta \boldsymbol{X} = \boldsymbol{X}(g + \delta g) - \boldsymbol{X}(g)$. For area perturbations of the mass and stiffness matrices in Equation 19 we get

$$\begin{aligned} \delta M_{ab} &\approx M_{ab}(A + \delta A) - M_{ab}(A) = \frac{1}{\rho_0 c_0^2}\int_0^L \delta A(x) N^a(x) N^b(x)\, dx, \\ \delta K_{ab} &\approx K_{ab}(A + \delta A) - K_{ab}(A) = \frac{1}{\rho_0}\int_0^L \delta A(x) \frac{dN^a(x)}{dx}\frac{dN^b(x)}{dx}\, dx. \end{aligned} \tag{27}$$

If we divide the duct into N_s elements of constant area A_s, we can expand the Lagrangian $\delta \mathcal{L}_n$ in the numerator of Equation 26 as

$$\begin{aligned} &\delta \mathcal{L}_n = \delta K_n - \delta U_n \\ &= \sum_{s=1}^{N_s} \frac{\delta A_s}{A_s} \sum_{a=1}^{N}\sum_{b=1}^{N} \frac{1}{2\omega_n^2 \rho_0} \Psi_n^{\mathrm{a}} \left[\int_{x_{s-1}}^{x_s} A_s(x) \frac{dN^a(x)}{dx}\frac{dN^b(x)}{dx} dx\right] \Psi_n^{\mathrm{b}} \\ &\quad - \sum_{s=1}^{N_s} \frac{\delta A_s}{A_s} \sum_{a=1}^{N}\sum_{b=1}^{N} \frac{1}{2\rho_0 c_0^2} \Psi_n^{\mathrm{a}} \left[\int_{x_{s-1}}^{x_s} A_s(x) N^a(x) N^b(x)\, dx\right] \Psi_n^{\mathrm{b}} \\ &\quad = \sum_{s=1}^{N_s} \frac{\delta A_s}{A_s}(K_n^s - U_n^s) = \sum_{s=1}^{N_s} \frac{\delta A_s}{A_s} \mathcal{L}_n^s, \end{aligned} \tag{28}$$

where K_n^s, U_n^s and $\mathcal{L}_n^s$ respectively denote the kinetic energy, potential energy and Lagrangian of the *s*-th element when the *n*-th acoustic resonance is excited.

For an element *s*, the sensitivity function in Equation 11 becomes

$$S_n^s = \int_{x_{s-1}}^{x_s} S_n(x) dx = \int_{x_{s-1}}^{x_s} \frac{\bar{\mathcal{L}}_n(x) A(x)}{\bar{E}_n} dx = \frac{\mathcal{L}_n^s}{\bar{E}_n}, \tag{29}$$

Finally, inserting Equation 28 into Equation 26 and taking into account Equation 29, we recover Equation 13 of section 2.2, which is at the basis of the iterative scheme to move the duct resonances by means of area perturbations. Namely,

$$\frac{\delta f_n}{f_n} = \sum_{s=1}^{N_s} \frac{\mathcal{L}_n^s}{E_n}\frac{\delta A_s}{A_s} = \sum_{s=1}^{N_s} S_n^s \frac{\delta A_s}{A_s}. \tag{30}$$

If one makes a further simplification and considers a discretization of the duct into $s = 1 \cdots N_s$ cylinders of areas A_s and lengths l_s, such that the acoustic velocity u_s and pressure p_s inside them can be taken as constant, the kinetic and potential energies for the n-th mode in $\mathcal{L}_n^s$ (and consequently in S_n^s) of Equation 30, can be easily computed as

$$K_n^s = \frac{1}{2}\rho_0 A_s l_s u_s^2 \quad \text{and} \quad U_n^s = \frac{1}{2\rho_0 c_0^2} A_s l_s p_s^2 \ . \tag{31}$$

4. SIMULATIONS

As summarized in Section 2.2, an iterative process can be followed to move a set of resonant frequencies (formants) to target ones by modifying the area functions of the duct (e.g., a human vocal tract). As a starting point to demonstrate the method, we have used the Matlab implementation SensMap[1] described by Story in [6] for speech synthesis.

SensMap relies on a discretization of the vocal tract into N_s concatenated tubes of constant cross-section A_s and length l_s, so the simplifications for the potential and kinetic energies in Equation 31 apply to Equation 30. For the examples below, the velocities and acoustic pressures of the various N_s tubes have been computed by means of an acoustic model based on the Transfer Matrix Method (TMM), similar to that of Sondhi and Schroeter [16]. However, we have replaced the matrix coefficients to resemble the boundary conditions one would encounter in 3D vocal tract acoustic simulations [17], given that our future goal will precisely consist in linking the current approach with 3D models.

To illustrate the method, two simple examples will be addressed in which the formants of vowel [ɑ] are moved to new target frequency values. The area functions in [18] have been used to obtain the vocal tract shape of vowel [ɑ]. In all simulations, the speed of sound has been set to $c_0 = 350$ m/s and the air density to $\rho_0 = 1.14$ kg/m^2. The vocal tract transfer function

$$H(f) = \frac{P_{out}(f)}{Q_{in}(f)} \tag{32}$$

has been computed using the TMM to obtain the formant frequencies at each iteration. $H(f)$ is computed as the ratio between the output acoustic pressure captured 3 cm from the mouth exit with respect to the input volume velocity at the vocal tract entrance (see [17]).

In the first example, the frequency of the third formant (F3) is decreased while the other formant frequencies are kept constant. To do so, the formant frequencies of the original vocal tract

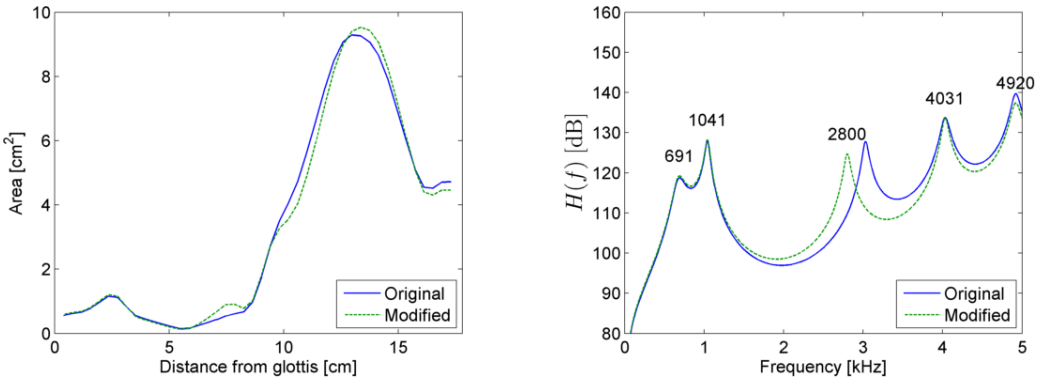
$$f = [691, 1041, 3030, 4032, 4920] \text{ Hz} \tag{33}$$

are moved to the target ones

$$f^{tg} = [691, 1041, \mathbf{2800}, 4032, 4920] \text{ Hz.} \tag{34}$$

[1] SensMap is a Matlab code freely distributed by the Speech Acoustics and Physiology Lab of the University of Arizona at http://sal.arizona.edu/node/27 (last access 20th February, 2019)

Figure 1 shows the area function and the vocal tract transfer function $H(f)$ for the original and modified vocal tracts. The achieved formant frequencies are shown in the top of the corresponding resonances in $H(f)$. As can be observed, the method succeeds in moving F3 from 3030 Hz to 2800 Hz. A total number of 117 iterations were needed to achieve an error of $\varepsilon \leq 1$ Hz (see Section 2.2 for its definition). Note, however, that although we intended to only move F3, the iterative procedure also produces some small shifts in the other formants, such as F4, with a small variation of the order of ε.



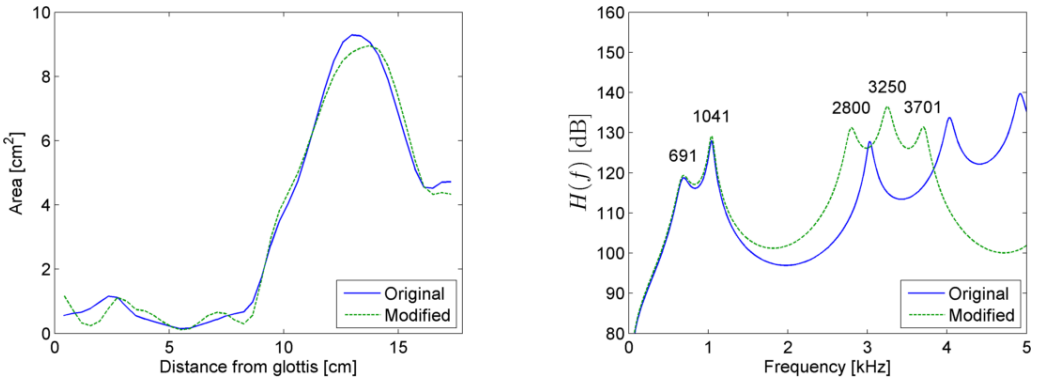
(a) (b)

Figure 1. (a) Vocal tract area function and (b) vocal tract transfer function $H(f)$ of vowel [ɑ] when the third formant is shifted down.

The second example deals with a more interesting application in what concerns the generation of computational expressive voice. In this case the third (F3), fourth (F4) and fifth (F5) formants are moved to form what is known as a cluster of formants. This can be typically found when changing the voice from a speaking style to a singing one. For this configuration, the target formant frequencies are selected to be

$$f^{tg} = [691\ 1041\ \mathbf{2800\ 3250\ 3700}]\ \text{Hz}. \tag{35}$$

The target values of F3, F4 and F5 in Equation 35 were taken from an example in [6], but preserving the first and second formants obtained for the current vocal tract configuration. For this example, a total of 305 iterations were ran to obtain a solution with an error smaller than 1 Hz, as in the preceding case. Figure 2 shows the corresponding area functions and vocal tract transfer functions, for the original and modified vocal tract. Observe that the first and second formants remain unaltered, whereas F3, F4 and F5 join together, almost reaching the desired target frequencies.



(a) (b)

Figure 2. (a) Vocal tract area function and (b) vocal tract transfer function $H(f)$ of vowel [ɑ] when the third, fourth and fifth formants are moved to form a cluster.

5. CONCLUSIONS

In this work, we have proved there is no need of resorting to the non-linear radiation pressure, to derive the area sensitivity functions for resonance tuning in ducts with varying cross-section. The propagation of sound waves inside such ducts is governed by the Webster equation. Applying a standard modal perturbation analysis to a finite element discretization of that equation allows one to recover the area sensitivity functions, without needing any non-linear phenomenon.

Two illustrative examples of the resonance iterative tuning procedure have been presented, in the context of speech synthesis. Those involved shifting a resonance (formant) to a particular target value and clustering a small set of resonances.

Ongoing work in the above framework of the discretized Webster equation contemplates the inclusion of damping in the duct walls, as well as deriving the sensitivity functions under length and wall admittance variations, using simple linear acoustics.

6. ACKNOWLEDGEMENTS

The authors would like to acknowledge the support provided by the Agencia Estatal de Investigación and FEDER, EU, through Project GENIOVOX TEC2016-81107-P. The first author would also like to thank l'Obra Social de la Caixa and the Universitat Ramon Llull for their support under grant 2018-URL-IR2nQ-031.

7. REFERENCES

1. Titze I, "*The Myoelastic Aerodynamic Theory of Phonation*", edited by the National Centre for Voice and Speech, Iowa City (2006).
2. Story BH, Titze IR, and Hoffman EA, The relationship of vocal tract shape to three voice qualities, J. Acoust. Soc. Am., 2001;109(4):1651-1667.
3. Chiba T and Kajiyama M, "*The vowel: Its nature and structure",* edited by Tokyo-Kaiseikan (1941).
4. Schroeder MR, Determination of the geometry of the human vocal tract by acoustic measurements, J. Acoust. Soc. Am., (1967);41(4):1002-1010.
5. Beissner K, The acoustic radiation force in lossless fluids in Eulerian and Lagrangian coordinates, J. Acoust. Soc. Am., (1998);103:2321–2332.
6. Story BH, Technique for "tuning" vocal tract area functions based on acoustic sensitivity functions, J. Acoust. Soc. Am., (2006);119(2):715-718.
7. Adachi S, Takemoto H, Kitamura T, Mokhtari P and Honda K, Vocal tract length perturbation and its application to male-female vocal tract shape conversion, J. Acoust. Soc. Am., (2007);121(6):3874-3885.
8. Fox R and Kapoor M, Rates of change of eigenvalues and eigenvectors., AIAA J (1968);6(12):2426-2429.
9. Adhikari S, Rates of change of eigenvalues and eigenvectors in damped dynamic system, AIAA J. (1999);37(11):1452-1458.
10. Arnela M, Guasch O, and Alías F, Effects of head geometry simplifications on acoustic radiation of vowel sounds based on time-domain finite-element simulations, J. Acoust. Soc. Am., (2013);134(4):2946-2954.
11. Arnela M, Blandin R, Dabbaghchian S, Guasch O, Alías F, Pelorson X,Van Hirtum A and Engwall O, Influence of lips on the production of vowels based on finite element simulations and experiments, J. Acoust. Soc. Am., (2016);139(5):2852-2859.

12. Arnela M, Dabbaghchian S, Blandin R, Guasch O, Engwall O, Van Hirtum A and Pelorson X, Influence of vocal tract geometry simplifications on the numerical simulation of vowel sounds, J. Acoust. Soc. Am., (2016);140 (3):1707-1718.
13. Dabbaghchian S. Arnela M, Engwall O and Guasch O, Reconstruction of Vocal Tract Geometries from Biomechanical Simulations, Int. J. Numer. Meth. Biomed. Eng. (2019), accepted.
14. Guasch O, Arnela M, Codina R and Espinoza H, A stabilized finite element method for the mixed wave equation in an ALE framework with application to diphthong production, Acta Acust. united Ac., (2016);102(1):94-106.
15. Dabbaghchian S. Arnela M, Engwall O and Guasch O, Synthesis of vowels and vowel-vowel utterances using a 3D biomechanical-acoustic model, (2019); submitted.
16. Sondhi MM Schroeter J, "A hybrid time-frequency domain articulatory speech synthesizer," IEEE Trans. Audio Speech Lang. Process., (1987); 35(7):955-967.
17. Arnela M, Dabbaghchian S, Guasch O, and Engwall O, MRI-based vocal tract representations for the three-dimensional finite element synthesis of diphthongs, Submitted (2019).
18. Story BH, "Comparison of magnetic resonance imaging-based vocal tract area functions obtained from the same speaker in 1994 and 2002", J. Acoust. Soc. Am., (2006); 123(1):327-335.