AN ATTEMPT TO SIMULATE FLUID-WALLS INTERACTIONS DURING VELAR STOPS.

Pascal Perrier[1], Yohan Payan[2], Joseph Perkell[3], Majid Zandipour[3],
Xavier Pelorson[1], Vincent Coisy[1]& Melanie Matthies[3]
1: Institut de la Communication Parlée, INPG & Université Stendhal, Grenoble, France
2: TIMC, Université Joseph Fourier, Grenoble, France
3: Speech Communication Group, R.L.E., Massachusetts Institute of Technology, Cambridge, MA, USA

ABSTRACT

This paper presents a quantitative assessment of the role of the interaction between the airflow in the vocal tract and the mechanical structures delimiting it (the "fluid-walls interaction") in the shaping of complex articulatory paths, called articulatory loops, that are observed during the production of a velar consonant C, in VCV sequences. The work is based on simulations made with a 2D biomechanical model of the tongue coupled with a model of the airflow. Our results suggest that for low to normal levels of subglottal pressure the contribution of the "fluid-walls interaction" is slight in comparison with the contribution of the biomechanics, especially for back vowels. But in case of a strong subglottal pressure this contribution could be significant and could, in particular, explain the forward loop observed for [k] in [ika].

1 INTRODUCTION

It has been many times suggested in the literature that the inference from neurophysiological, articulatory or acoustic data of speech production mechanisms and control strategies requires beforehand a good description of the physical phenomena underlying the speech production. Indeed, it has been shown that the biomechanical and anatomical properties of the motor system are likely to have significant influence on the kinematics (trajectory shapes, velocity profiles) of human movements in general (see for example [1]) and of speech movements in particular ([2], [3]). Moreover, the interaction between the airflow in the vocal tract and the mechanical structures delimiting the vocal tract (called the "walls" in here), has also been hypothesized to contribute to the complex articulatory patterns, called the articulatory loops, observed during the production VCV sequences where C is a velar consonant ([4], [5], [6]). In a preceding work [7], we studied the possible contribution of the muscle arrangements within the tongue to this articulatory pattern. We concluded that the biomechanical and anatomical properties of the tongue musculature could be the main factor responsible for the observed loops ; we did though not preclude the hypothesis of a potential role for air pressure, which would, in addition to the biomechanics, contribute to the observed forward movement of the tongue during velar consonant closure. The aim of the present work is to provide a quantitative assessment of the possible role of the "fluid-walls interaction" in the shaping of the articulatory paths, based on simulations made with a 2D biomechanical model of the tongue coupled with a model of the airflow.

2 THE TONGUE MODEL

2.1 Biomechanical structure

The tongue model (an improved version of the model of Payan and Perrier [3]) includes the main muscles responsible for shaping and moving the tongue in the midsagittal plane (posterior and anterior parts of the genioglossus, styloglossus, hyoglossus, inferior and superior longitudinalis and verticalis). Elastic properties of tissues are accounted for by finite-element (FE) modeling of the tongue mesh in 2D defined by 221 nodes and 192 isoparametric elements. Muscles are modeled as general force generators that (1) act on anatomically specified sets of nodes of the FE structure, and (2) modify the stiffness of specific elements of the model to account for muscular structure within tongue tissues. Curves representing the contours of the lips, palate and pharynx in the midsagittal plane are added. The jaw and the hyoid bone are represented in this plane by static rigid structures to which the tongue is attached. Changes in jaw height can be simulated through a single parameter that modifies the vertical position of the whole FE structure as compared to the palate. Jaw height is kept constant during the whole duration of a simulated speech sequence (like in the bite block condition).

Collisions between tongue surface and palatal or velar contours are also modeled. The computation of the force generated during the contact is based on "a penalty method" [8], modeling a non-linear relationship between contact force and position/velocity of points located on the tongue surface (see [7] for more details).

2.2 Control of the model

The model is controlled according to Feldman's Equilibrium Point Hypothesis [9]. This theory of motor control, grounded in basic neurophysiological mechanisms of muscle force generation, suggests that the central nervous system controls movements by selecting, for each acting muscle, a threshold muscle length, λ, where the recruitment of α motoneurons (responsible for active forces) starts. If the muscle length is larger than λ, muscle force increases exponentially with the difference between the two lengths. Otherwise no active muscle force is generated. Moreover, Feldman's basic suggestion is that movements are produced from posture to posture, a posture being a stable mechanical equilibrium state of the motor system associated to a specific set of λ values. Hence, in the model, a discrete sequence of control variable values (λs), those specifying the successive postures, underlies a continuous movement. In the current version of our control model, the transition from one λ set to the next is made with constant λ rate shifts, and the onsets and offsets of the λ shifts are the same for all muscles. In other words, we hypothesize that the recruitments of all tongue muscles are synchronized. This is probably partially wrong, but, in absence of more accurate neurophysiological data, we adopted this approach.

3 MODELING OF FLUID-WALL INTERACTION

The modeling of the fluid-wall interaction implies at each time step the computation of (1) the area function from the sagittal distances generated by the 2D tongue model; (2) the volume velocity of the airflow through the vocal tract (3) the distribution of pressure within the vocal tract; (4) the pressure forces at each node of the tongue model, which are then added to the muscle forces to calculate the global forces shaping the tongue.

The area function is computed using an adapted version of the original $\alpha\beta$ model of Heinz & Stevens, where β=1.5 and α varies from the glottis to the lips according to a division of the vocal tract into 7 sections and with the value of the sagittal distance, in order to provide realistic vocal tract cross-sectional areas (see [11] for more details).

Flow velocity and pressure distribution are calculated with a flow model [12] based on a simple 2D potential flow theory, accounting for viscous losses as a perturbation of the inviscid solution. In addition, flow separation effects within a vocal tract constriction are taken into account. For the sake of simplicity, the flow separation position is estimated as the point downstream of the constriction where the cross-sectional area is 20% larger than the minimum area in the constriction [13]. The flow model is driven by a single parameter: the pressure difference $\Delta P = P_0 - P_{out}$, where P_0 and P_{out} are respectively the pressure past the glottis and at the lips.

4 SIMULATIONS

4.1 Control of tongue movements

To generate VCV sequences with the tongue model, the following simplifying strategies have been adopted for the initial simulation:

- The studied velar consonant is the non voiced stop [k]
- The production of this consonant consists in movements toward and from a virtual articulatory target that is located just beyond the palate and cannot, therefore, be actually reached. The duration of the contact against the palate depends on the "hold time" command for the virtual target and on the distance between the target and the palatal contour. The contact force acting on the tongue depends on the general level of muscular force and on the distance between the virtual target and the palate.
- Symmetrical temporal patterns of λ shifts have been chosen for the movements from the vowel toward and from the consonant toward the vowel.
- Since we consider that the target of a phoneme is context-dependent and results from a higher-level planning process (see [14]), two different targets have been used for [k]: a front one with front vowel contexts, and a back one with back vowel contexts:
- The virtual targets were chosen in order to correspond to tongue contours during the contact that are similar to the data published in the literature for [k] (for French see [15]). Both realizations only involve the recruitment of the posterior Genioglossus (GGp) and the Styloglossus (SG). In the front target the GGp generate more force and the SG less than is the posterior one.

4.2 Defining the control parameter

As explained in a previous paragraph, the aerodynamical model is entirely controlled by the single parameter $P_0 - P_{out}$. While, the pressure at the lips, P_{out} can reasonably be assumed to be constant (i.e. equals the atmospheric pressure), the value of P_0 and its evolution with time is much more complex to determine. Indeed, in a vowel-plosive-vowel configuration, the value P_0 is closely linked with the onset-offset of vocal folds vibrations.

In principle, this problem could be solved by coupling a model of the vocal folds (such as a two-mass model, for instance) to the present one. However, such a coupling is mathematically complex and is likely to generate computational instabilities.

For this reason, the temporal course of the pressure P_0 is rather generated using the pressure-flow measurements performed on a French speaker producing stop consonants. A stylized temporal function for P_0 was then derived by fitting the measured data. The pressure increases prior to the full occlusion from P_{out} to a maximum value (the subglottal pressure P_s for the unvoiced consonant [k]) within 20ms. P_0 is then maintained constant and decreases to P_{out} within 10 ms after the release of the occlusion as shown, on an example, in Figure 1.

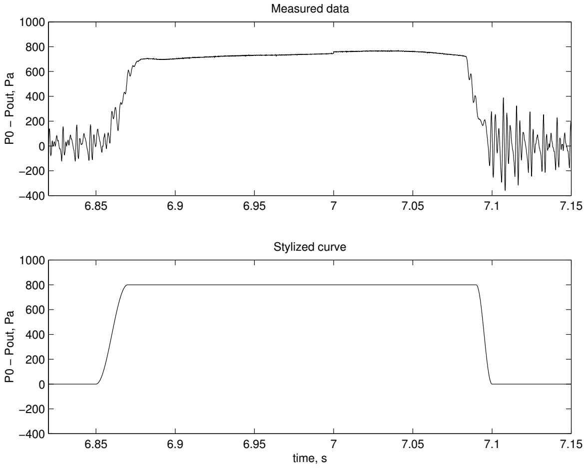


Figure 1: Time variation of the command parameter P_0-P_{out}. Top curve: measured data. Bottom curve: stylized command.

The time at which the pressure started to increase was determined by a trial procedure. A good control parameter for this task was found to be the calculated volume flow velocity.

5 RESULTS

5.1 Sequence [aka]

Figure 2 presents the trajectories of 4 nodes located on the upper contour of the tongue in the velar region, as generated by the model.

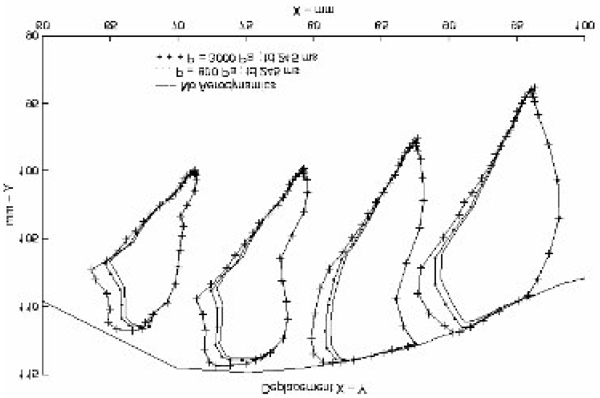


Figure 2 : Impact of the aerodynamics on the loops for [aka]

The simple solid line (SL line in here) was obtained when no account was given for the aerodynamics (NA condition); the solid line superimposed with dots (SLD line in here) depicts the trajectories obtained for a medium P_s value (800 Pa); the solid line superimposed with crosses correspond to a very high, quite unrealistic, P_s value (3000 Pa). The solid line on the top of the figure represents the hard and soft palates contour. Front is on the left of the figure. As already shown in [7], large forward loops (maximal size around 8 mm) are observed even in the NA condition. In the other conditions, the size of the loop is slightly larger, around 1/3 of mm for 800 Pa, and 1.5 mm for 3000 Pa. It should also be noted that while in the NA condition the two most posterior nodes (on the right of the figure) do not stay in contact with palate during the whole occlusion, for the two other cases with aerodynamics, the tongue is apparently pushed up more strongly against the palate.

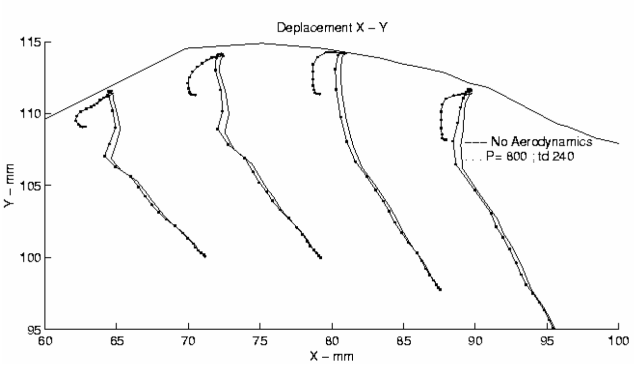
The time td mentioned on the figure is the onset time of the pressure increase, just before the occlusion. In the NA condition the occlusion starts at time 260 ms. Three values of td were then tested, 240, 245 and 255ms, and no impact was observed on the articulatory trajectories. This suggests that, for [aka], the Bernouilli effect is extremely low, and that the main effect of the aerodynamics is then essentially to push the tongue forward, but slightly in comparison to the impact of the biomechanical structure of the tongue.

5.2 Sequence [ika]

In a preceding paper [7], we mentioned for [ika] that the biomechanical model, controlled in the simple way described above, was not able to predict the experimental observations of a forward loop during the consonant, while the vowel-to-vowel transition generates typically a backward movement. Figures 3a and 3b show the results obtained with the simulation of the fluid-walls interaction as compared to the NA condition. In both cases the SL line corresponds to the NA condition and the SLB line respectively to the 800 Pa and the 3000 Pa condition.*

It can be noticed that the backward movement observed in the NA condition is reduced if $P_S = 800$Pa, and, even, that, if P_S is as strong as 3000Pa, forward loops are observed on the nodes that are not always in contact with the palate (actually very small loops can be also observed for 800Pa). These loops start with an upward deviation of the trajectory from the original one obtained in the NA condition: the tongue moves faster upward toward the palate. This is due to the Bernouilli effect. This effect is strongly depending on the relation between the level of pressure drop within a tube and the cross-sectional area of this tube. Hence the articulatory trajectories should be sensitive to changes in td values. This is, indeed, what can be observed on figure 4 that depicts the trajectory of the second node from the left, when P_S=3000Pa, for 2 different values of td, 240 and 255ms. If td=240ms, since the vocal tract pressure increases earlier, the upward deviation of the trajectory happens earlier and generate a larger loop than for td=250ms.

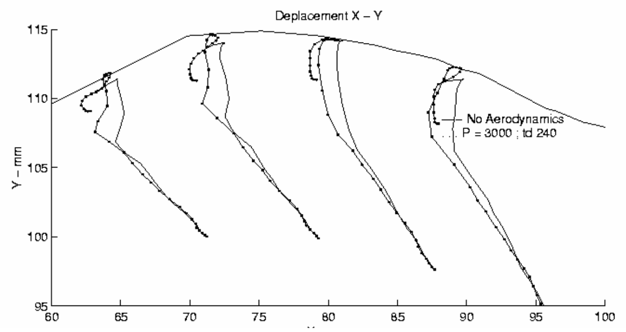
2.a.



2.b.

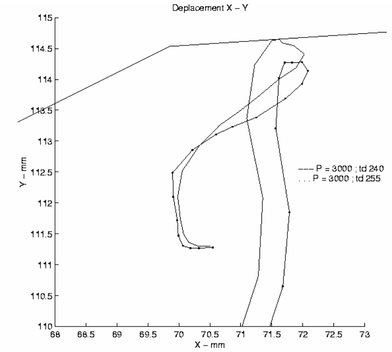
Figure 3 : Impact of the aerodynamics on the loops for [ika].

a. P=800 Pa *versus* NA condition



b. P= 3000 Pa *versus* NA condition

Figure 4 : Influence of the onset time td of the pressure



increase on the articulatory trajectories for [ika]
P_S = 3000Pa ; td = 240ms (SL line), td = 255ms (SLD line)

5.3 Sequence [iki]

For [iki], we observed, in the absence of pressure very small backward movements during the consonant. The impact of the aerodynamics (shown for P_S = 800Pa on figure 5) is small and similar to the one observed for the [ika] sequence where the amplitude of the backward movement is reduced, and a small forward loop can be observed for the nodes that are not in contact with palate. It should be noted that for theses two sequences, since the first vowel is a front vowel, the front configuration of [k] was used for the simulation.

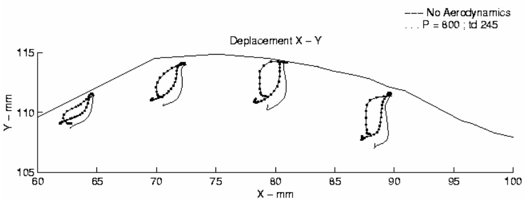


Figure 5 : Impact of the aerodynamics on the loops for [iki]

6 CONCLUSIONS

Our simulations, which were run with a 2D biomechanical tongue model coupled with a flow model, provided a quantitative assessment of the role of the aerodynamical factors on the articulatory trajectories in the production of velar consonants. Three main conclusions can be drawn out from our results:

- Large loops are mainly due to biomechanical factors such as muscle arrangements within the tongue; they are observed for posterior articulations of the consonant, and with back vowels; for these articulations, the impact of the aerodynamics seems to be negligible.
- In presence of front vowels, i.e. for the anterior articulations of the consonant, the effects of the biomechanics and of the aerodynamics are comparable, and small loops can be generated as a result of the aerodynamical factors, if the level of pressure is large enough.
- The Bernoulli effect seems to be the major aerodynamical factor that explains the generation of these loops, much more than the pressure forces that could push the tongue forward, but that are small in comparison to the muscle forces.

These conclusions have to be further carefully evaluated especially by refining the aerodynamical modeling: both the synchronization between tongue movements and pressure increase in the vocal tract, and the determination of the flow separation location are, indeed, likely to modify the amplitude of the Bernoulli effect. These two parameters have then to be very carefully chosen.

7 REFERENCES

[1] Flanagan J.R., Ostry, D.J. & Feldman, A.G. (1993). Control of trajectory modifications in target-directed reaching. *Journal of Motor Behavior, 25(3*), 140-152.

[2] Ostry, D.J., Gribble, P.L., & Gracco, V.L. (1996). Coarticulation of Jaw Movements in Speech Production: Is Context Sensitivity Centrally Planned? *Journal of Neuroscience, 16*, 1570-1579.

[3] Payan, Y. & Perrier, P. (1997). Synthesis of V-V Sequences with a 2D biomechanical tongue model controlled by the Equilibrium Point Hypothesis. *Speech Communication, 22 (2/3*), 185-205.

[4] Houde, R. (1967). *A study of tongue body motion during selected speech sound*s, Doctoral Dissertation, University of Michigan, Ann Arbor.Ostry and Flanagan, 1993

[5] Mooshammer, *C., Hoole, P. & Kuhnert, B. (1995). On loops, J.* Phonetics 23, 3-21.

[6] Hoole, P., Munhall, K. & Mooshammer, C (1994). Do air-stream mechanisms influence tongue movement paths?, *Journal of the Acoustical Society of America, 95*, 2821(A).

[7] Perrier, P., Payan, Y., Perkell, J., Zandipour, M. & Matthies, M. (1998). On Loops and Articulatory Biomechanics. *Proceedings of the 5th International Conference on Spoken Language and Processing* (Vol2., pp. 421-424). Sidney Australia.

[8] Marhefka, D.W., & Orin, D.E. (1996). "Simulations of Contact Using a Non-Linear Damping Model," in *Proc. of IEEE International Conference on Robotics and Automation, Vol. 2.*(pp. 1662-1668) (Minneapolis, MN).

[9] Feldman, A.G. (1986). "Once More on the Equilibrium-Point Hypothesis (λ Model) for Motor Control," *Journal of Motor Behavior*, *Vol. 18 (1),* pp. 17-54, 1986.

[10] Heinz, J.M. & Stevens, K.N. (1965). On the relations between lateral cineradiographs, area functions and acoustic spectra of the speech. *Proceedings of the 5th International Congress of Acoustic,* A44.

[11] Perrier P., Boë L.J. & Sock R. (1992). Vocal Tract Area Function Estimation From Midsagittal Dimensions With CT Scans and a Vocal Tract Cast: Modeling the Transition With Two Sets of Coefficients. *Journal of Speech and Hearing Research, 35*, 53-67

[12] Pelorson, X., Hirshberg, A., Wijnands, A.P.J. & Bailliet H.M.A. (1995). Description of the flow through in-vitro models of the glottis during phonation. Acta Acustica, 3, 191-202.

[13] Pelorson X., Liljencrants J., Kröger B. (1995). On the aeroacoustics of voiced sound production. *Proc. 15 th International Congress on Acoustics, Trondheim, Norway, vol 4,* 501-504.

[14] Perrier P., Lœvenbruck H. & Payan Y. (1996). Control of tongue movements in speech: The Equilibrium Point hypothesis perspective. *Journal of Phonetics, 24*, 53-75.

[15] Bothorel, A., Simon, P., Wioland, F., & Zerling, J.P. (1986). *Cinéradiographie des voyelles et consonnes du français*. Trav. de l'Inst. de Phonétique de Strasbourg, University of Strasbourg.

Acknowledgements

This work was supported by the CNRS, NSF and NIH.