Acoustic simulation of alveolar trills using lumped self-oscillating models of the tongue tip

Benjamin Elie, Yves Laprie

Acoustic simulation of alveolar trills using lumped self-oscillating models of the tongue tip

Benjamin Elie[1, a)] and Yves Laprie[1]
LORIA, INRIA/CNRS/Université de Lorraine, Nancy, France

(Dated: March 16, 2017)

This paper investigates the possibility to model the self-sustained oscillations of the tongue tip during the production of alveolar trills. Using a realistic geometry of the vocal tract, derived from cineMRI data of a subject, it first compares the mechanical behavior of two lumped mechanical models of the tongue tip: a single-mass and a two-mass model. Simulations show that the two-mass model generates small phase shifts between the upstream and downstream parts of the tongue tip, suggesting a behavior similar to the single-mass model. The two-mass model is also less responsive than the single-mass model to the variations of the instantaneous acoustic pressure upstream of the lingual constriction due to the oscillations of the vocal folds. For this reason, the two-mass model should be favored when dealing with acoustic simulations of alveolar trills. The paper also proposes a solution to simulate an incomplete occlusion of the vocal tract during linguopalatal contacts by adding a lateral acoustic waveguide. This allows all kinds of situation experimentally observed during the production of alveolar trills to be simulated. The simulation framework is used to study the impact of a set of parameters on the characteristics of the produced alveolar trills. It shows that the production of trills is favored when the tongue tip position is slightly away of the hard palate, and when the glottis is fully adducted.

PACS numbers: 43.70.Bk, 43.72.Ja

I. INTRODUCTION

During the production of trill consonants (bilabial, alveolar, or uvular), aerodynamic forces applied to a particular articulator (lips, tongue tip, and uvula, respectively) are known to be responsible for the oscillations of the target articulator. The oscillations of the articulators lead to periodic amplitude modulations of the acoustic pressure waveform radiated at the lips. This periodic amplitude modulation is the main characteristics of trills.

This paper focuses on the mechanical and acoustic modeling of the production of alveolar trills, which are produced by the oscillation of the tongue tip in the alveolar region. This trill is the most common across languages[1,2]. However, this is a hard-to-produce sound that is mastered late in the acquisition process[3]. This suggests that it requires specific articulatory and aerodynamic conditions, which have been reported to be the reason of assimilation in neighboring phonetic segments[4]. Curiously, only a few studies have investigated the acoustic and aerodynamic characteristics of the production of alveolar trills. Acoustic properties have been deeply detailed via analysis of the speech signal[5] and the intraoral pressure[6]. Linguopalatal contacts, along with the speech signal characteristics have also been investigated[7].

Despite these works, there is still a lack of acoustic and/or mechanical models of the production of trills. To the best of our knowledge, the work by McGowan[8] is the sole attempt to study the production of trills using numeric simulations. He investigated the mechanisms involved in the self-oscillations of the tongue tip with a very simplified vocal tract. Simulations showed that the oscillations of the tongue tip were sustained thanks to the wall compliance of the vocal tract, hence a behavior similar to a one-mass system. The author has also briefly discussed the potential impact of classic two-mass lumped models, similar to those used for the vocal folds[9,10]. Indeed, they present the advantage of simulating phase shifts between the upstream and downstream parts of the vocal folds, as observed in real world experiments[9], and also to take a mobile flow separation point[10,11] into account. Besides, simulating self-oscillations of the vocal folds using a single-mass model is possible only if the oscillation frequency lies below the first resonance frequency of both the subglottal and supraglottal systems[12]. One might expect the same behavior for the tongue tip, and this paper investigates these questions, in Sec. IV, by comparing the mechanical behavior of a one-mass system, similar to the one introduced by Flanagan and Langraf[12], and a two-mass system with smooth contours and a mobile flow separation point[10], as detailed in Sec. III. In this work, the glottal source is connected to the classic transmission line paradigms[13], considering recent improvements[14] to deal with self-oscillating vocal folds with a membranous gap. Thus, in comparison with the work of McGowan[8], it supports realistic shapes of the vocal tract, and also the connection with self-oscillating models of the vocal folds with a glottal chink. This is an important contribution of the paper, since it enables the influence of the acoustic pressure variations upstream of the lingual constriction due to the vocal folds motion to be investigated with real anatomic data. Area functions have been derived from real-time cineMRI using a sparse reconstruc-

[a)] benjamin.elie@inria.fr

tion technique[15], detailed in Sec. II.

This paper also addresses the question of the different situations that may be observed in real subjects producing alveolar trills. Indeed, Recasens and Pallarès[7] have shown that the contact between the tongue tip and the hard palate is not systematic, depending on the speaker, but may occur, and results from Solé[6] suggest that there is no complete occlusion during the production of consonant trills, since the low frequency component of the mouth airflow is constantly strictly positive. Consequently, when dealing with acoustic simulations of the alveolar trills, one should guarantee that tongue tip oscillations with, or without contacts can be simulated, and also that, in some cases, linguopalatal contacts do not completely stop the main air path in the vocal tract. The paper presents in Sec. V a solution to deal with these situations by adding a lateral acoustic waveguide connected both at the upstream and the downstream parts of the tongue. This lateral connection is made possible thanks to the recent Extended Single-Matrix Formulation (ESMF) of the vocal tract[14]. The simulation framework including our self-oscillating tongue tip model is finally used to investigate the influence of various parameters on the production of alveolar trills. Results of the simulations are shown and discussed in Sec. VI.

II. ACOUSTIC AND ARTICULATORY CHARACTERISTICS OF ALVEOLAR TRILLS

A. Acoustic characteristics

Alveolar trills ([r]) are sounds that are characterized by periodic oscillations of the tongue tip. In world's languages, other trills exist according to the vibrating articulator, such as the uvula (uvular trills [ʀ]), or the lips (bilabial trills [ʙ] and labiodental trills [ʙ̪]), for instance. The vibrations of the articulators are not directly controlled by the speaker, namely they are not due to an active muscular control, but to the aerodynamic forces applied at their vicinity[8]. Studies about the production of alveolar trills include the mastering process[3] by children, occurrence in world's languages[1,2], coarticulatory effects[4,16,17], acoustic characteristics[5,6], and air-tissue interaction modeling[8].

From these studies, it appears that alveolar trills are difficult to produce[3], they impact the neighboring phonetic segments[4]. Aerodynamic measurements on speakers show that a high intraoral pressure is required to produce trills, and even higher for voiceless alveolar trills[6]. The same paper reports significant variability of the mean intraoral pressure among the two subjects of the study. Same subjects uttered trills with oscillation frequency in the range 28-33 Hz for voiceless trills and 26-29 Hz for voiced trills[6]. Their open phase is 1.94 times longer than the closed phase for voiceless trills, and 1.34 times longer for voiced trills.

Although the aforementioned studies have provided useful information about the mechanisms involved in the production of alveolar trills as well as their acoustic characteristics, several questions have not been fully addressed. Electro-Palatography data (EPG) have shown that contacts occur during the production of consonant trills[4,7,17]. However, contact patterns significantly vary across speakers, and even across repetitions of a single speaker. For a few speakers, contacts occur only at the edge of the tongue, and not at the tip. Hence, there is no evidence of systematic contact between the tongue and the hard palate, or the teeth, during the oscillation cycle. It is not unlikely that the tongue tip freely vibrates slightly away from the hard palate. Besides, in the case of contact between the tongue tip and the hard palate, is there a complete occlusion of the air path, or is there narrow channels around one or both sides of the tongue that permit the main airpath to not be stopped by the obstacle? These questions are particularly raised by the characteristics of the acoustic pressure waveforms of the recorded trills. In all studies[5,6], the acoustic waveform during the so-called closed phase is not null. This suggests that trills do not involve complete occlusion. It is also clear in Fig. 3 of the paper of Solé[6] that the mouth airflow is systematically strictly positive during the trill production. However, this is never discussed. Consequently, when dealing with numeric simulations, one should consider the three following situations: i) there is no contact between the tongue tip and the hard palate in the alveolar region, ii) contacts between the tongue tip and the hard palate occur, but not on all the edges of the tongue so that there is no total occlusion of the vocal tract, i.e. the intraoral airflow is not stopped, and iii) there are contacts everywhere on the edges of the tongue so that the occlusion is complete. Another major contribution of the paper is to present an acoustic model that is able to reproduce all of these situations.

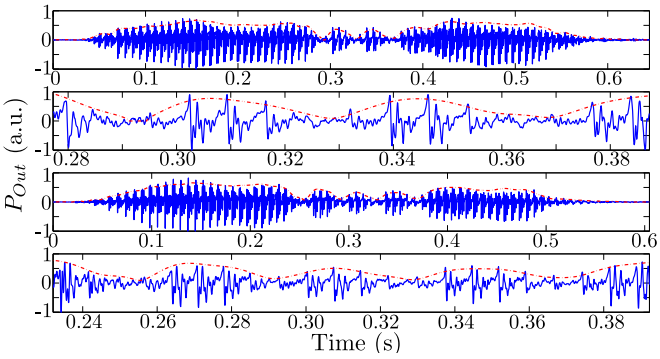


Figure 1. Audio speech signal of two utterances of the pseudoword /ara/, with a zoomed-in version of the alveolar trill [r]. The time envelope is plotted as a dashed line

Fig. 1 shows two examples of acoustic speech signal of alveolar trills uttered in the pseudoword /ara/. The time envelope of the acoustic pressure waveform shows periodic cycles during the production of the alveolar trill [r]. It oscillates between a maximum, at a level slightly less

than the adjacent vowels, and a minimum, which is not null. The non-zero value of the acoustic waveform minimum suggests that there is a permanent acoustic flow at the lips, going either through the lingual constriction, or along the sides of the tongue, or both ways.

B. Articulatory characteristics

Observing tongue tip movements during alveolar trills is very difficult. Indeed, since the oscillation has small amplitude and is relatively fast, it requires high spatiotemporal resolutions. Thus, high-speed Ultrasound imaging and Electromagnetic articulography (EMA)[18] have been used to observe the time evolution of the tongue tip position during alveolar trills. These studies provide data from alveolar trill realizations of German[18], Czech[19], Spanish, and Russian speakers[20]. All of them show oscillation amplitudes of a few millimeters. EMA measurements also show that the tongue exhibit a flatter shape in the coronal plane during trills in comparison with fricatives[19].

For our simulations, the whole area function is needed. For that purpose, the vocal tract geometry has been observed via real-time cineMRI techniques. It consists in a reconstruction of the midsagittal view of the vocal tract at a reasonably high frame rate, using sparse reconstruction methods[15]. The frame rate is 29 frames per second, with a spatial resolution of 1×1 mm^2. MRI experiments were performed on a 3T Signa HDxt MR system (General Electric Healthcare, Milwaukee, WI). Speech MRI data were obtained from a healthy volunteer with written informed consent and approval of local ethics committee. He is a 35 years old male speaker. The data were collected with a 16-channel neurovascular coil array. The protocol consisted in a sagittal slice through the middle of the vocal track acquired with a custom modified Spoiled Fast Gradient Echo (Fast SPGR, TR 3.5ms, TE 1.1ms, line BW 83.33 kHz, flip angle 30 degrees, half echo in frequency direction, matrix 256×256, 512 temporal frames). For each temporal frame the modification consisted to acquire only a randomized fixed size sample (10 lines) of the 256 phase lines of the k-space. The alveolar trill is pronounced in the intervocalic context /ara/.

We considered that acquiring MR images at 29 frames per second, namely similar to the trill frequency, is sufficient to get a good idea of the global position of the tongue tip during oscillations along with the rest of the vocal tract since we acquire several images per trill. For the study, the speaker was asked to slightly lengthen the trill duration, so that between 5 and 10 images per trill realization are available.

Fig. 2 shows an example of the midsagittal view of the vocal tract during the production of an alveolar trill in the sequence /ara/. Contours of the vocal tract in each reconstructed MR image of the sequence have been semi-automatically extracted by hand, as detailed in our previous paper[15], and area functions were obtained by, first, dividing the vocal tract shape in tubelets perpendicular to the centerline, and then applying α β transformations to recover the area[21]. This transformation is a power function that gives the cross-sectional area $a(x)$ as a function of the midsagittal distance $d(x)$:

$$a(x) = \alpha(x) d(x)^{\beta(x)}, \tag{1}$$

where α and β are ad hoc coefficients. Their values depend on the region of the vocal tract, and x is the distance to the glottis along the centerline. Several studies have provided different values for α and β. We kept those given by Soquet *et al.*[21] for the male speaker. The determination of the centerline plays a critical role in synthesis since it influences the length of the vocal tract. A specific algorithm was designed purposely[22].

As previously explained, the real-time cineMRI reconstruction enables several area functions to be recovered for a single trill realization. Since the main aim of the paper is to investigate the possibility to simulate self-oscillations of the tongue tip in alveolar trills with lumped mass-spring models integrated into realistic shapes of the vocal tract, using area functions of alveolar trills on different contexts is not in the scope of the paper. Consequently, for the sake of brevity, we chose to use only one area function for the rest of the paper: it corresponds to the median area function extracted from the acquired sequence of images corresponding to the production of the examined trill. The choice of a median area function is mainly for the purpose of compensating the potential errors due to the manual delineation, as well as the variations in the alveolar constriction region due to the tongue tip movement during the trill production. The resulting area function is shown in Fig. 2. In the simulations, this area functions corresponds to that of the vocal tract when the tongue is at its rest position. Oscillations of the tongue tip will then modify the cross-sectional area of the lingual constriction, namely the area of the tubelet having the smallest area (at 15.9 cm in Fig. 2 (c)). Note the narrowing at the middle of the vocal tract, which is due to the presence of the velum in the area, as shown in the MR image in Fig. 2.

III. ACOUSTIC MODEL

The simulation framework used to compute the acoustic propagation inside the vocal tract is derived from the *transmission line circuit analog model* (TLCA) approach[13]. It considers plane waves propagating along a spatially sampled vocal tract, modeled as a set of connected acoustic tubes, or *tubelets*. Unlike the other widely used approach, the *reflection type line analog* (RTLA) model[23,24], it easily deals with time-varying lengths of the vocal tract, and also with uneven spatial sampling of the vocal tract. For further information, the reader may find a detailed review of existing techniques for speech synthesis in Ref.[25]. The framework that is

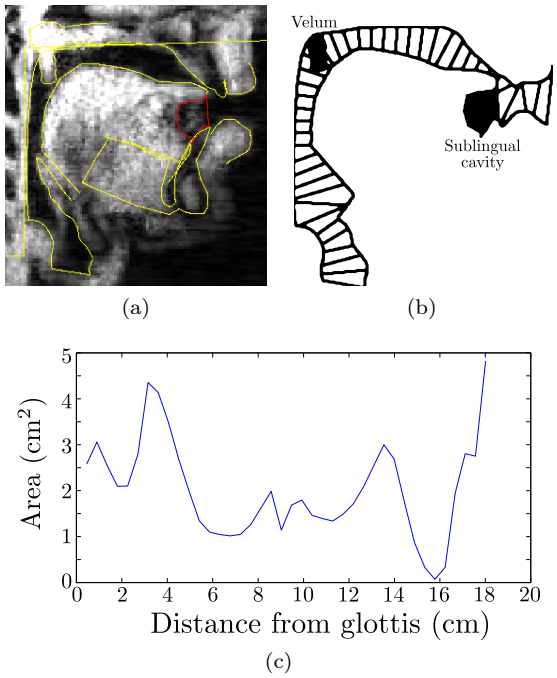


Figure 2. Top: (a) MR image of a midsagittal view of the vocal tract during the production of an alveolar trill in the sequence /ara/. Contours of the articulators are shown in the MR image, as well as (b) the spatial sampling of the vocal tract into acoustic tubelets. (c) resulting median area function obtained from real-time cineMRI data.

used in this paper, called ESMF[14], considers recent improvements of TLCA-based techniques, such as the possibility to connect self-oscillating models of the vocal folds with a glottal chink. In this paper, the acoustic model is extended to include the potential oscillations of the tongue tip.

A. Acoustic propagation

The vocal tract is seen as a waveguide network, where each waveguide represents a side cavity[14,26]. In our study, side cavities include the sublingual cavity (*cf.* Fig. 2) and the lateral channels around the sides of the tongue. It has been shown, following the recent ESMF paradigm[14], that the wave propagation inside such networks, connected to a self-oscillating model of the vocal folds, is driven by the following system of equations

$$\mathbf{f} = \mathbf{Z}\mathbf{u}_Z + \mathbf{Q}\mathbf{u}_Q, \tag{2}$$

where $\mathbf{f} \in \mathbb{R}^{(N+1)}$ is a vector containing pressure forces, $\mathbf{Z} \in \mathbb{R}^{(N+1)\times(N+1)}$ is a tridiagonal matrix containing impedance and loss terms associated to each tubelet, where $\mathbf{Q}$ is a square matrix the same size as $\mathbf{Z}$ having only one non-zero element, that is $Q_{(1,1)} = R_b$, $\mathbf{u} \in \mathbb{R}^{N+1}$ is the vector containing the volume velocities inside each tubelet, and $\mathbf{u}_Q \in \mathbb{R}^{(N+1)} = [U_1^2, U_2^2, \ldots, U_N^2]^T$ is the vector containing the square power of the volume velocities. The term R_b is the Bernoulli resistance, as defined in Ref.[14].

The acoustic propagation model supports the connection of self-oscillating models of the vocal folds with a glottal chink[14]. In such case, the membranous portion of the glottis is composed of two distinct parts along the length of the vocal folds: an adducted part, modeled with a classic two-mass self-oscillating system[10], and an abducted portion, i.e. the so-called *glottal chink*, modeled as an acoustic branch. This allows incomplete closure of the glottis during oscillation cycles of the vocal folds to be modeled, and has been found to be important for simulating breathy voice[27], and voiced fricatives[28]. In this paper, the chink length opening, denoted by l_{ch}, is defined as the percentage of the vocal folds length that corresponds to the glottal chink. Consequently, $l_{ch} = 0\%$ when there is no glottal chink, and $l_{ch} = 100\%$ when the glottis is fully abducted.

B. Self-oscillation models of the tongue tip

In order to reproduce self-oscillations of the tongue tip, our models consider the lower wall of the tubelet corresponding to the lingual constriction to be lumped mass-spring systems, either a single-mass, as introduced by Flanagan and Landgraf[12], or the two-mass model, derived from the classic two-mass mechanical system for the vocal folds by Ishizaka and Flanagan[9]. In this paper, self-oscillations are used to update the cross-sectional area of the tubelet that corresponds to the lingual constriction, namely the tubelet that presents the minimal cross-section area. Consequently, only the extremity of the tongue tip is assumed to oscillate, without modifying the cross-section areas of the neighboring tubelets. In the following presentation of the models, the upstream and downstream sections, as in Figs. 3 and 4, represent the tubelets located just downstream and upstream of the tubelet with the minimal cross-section area.

1. One-mass model

When dealing with the vocal folds, one-mass models have provided satisfying results as long as the oscillation frequency was below the first formant frequency of the subglottal and supraglottal systems[29,30], and the subglottal pressure was varied[31]. However, only a transverse oscillation mode can be simulated. For tongue tip oscillations during alveolar trills, the observed range[6] of the oscillation frequency lies between 25 and 35 Hz, which is much below the formant frequencies of the vocal tract. Also there is no evidence, so far, of the existence of important contributions of tongue tip's oscillation modes higher than the transverse mode. As a consequence, the single-mass model may be sufficient to accurately simulate alveolar trills[8]. One major drawback would be the

fact that single-mass models have been shown to lead to unrealistic amount of acoustic interactions between the vocal folds and the vocal tract[31], which could also be true between the tongue tip and the rest of the vocal system. This will be discussed in the Sec. IV.

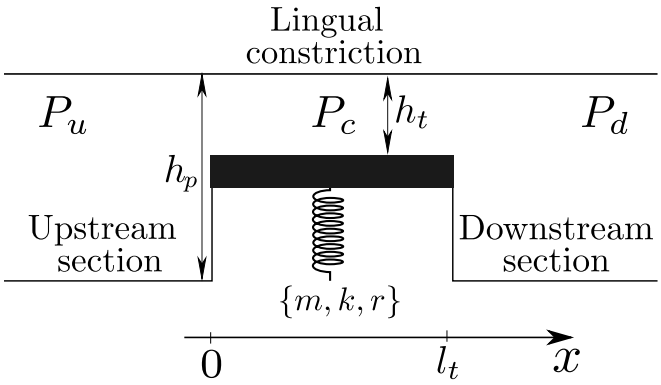


Figure 3. Single-mass model for the tongue tip. P_d, P_u, and P_c denote the pressure downstream, upstream and inside the lingual constriction, respectively. The mass, stiffness and the damping of the tongue are denoted by the triplet $\{m, k, r\}$. The height of the lingual constriction is denoted by h_t.

The system is represented in Fig. 3. It is a classic forced oscillating system with a single degree of freedom. The equation of the lingual constriction height h_t writes

$$m\ddot{h}_t(t) + r\dot{h}_t(t) + k\Delta h_t(t) = F(t), \tag{3}$$

where m, r, k are respectively the mass, the damping, and the stiffness of the tongue tip, and F is the pressure forces that are applied to the system. The elongation is $\Delta h_t = h_t - h_0$, where h_0 denotes the rest position of the tongue.

Pressure forces are defined as

$$F = \int_0^{l_t} P(x)dx, \tag{4}$$

where l_t is the length of the lingual constriction. The pressure distribution along the lingual constriction is given by

$$P(t) = P_u(t) + Be(t) + Po(t) + In(t) \tag{5}$$

where $Be(t)$, $Po(t)$, and $In(t)$ are respectively the steady term of the Bernoulli equation, the Poiseuille corrective term and the unsteady term of the Bernoulli equation. They are defined as:

$$\begin{aligned} Be(t) &= -\frac{\rho_a U_t^2(t)}{2w_t^2}\left[\frac{1}{h_t^2(t)} - \frac{1}{h_p^2(t)}\right], \\ Po(t) &= -\frac{12\mu_a l_t U_t(t)}{w_t h_t^3(t)}, \\ In(t) &= -\frac{\rho_a l_t}{w_t}\frac{\partial}{\partial t}\left[\frac{U_t(t)}{h_t(t)}\right], \end{aligned} \tag{6}$$

where w_t and l_t are respectively the width and the length of the lingual constriction, ρ_a and μ_a are respectively the mass density and the shear viscosity of the air. The height h_p is the height of the section just downstream of the lingual constriction.

2. Two-mass model

For simulating the self-oscillations of the vocal folds, the choice of a two-mass system instead of a single-mass one is motivated by the possibility to reproduce higher modes of oscillation, and especially phase shifts between the upstream and the downstream parts of the vocal folds. Indeed, at least two modes of oscillation have been extensively observed and shown to be important to consider in the simulations[29]. For tongue tip oscillations, this has been previously discussed by McGowan[8], but there is, so far, no evidence of such behavior of the tongue tip during trill productions.

The two-mass model used for the tongue tip is then similar to the one used for vocal folds. It includes smooth contours, a mobile flow separation point[10,11], and the computed pressure forces consider viscous losses and unsteady flow effects[32,33].

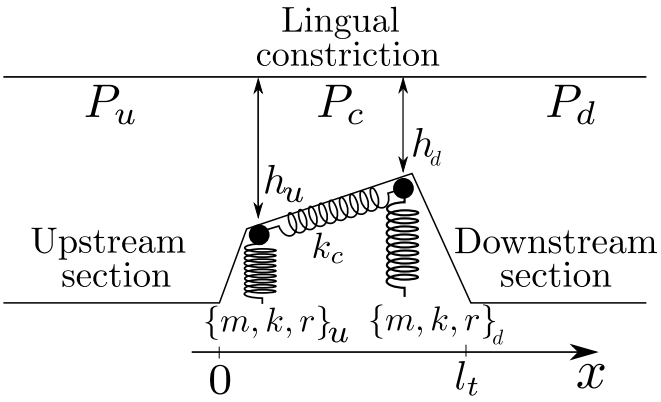


Figure 4. Two-mass model for the tongue tip. P_d, P_u, and P_c denote the pressure downstream, upstream and inside the lingual constriction, respectively. The mass, stiffness and the damping of the tongue are denoted by the triplet $\{m, k, r\}$. The height of the lingual constriction is denoted by h.

The geometry of the tongue tip is then represented in Fig. 4. In this paper, indices u and d denote the upstream and downstream parts respectively. The equation of the tongue motion is

$$\begin{cases} m_u\ddot{h}_u(t) + r_u\dot{h}_u(t) + (k_u + k_c)\Delta h_u(t) - k_c\Delta h_d(t) = F_u(t) \\ m_d\ddot{h}_d(t) + r_d\dot{h}_d(t) + (k_d + k_c)\Delta h_d(t) - k_c\Delta h_u(t) = F_d(t) \end{cases}, \tag{7}$$

where the expressions of the pressure forces and the pressure distribution are similar to Eq. (4) and Eq. (5). However, the potential phase shift between both masses modifies the expressions of $Be(x,t)$, $Po(x,t)$, and $In(x,t)$

in Eq. (6), which depend on the position along the x-coordinate:

$$\begin{aligned} Be(x,t) &= -\frac{\rho U_t^2(t)}{2w_t^2}\left[\frac{1}{h_t^2(x,t)} - \frac{1}{h_t^2(x_0,t)}\right], \\ Po(x,t) &= -\frac{12\mu U_t(t)}{w_t}\int_{x_0}^{x}\frac{dx}{h_t^3(x,t)}, \\ In(x,t) &= -\frac{\rho}{w_t}\frac{\partial}{\partial t}\left[U_t(t)\int_{x_0}^{x}\frac{dx}{h_t(x,t)}\right]. \end{aligned} \tag{8}$$

Eq. (8) is valid for x located upstream of the separation point. Indeed, in this paper, similarly to the vocal folds model, the two-mass model of the tongue tip includes a mobile separation point x_s, where x_s is such that $h_s = h(x_s) = 1.2h_u$ if $1.2h_u < h_d$, and $x_s = x_d$ otherwise. The value of 1.2 is a ad hoc criterion[10]. The pressure $P(x,t) = P_d(t)$ if $x > x_s$.

The collision model use artificial stiffening of the tongue tip, namely the stiffness parameters k_u and k_d are multiplied by a factor 4 if h_u, or h_d, are equal or less than 0, respectively, and the damping is multiplied by a factor 2.2. These factors are also ad hoc criteria, chosen following the original two-mass model by Ishizaka and Flanagan[9].

IV. COMPARISON BETWEEN LUMPED MECHANICAL MODELS

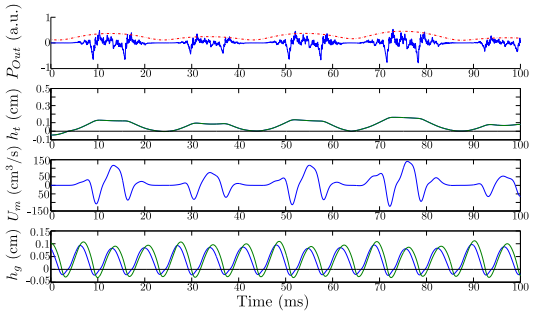
A. Simulation of alveolar trills

Using the area function derived from cineMRI data of a real speaker, sustained alveolar trills are simulated, first with the single-mass model, and then with the two-mass model. The parameters chosen for the two-mass system are summarized in Tab. I. The width of the lingual constriction w_t and the rest height h_0 are derived from the MRI data. Since lumped mass-spring systems models consider the shape of the constriction as rectangular, w_t is estimated such that the constriction area $a_t = h_0 w_t = \alpha_t h_0^{\beta}$ matches the cross-sectional area given by the power transformation of Eq. (1), hence

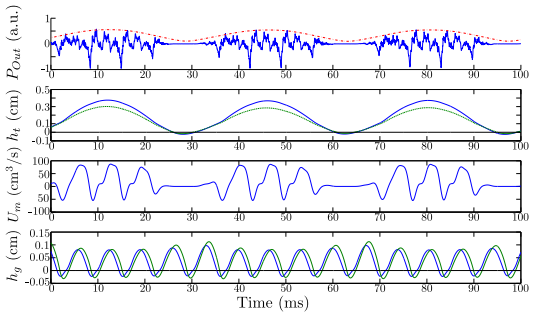
$$w_t = \alpha_t h_0^{\beta_t - 1}, \tag{9}$$

where α_t and β_t are the transformation coefficients from Soquet *et al.*[21] in the alveolar region, namely $\alpha_t = 1.92$ and $\beta_t = 1.2$ when w_t and h_0 are expressed in cm. The stiffness values have been chosen according to the mass values so that the resonance frequency of the uncoupled mechanical system is 25 Hz, which is in the range of typical values observed in real speakers[6]. Assuming the mass density of the tongue ρ_t is slightly more than water[34], i.e. $\rho_t \simeq 1.04\times10^6$ g.m^{-3}, the distance between the base and the extremity of the tongue tip h_{tt}, measured in MR images, is around 1.1 cm, and the tongue tip is a parallelepiped, each mass is then

$$m = \frac{1}{2}\rho_t \times l_t \times h_{tt} \times w_t \simeq 0.26 \text{ g}.$$



(a)



(b)

Figure 5. Simulation of alveolar trills with linguopalatal contacts by using (a) a single-mass model, and (b) a two-mass model. For each figure, from top to bottom are plotted the acoustic pressure radiated at the lips (P_{Out}), as well as the time envelope (dot-dashed line), the time evolution of the lingual constriction height h_t at the upstream (solid line) and downstream (dashed line) sections, the low frequency filtered airflow at the mouth U_m. The bottom figure plots the displacement h_g of the upstream (solid line) and downstream (dashed line) parts of the vocal folds.

Results of the simulations are displayed in Fig. 5. It shows the oscillations of the lingual constriction height around its rest position. In this example, periodic contacts between the tongue tip and the hard palate occur. In this configuration, namely when there is no lateral channel, linguopalatal contacts completely stop the air flow inside the vocal tract, hence the zero values in the airflow waveform U_m, and in the acoustic pressure waveform P_{Out}. In any case, the time envelope of the acoustic pressure is then a periodic function, which is at its minimum when the tongue tip is close to the hard palate (or even is in contact with), and at its maximum when the tongue tip is far from the hard palate. Simulations also reveal that the tongue tip oscillations may disturb the motion of the vocal folds. Indeed, the amplitude of the oscillation of the vocal folds increases when the tongue tip is close to, or in contact with, the hard palate. More interestingly, the downstream part of the vocal folds are more impacted, since it exhibits much larger oscillation

amplitude than the upstream part in the closing phase of the lingual constriction. Both the single-mass and the two-mass models simulate this phenomenon.

Simulations also reveal significant differences in their behavior. Firstly, the tongue oscillation frequency f_t of the single-mass model appears to be much greater than the one of the two-mass model (45 Hz versus 29 Hz). Secondly, the tongue tip displacement waveform computed with the single-mass model is much less smooth than the waveform given by the two-mass model. Indeed, high-order harmonics seem to significantly contribute to the displacement waveform. The main acoustic consequence is a more erratic time envelope of the acoustic waveform. One possible explanation of the presence of these high-order harmonics could be an excessively high coupling between the tongue tip and the vocal folds. Indeed, since the oscillations of the tongue tip are driven by the acoustic pressure upstream of the lingual constriction, it is very likely that the oscillations of the vocal folds disturb the oscillations of the tongue tip. This hypothesis is investigated in the next section.

Table I. Parameters of the two-mass model of the tongue tip

Parameter	Unit	Value
Subglottal pressure P_{sub}	Pa	1000
Rest position of the tongue tip h_0	mm	0.7
Length of the lingual constriction l_t	mm	4.4
Width of the lingual constriction w_t	mm	10.5
Mass $m = m_1 = m_2$	g	0.26
Stiffness $k = k_1 = k_2$	N/m	228
Nominal damping coefficient r_i	kg.rad.s^{-1}	$0.02\sqrt{k.m/2}$
Coupling spring k_c	N/m	$k/2$

B. Coupling between the glottis and the tongue tip oscillations

1. *Contributions of the vocal folds oscillations in the tongue tip displacement*

Fig. 6 shows the spectra of the waveform corresponding to the lingual constriction opening, computed from the simulations presented in the previous section, namely using the single-mass and the two-mass models. Results evidence the hypothesis of a stronger acoustic coupling between the tongue tip and the vocal folds in the case of the single-mass model. Indeed, in that case, peaks at the oscillation frequency of the glottal opening ($f_0 = 137$ Hz) and the following harmonics are clearly visible in the spectra of the lingual constriction opening. The spectrum for the single-mass also exhibits heterodyne peaks: considering the tongue oscillation frequency $f_t = 45$ Hz and

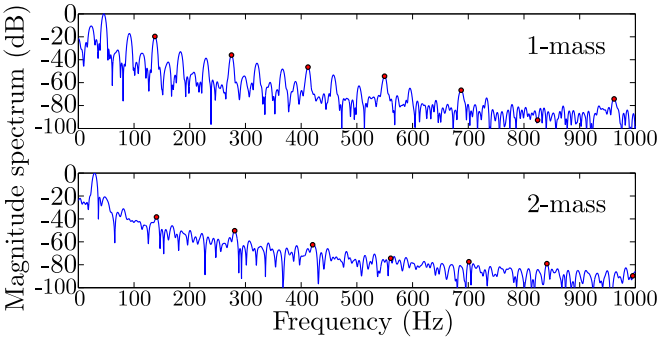


Figure 6. Magnitude spectrum of the lingual constriction opening. Top is the single-mass model, bottom is the two-mass model. The oscillation frequency of the glottal opening and its harmonics are denoted by circles.

the fundamental frequency of the vocal folds $f_0 = 137$ Hz, the heterodyne frequency, at the frequency $f_0 - f_t = 91$ Hz, and the corresponding harmonic peaks are prominent. The spectrum of the lingual opening obtained with the two-mass model does exhibit additional contributions of the oscillation frequencies of the vocal folds, but peaks are much smaller than in the case of the single-mass model.

2. *Impact of the vocal folds on the characteristics of the tongue tip oscillations*

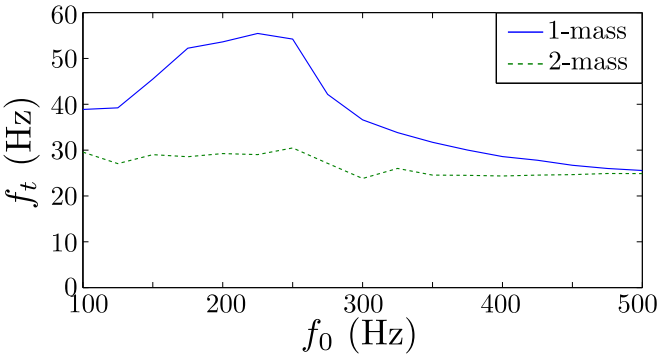


Figure 7. Trill frequency (denoted f_t) as a function of the fundamental frequency of the vocal folds (denoted f_0) of the single-mass (solid line) and the two-mass model (dashed line).

To investigate the impact of the acoustic coupling between the vocal folds and the tongue tip, simulations are carried out at various oscillation frequencies of the vocal folds, ranging from 100 to 500 Hz. Values of the mechanical parameters of the tongue tip models are still the same as in the previous sections, and they are still such that the resonance frequency of the mechanical system is 25 Hz. Fig. 7 shows the variation of the oscillation frequency of the tongue tip as a function of the fundamental frequency of the glottal opening. The trill frequency f_t corresponds

to the inverse of the mean value of the individual oscillating cycle periods of the simulated trills. Again, the strong coupling between the vocal folds and the tongue tip is highlighted by the simulations. Indeed, whereas the trill frequency f_t, obtained with the two-mass model barely varies with the oscillation frequency of the vocal folds f_0 (f_t remains in the range 25-30 Hz), f_t obtained with the single-mass model significantly varies with f_0, between 25 and 55 Hz. Note that this effect vanishes for high values of f_0.

When dealing with the simulation of tongue tip oscillations during the production of voiced alveolar trills, it seems that the two-mass model should be considered instead of the single-mass. Indeed, it guarantees better stability in regards with the fundamental frequency of the glottal source. The acoustic characteristics of the simulated speech signal are also less erratic, which is more in agreements with the observed speech signals of real speakers.

V. SIMULATIONS OF THE INCOMPLETE CLOSURE OF THE VOCAL TRACT AT THE TONGUE PALATE CONTACT INSTANTS

As discussed in Sec. II, there is no evidence of complete closure of the vocal tract during contacts of the tongue tip with the hard palate. Indeed, the air flow at the mouth and the acoustic pressure waveform are not null. In order to simulate the incomplete closure of the vocal tract during linguopalatal contacts, we suggest to add a lateral acoustic branch to simulate the air path going along the side of the tongue. The ESMF paradigm that is used in this paper enables such situations[14] to be taken into account in the acoustic simulations. Due to the lack of data (cineMRI films provide only midsagittal views of the vocal tract), the area function of the lateral branch is estimated according to these assumptions: i) connections at the upstream and at the downstream parts are located respectively at 14.0 cm and 16.2 cm from the glottis, ii) the cross-sectional area of the lateral branch is taken at a certain ratio r_l of the vocal tract area function located between the connections, iii) the length of the lateral branch is the same as the length of the main air path. The presented simulation uses the two-mass model of the tongue tip with the same mechanical parameters as in Tab. I, integrated into a vocal tract modeled with a lateral branch that is connected to the main path with an area ratio r_l of 50%.

The simulation with a ratio $r_l = 50\%$, displayed in Fig. 8, shows that this vocal tract configuration enables linguopalatal contact without complete closure of the vocal tract to be simulated. When the tongue tip and the hard palate are in contact ($h_t \leq 0$), the air flow U_m is not null, nor is the acoustic signal pressure. This can be compared with Fig. 5 (b), which shows the results of a simulation with similar parameters, but without the lateral waveguide. The acoustic waveform obtained with

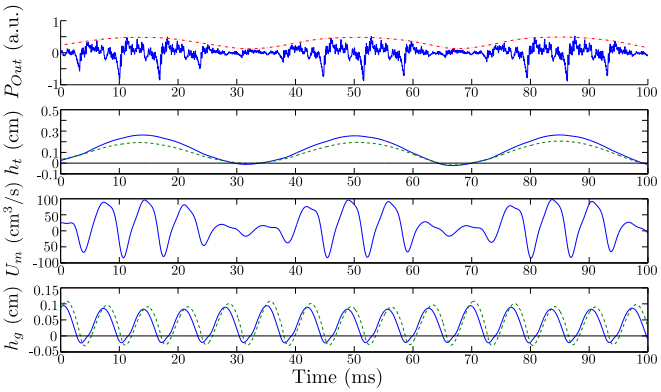


Figure 8. Results of the simulation of a sustained alveolar trill with a lateral configuration, using the two-mass model, and $r_l = 50\%$. From top to bottom: acoustic pressure radiated at the lips P_{Out} (solid line) and time envelope (dash-dotted line), lingual constriction height h_t at the upstream (solid line) and at the downstream (dashed line) parts, and air flow U_m at the mouth. The bottom plot represents the displacement h_g of the vocal folds at the upstream (solid line) and at the downstream (dashed line) sections.

the incomplete closure configuration is more in agreement with the acoustic waveform observed in real trills (see Fig. 1, for instance).

VI. CONDITIONS OF PRODUCTION OF ALVEOLAR TRILLS

This section investigates the influence of various parameters of the presented model on the simulation of alveolar trills. Studied parameters include the subglottal pressure, the mass of the tongue tip, the height of the lingual constriction at rest, the glottal opening, and the lateralization of the air flow.

The computed trill characteristics are the trill frequency f_t, the trill amplitude, denoted $\hat{h}_t$, and the trill contact ratio, denoted C_r. For each simulation of sustained trills, a group of 10 successive periods are considered. The trill amplitude $\hat{h}_t$ is then the mean value of the difference between the maximal and the next minimal lingual constriction height in the group of the 10 successive oscillating cycles. The trill contact ratio C_r is the percentage of oscillating cycles where there is a linguopalatal contact among 10 successive periods.

A. Effect of the mass and the subglottal pressure

For these simulations, the glottal chink is entirely closed, i.e. $l_{ch} = 0\%$, and there is no lateralization, i.e. $r_l = 0\%$. The other parameters, namely the rest position h_0, the length and the width of the lingual constriction are set to the values shown in Tab. I. The masses m_1 and m_2 vary from 10 mg to 1 g, and the subglottal pressure

from 500 to 2000 Pa.

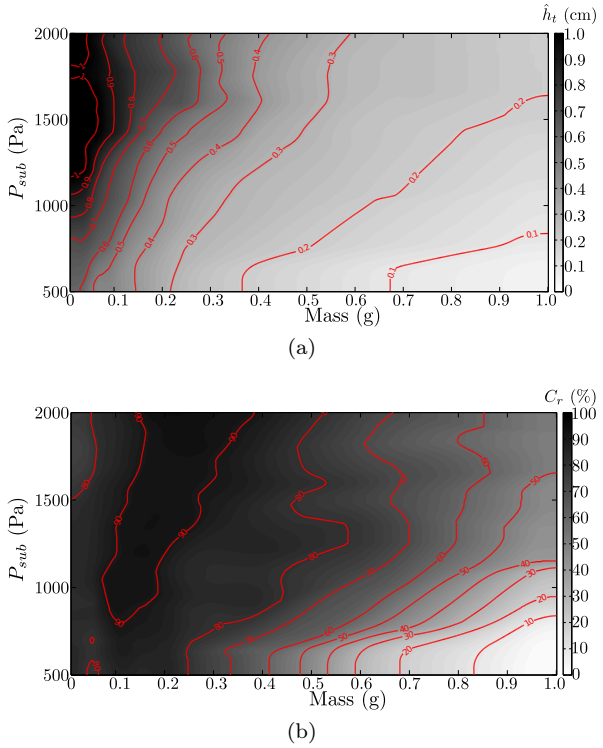


Figure 9. (a) Trill amplitude and (b) trill contact ratio, as a function of the mass $m_1 = m_2$ and the subglottal pressure P_{sub}. Contour lines are displayed.

Fig. 9 shows the trill contact ratio and the trill amplitude as a function of the mass of the tongue tip and the subglottal pressure. Not surprisingly, the trill amplitude is the largest for small values of the tongue mass and for large values of the subglottal pressure P_{sub}. For mass values smaller than 0.1 g, the amplitude of the tongue tip oscillation may be larger than 1 cm, which seems unrealistically high, since it is out of the range of reported data[18,19,35]. This leads to area variations of approximately 1.4 cm^2, according to Eq. (1), and is much larger than data provided by McGowan[8], who reported maximal constriction area between 0.27 and 0.67 cm^2. There areas correspond to maximal constriction heights of 0.44 cm and 0.21 cm, respectively. In the range of subglottal pressure values usually observed in natural speech[36], i.e. under 1000 Pa, the trill amplitude values between 0.21 cm and 0.44 cm are reached for mass values between 0.2 g and 0.4 g. In this range, linguopalatal contacts occur almost systematically ($C_r > 80\%$). Note that the contact ratio C_r follows the same evolution than the trill amplitude. For the rest of the paper, the mass of the tongue will be set to 0.26 g, as in Sec. IV, since it both roughly corresponds to anatomical data, and lies in the range of mass values that simulates realistic oscillations of the tongue tip.

B. Effect of the tongue initial position

For these simulations, the glottal chink is entirely closed, i.e. $l_{ch} = 0\%$, and there is no lateralization, i.e. $r_l = 0\%$. The length and the width of the lingual constriction are set to the values shown in Tab. I. The masses m_1 and m_2 are set to 0.26 g. The rest position varies from 0 to 0.4 cm and the subglottal pressure from 500 to 2000 Pa.

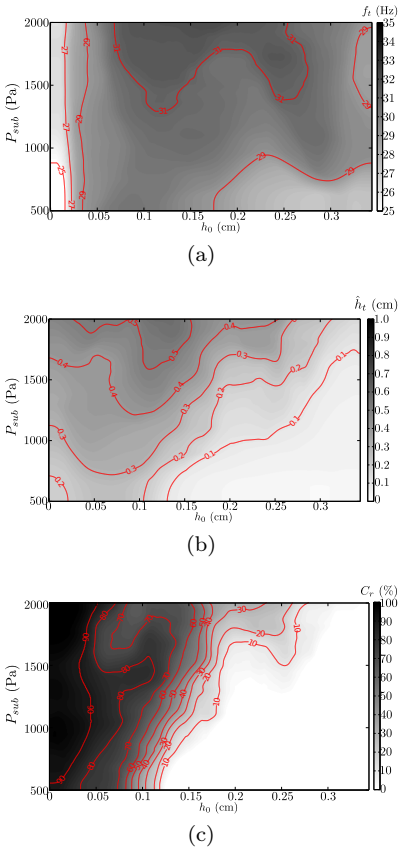


Figure 10. (a) Trill frequency f_t, (b) trill amplitude $\hat{h}_t$, and (c) trill contact ratio C_r, as a function of the subglottal pressure P_{sub} and the rest position of the tongue h_0. Contour lines are displayed.

The tongue initial position h_0 is set to different values to study its impact on the self-oscillating motion of the tongue tip. Fig. 10 shows the trill frequency f_t, the trill amplitude $\hat{h}_t$, and the trill contact ratio C_r as a function of the subglottal pressure P_{sub} and the initial tongue position h_0. It clearly shows that the inital tongue position has a significant influence on the production of alveolar trills. For instance, for a given P_{sub}, the trill amplitude as a function of h_0 exhibits a local maximum, i.e. for h_0 between 0.5 and 1 mm. This implies that the tongue should not be in contact with the hard palate at the beginning of the trill. Results also show that when h_0 is larger than approximately 1.25 mm, linguopalatal contacts occur only when P_{sub} is larger than a certain threshold.

This threshold value increases as h_0 increases. It is also worth noting that larger rest positions of the tongue tend to increase the trill frequency. The trill frequency is the most sensitive to the rest position when h_0 is smaller than the local maximum, i.e. when $h_0 < 0.5$ mm. This confirms the fact that alveolar trills is difficult to realize since it requires a very fine adjustment of the tongue tip position.

C. Effect of the incomplete closure of the vocal tract

For these simulations, the glottal chink is entirely closed, i.e. $l_{ch} = 0\%$, the rest position, the length and the width of the lingual constriction are set to the values shown in Tab. I. The masses m_1 and m_2 are set to 0.26 g. The lateral area ratio r_l varies from 0 to 0.5, and the subglottal pressure from 500 to 2000 Pa.

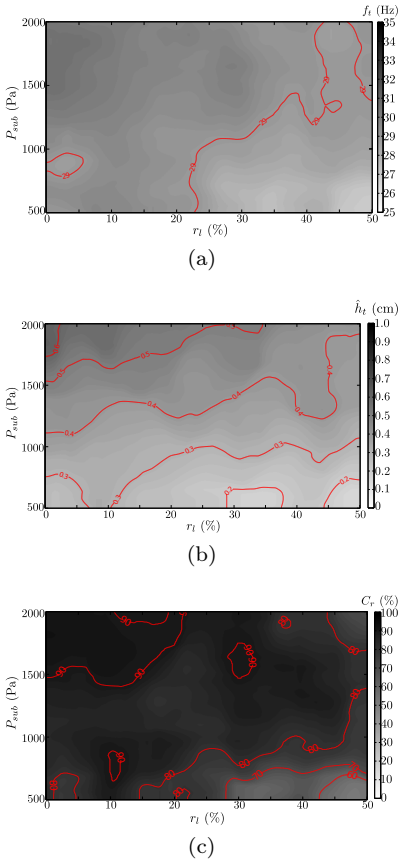


Figure 11. (a) Trill frequency f_t, (b) trill amplitude $\hat{h}_t$, and (c) trill contact ratio C_r, as a function of the subglottal pressure P_{sub} and the lateralization ratio r_l. Contour lines are displayed.

The incomplete closure of the vocal tract is studied via the lateralization ratio r_l. Fig. 11 shows that this has no significant effect on the trill characteristics: the trill frequency is barely lowered by 1 Hz when there is an incomplete closure of the vocal tract, and, for a given P_{sub}, the trill amplitude and the trill contact ratio exhibit a slight tendency to decrease as the incomplete closure increases (smaller r_l). However, although the incomplete closure of the vocal tract during linguopalatal contacts has a very small impact on the mechanical behavior of the tongue tip, it has been shown in Sec. V that the acoustic impact is not negligible. Consequently, when dealing with acoustic synthesis of alveolar trills using physical models, such as this one, one should consider the lateralization of the airflow.

D. Effect of the glottal chink

For these simulations, there is no lateralization, i.e. $r_l = 0\%$, the rest position, the length and the width of the lingual constriction are set to the values shown in Tab. I. The masses m_1 and m_2 are set to 0.26 g. The glottal chink opening varies from 0 to 100 %, and the subglottal pressure from 500 to 2000 Pa.

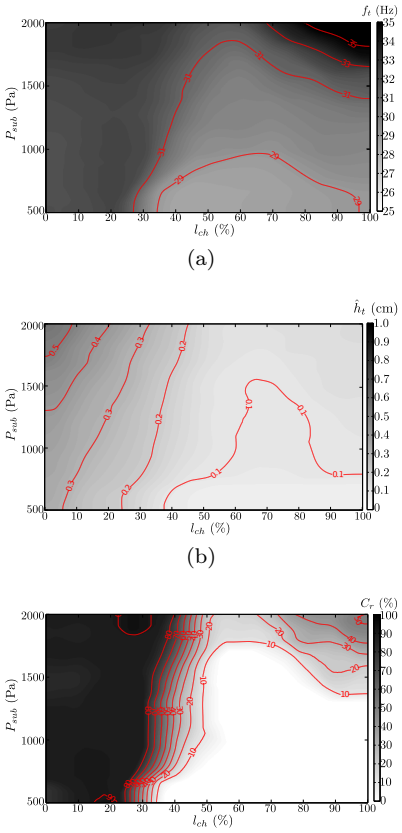


Figure 12. (a) Trill frequency f_t, (b) trill amplitude $\hat{h}_t$, and (c) trill contact ratio C_r, as a function of the subglottal pressure P_{sub} and the glottal chink opening l_{ch}. Contour lines are displayed.

Fig. 12 shows that the presence of the glottal chink has a significant impact on the trill characteristics. Indeed, it has for main effect, for a given P_{sub}, to lower

the trill frequency, the trill amplitude, and also the trill contact ratio. Indeed, when the chink opening is greater than approximately 30%, the trill amplitude is smaller than 2 mm, and the trill contact ratio shows that there is almost no linguopalatal contacts. There is a transition zone, when l_{ch} is between 25 and 35%, where the trill contact ratio is very sensitive to chink length perturbations. Before this transition, for $l_{ch} < 25\%$, the trill contact ratio is 100%, and it vanishes for $l_{ch} > 40\%$, except when P_{sub} is larger than 1700 pa, approximately.

Note that the decrease of the trill frequency for voiceless trills is in contradiction with the experimental observations by Solé[6]. However, since the trill frequency may also be increased with the subglottal pressure, the fact that voiceless trills exhibit a higher mean frequency in natural speech may also be due to an increase of the subglottal pressure in comparison with their voiced counterpart.

VII. CONCLUSION AND FURTHER WORKS

This paper has discussed the possibility to model the self-oscillations of the tongue tip in alveolar trills via classic lumped mass-spring systems. Based on models similar to those used for vocal folds, it has been shown that periodic oscillations of the tongue tip could be reproduced to simulate the production alveolar trills. In comparison with previous works[8], the acoustic model presented in this paper can support physiologically realistic area functions of the vocal tract, and also the connection with a self-oscillating model of the vocal folds with a glottal chink to simulate voiced and voiceless alveolar trills. It can also reproduce an incomplete closure of the vocal tract during linguopalatal contacts due to lateral leakage of the airflow around the tongue.

Thus, it allows the potential acoustic interactions between the vocal folds oscillations and the tongue tip to be investigated. Our simulations show that the single-mass model is subject to extremely strong interactions with the vocal folds oscillations. It is also shown that, in this case, the trill oscillation frequency is out of range of the experimentally observed trill frequency on real subjects (around 40 Hz, while it is in the range 25-30 Hz in natural speech[6,8]). The two-mass system exhibits less interactions with the periodic oscillations of the vocal folds and can oscillate at frequencies similar to the trill frequency observed in natural speech. In our simulations, the two masses do not show significant phase shifts, which suggests that, unlike the vocal folds, the main oscillation mode is the transverse mode.

The simulation framework presented in this paper also deals with the case of an incomplete closure of the vocal tract during linguopalatal contacts to be modeled thanks to the connection of lateral branches around the tongue. Although our simulations show that the connection of lateral branches does not significantly modify the oscillation characteristics of the tongue tip, it has a strong acoustic impact: the intraoral airflow and the acoustic pressure waveforms are not completely null in the closed phase of the trill, as it has been observed in natural speech in all of the previous studies[6,8]. Consequently, the incomplete closure of the vocal tract should be taken into account when dealing with the acoustic synthesis of alveolar trills using physical models.

Additionally, our simulations have revealed the fact that there is an initial position of the tongue tip that favors the production of alveolar trills. Indeed, for a given subglottal pressure, the trill amplitude as a function of the rest position of the tongue admits a local maximum, located between 0.5 and 1 mm. When the rest position of the tongue tip is greater than 1 mm out of the hard palate, the trill amplitude significantly decreases, making the production of alveolar trills more difficult. This might be one of the reason why consonant trills are hard-to-produce sounds in human spoken languages.

Finally, it has been shown that the amplitude of the tongue oscillation is significantly reduced in voiceless configurations, i.e. when the vocal folds is partially, or completely abducted. This is especially true when the length of the glottal chink exceeds a critical value, between 30 and 40% of the total length of the glottis, depending on the value of the subglottal pressure. This confirms the fact that voiceless trills are more difficult to pronounce than voiced trills, and also this explains why voiceless trills are rare among human language,

This first attempt to model the tongue tip oscillations with classic lumped mass-spring systems opens new ways of investigations of the production of trills. Indeed, one could imagine similar models applied to other trills, such at the lips for the bilabial trills, or at the uvula for the uvular trills. The possibility to support realistic geometries of the vocal tract derived form cineMRI acquisitions of real speakers is a useful tool to study the impact of various articulatory configurations on the production of trills, and consequently to investigate the reasons of the difficulty encountered by numerous speakers to produce trills. Additional data, considering phonological contexts, are planned to be acquired in the next future to complete this investigation.

ACKNOWLEDGEMENTS

This study is supported by the ANR (*Agence Nationale de la Recherche*) ArtSpeech project.

REFERENCES

[1] M. Ruhlen, A Guide to the Languages of the World., Stanford University Press, Stanford, CA, 1987.
[2] I. Maddieson, S. F. Disner, Patterns of sounds, Cambridge university press, 1984.
[3] B. C. Jimenez, Acquisition of spanish consonants in children aged 3–5 years, 7 months, Language, Speech, and Hearing Services in Schools 18(4) (1987) 357–363.

[4]M.-J. Solé, Production requirements of apical trills and assimilatory behavior, in: Proceedings of the XIVth International Congress of Phonetic Sciences, Vol. 1, 1999, pp. 487–490.
[5]E. Mendoza, Gloria Carballo, Acoustic characteristics of trill productions by groups of spanish children, clinical linguistics & phonetics 14 (8) (2000) 587–601.
[6]M.-J. Solé, Aerodynamic characteristics of trills and phonological patterning, J. of Phon. 30 (2002) 655–688.
[7]D. Recasens, M. D. Pallarès, A study of/j/and/r/in the light of the "DAC" coarticulation model, Journal of Phonetics 27(2) (1999) 143–169.
[8]R. S. McGowan, Todayngue-tip trills and vocal-tract wall compliance, J. Acoust. Soc. Am. 91(5) (1992) 2903–2910.
[9]K. Ishizaka, J. L. Flanagan, Synthesis of voiced sounds from a two-mass model of the vocal cords, Bell Syst. Tech. J. 51(6) (1972) 1233–1268.
[10]N. J. C. Lous, G. C. J. Hofmans, R. N. J. Veldhuis, A. Hirschberg, A symetrical two-mass vocal-fold model coupled to vocal tract and trachea, with application to prothesis design, Acta Acustica 84 (1998) 1135–1150.
[11]X. Pelorson, A. Hirschberg, R. R. van Hassel, A. P. J. Wijnands, Y. Auregan, Theoretical and experimental study of quasisteady-flow separation within the glottis during phonation. Application to a modified two-mass model, J. Acoust. Soc. Am. 96(6) (1994) 3416–3431.
[12]J. Flanagan, L. Landgraf, Self-oscillating source for vocal-tract synthesizers, IEEE Transactions on Audio and Electroacoustics 16 (1) (1968) 57–64.
[13]S. Maeda, A digital simulation method of the vocal-tract system, Speech Communication 1 (1982) 199–229.
[14]B. Elie, Y. Laprie, Extension of the single-matrix formulation of the vocal tract: Consideration of bilateral channels and connection of self-oscillating models of the vocal folds with a glottal chink, Speech Communication 82 (2016) 85–96.
[15]B. Elie, Y. Laprie, P.-A. Vuissoz, F. Odille, High spatiotemporal cineMRI films using compressed sensing for acquiring articulatory data, in: Eusipco, Budapest, 2016, pp. 1353–1357.
[16]A. M. Lewis, Coarticulatory effects on spanish trill production, in: Proceedings of the 2003 Texas Linguistics Society Conference, Vol. 116, 2004, p. 127.
[17]D. Recasens, Coarticulation in catalan dark [/] and the alveolar trill: general implications for sound change, Language and speech 56 (1) (2013) 45–68.
[18]P. W. Schönle, K. Gräbe, P. Wenig, J. Höhne, J. Schrader, B. Conrad, Electromagnetic articulography: Use of alternating magnetic fields for tracking movements of multiple points inside and outside the vocal tract, Brain and Language 31 (1) (1987) 26–35.
[19]P. Howson, A. Kochetov, P. van Lieshout, Examination of the grooving patterns of the czech trill-fricative, Journal of Phonetics 49 (2015) 117–129.
[20]M. Proctor, Towards a gestural characterization of liquids: Evidence from spanish and russian, Laboratory Phonology 2 (2) (2011) 451–485.
[21]A. Soquet, V. Lecuit, T. Metens, D. Demolin, Mid-sagittal cut to area function transformations: Direct measurements of mid-sagittal distance and area with MRI, Speech Communication 36(3) (2002) 169–180.
[22]Y. Laprie, M. Loosvelt, S. Maeda, E. Sock, F. Hirsch, Articulatory copy synthesis from cine X-ray films, in: Interspeech 2013 (14th Annual Conference of the International Speech Communication Association), Lyon, France, 2013, pp. 1–5.
[23]J. L. Kelly, C. C. Lochbaum, Speech synthesis, in: Proceedings of the Fourth International Congress on Acoustics, 1962, pp. 1–4.
[24]B. H. Story, Phrase-level speech simulation with an airway modulation model of speech production, Computer Speech & Language 27(4) (2013) 989–1010.
[25]B. J. Kröger, P. Birkholz, Articulatory synthesis of speech and singing: State of the art and suggestions for future research, in: Multimodal Signals: Cognitive and Algorithmic Issues, Springer, 2009, pp. 306–319.
[26]P. Mokhtari, H. Takemoto, T. Kitamura, Single-matrix formulation of a time domain acoustic model of the vocal tract with side branches, Speech Communication 50(3) (2008) 179 – 190.
[27]M. Zañartu, G. E. Galindo, B. D. Erath, S. D. Peterson, G. R. Wodicka, R. E. Hillman, Modeling the effects of a posterior glottal opening on vocal fold dynamics with implications for vocal hyperfunction, J. Acoust. Soc. Am. 136(6) (2014) 3262–3271.
[28]B. Elie, Y. Laprie, A glottal chink model for the synthesis of voiced fricatives, in: 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2016, pp. 5240–eli–5244.
[29]I. R. Titze, The physics of small-amplitude oscillation of the vocal folds, J. Acoust. Soc. Am. 83(4) (1988) 1536–1552.
[30]Z. Zhang, J. Neubauer, D. A. Berry, Aerodynamically and acoustically driven modes of vibration in a physical model of the vocal folds, The Journal of the Acoustical Society of America 120 (5) (2006) 2841–2849.
[31]E. Cataldo, F. R. Leta, J. Lucero, L. Nicolato, Synthesis of voiced sounds using low-dimensional models of the vocal cords and time-varying subglottal pressure, Mechanics Research Communications 33(2) (2006) 250–260.
[32]C. Vilain, X. Pelorson, C. Fraysse, M. Deverge, A. Hirschberg, J. Willems, Experimental validation of a quasi-steady theory for the flow through the glottis, J. of Sound and Vibration 276(3-5) (2004) 475 – 490.
[33]L. Bailly, X. Pelorson, N. Henrich, N. Ruty, Influence of a constriction in the near field of the vocal folds: Physical modeling and experimental validation, J. Acoust. Soc. Am. 124(5) (2008) 3296–3308.
[34]F. A. Duck, Physical properties of tissues: a comprehensive reference book, Academic press, 2013.
[35]Y.-C. Kim, J. Kim, M. Proctor, A. Toutios, K. Nayak, S. Lee, S. Narayanan, Toward automatic vocal tract area function estimation from accelerated three-dimensional magnetic resonance imaging.
[36]A. Giovanni, D. Demolin, C. Heim, J.-M. Triglia, Estimated subglottic pressure in normal and dysphonic subjects, Annals of Otology, Rhinology & Laryngology 109(5) (2000) 500–504.