

MODELING OF THE FRONT CAVITY AND SUBLINGUAL SPACE IN AMERICAN ENGLISH RHOTIC SOUNDS

Zhaoyan Zhang[†], Carol Espy-Wilson[†], Suzanne Boyce[‡], Mark Tiede^Y

[†]University of Maryland, College Park

[‡]University of Cincinnati

^YHaskins Laboratories

ABSTRACT

The production of American English (AE) /r/ sounds is variable but generally involves a large-volume front cavity. In some cases, there is also a sublingual cavity. Previous work has shown that the large front-cavity volume and the sublingual cavity are directly or indirectly responsible for the characteristically low frequency of the third formant (F3) of /r/. The entire front cavity is normally modeled as a single tube. The sublingual cavity, if present, is modeled as a side branch to the front cavity. However, given the dimensions of the front cavity, it is possible that high order acoustic modes are excited which may produce zeros and affect formant locations. Detailed information of the flow field involved will help to better understand and accurately model the front cavity acoustics. In this study, a finite element study of the flow field in the front cavity and sublingual cavity is described, using dimensions measured from MRI studies of subjects producing AE /r/. The results show that the large-volume front cavity is better modeled as a single tube with a side branch rather than as a single tube alone. The effective length of this side branch is further increased by the presence of a sublingual cavity, giving a zero in the range of F5 in the resulting spectrum.

1. INTRODUCTION

The American English (AE) /r/ sound is characterized by a lowering of the third formant (F3) so that it is close in frequency to F2. This F3 lowering can be achieved by forming a large-volume front cavity (Fig. 1a) or a large-volume front cavity plus a sublingual cavity (Fig. 1b). The sublingual cavity is generally defined as the space bounded vertically by the tongue and buccal floor. In previous work, the front cavity was modeled as a single tube of varying area functions with waves propagating in the anterior-posterior direction. If a sublingual cavity is present, it can be modeled either as an increment to the length and volume of the front cavity, or as a side branch

to the front cavity [1]. Accordingly, the lengths and area functions of the front cavity and the sublingual cavity are normally obtained by thresholding the air-tissue boundary observed from a vocal tract bracketing set of MR images scanned during sustained /r/ production [2].

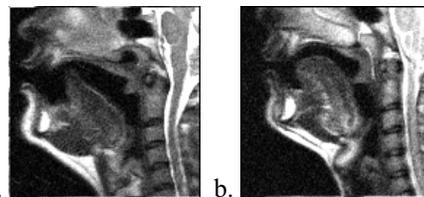


Figure 1: Midsagittal profile of two subjects producing sustained AE /r/ sounds.

It is possible that two-dimensional waves are excited in the front cavity and the sublingual cavity. Considering the comparatively large dimension involved in AE /r/ (about 1-2 cm of depth for the front cavity), and the direction of airflow, there may be air trapped at the bottom of the front cavity. This effect may exist independent of the existence of a sublingual cavity. This part of the front cavity volume would most likely join the sublingual cavity, if present, as a side branch to the main flow through the upper portion of the front cavity. This would give a greater side-branch effective length than if only the sublingual cavity is considered as a side branch. Espy-Wilson *et al.* [1] have shown that a fairly close prediction of F3 is obtained whether the sublingual cavity is modeled as an increment to the front cavity length and volume or as a side branch to the front cavity. However, there are differences in the obtained higher formants and zeros. With an increased side-branch effective length, these differences can be expected to be even greater.

The goal of this study is to provide a detailed characterization of the flow field and acoustics of the front cavity for AE /r/ production. A finite element (FEM) analysis was conducted to investigate the flow field and acoustics involved. For simplicity, two-dimensional models of the front cavity were used instead of the complete three-dimensional geometries. Two

typical front-cavity geometries (with and without a sublingual cavity) were used in this study, representing those shown in Figs. 1a and 1b respectively. For each geometry, harmonic response analysis was performed and the pressure distribution at various frequencies was calculated. The results and their implications for acoustic modeling and area function extraction from MR image sets are discussed below.

2. RESULTS

2.1. Front cavity without sublingual cavity

The front-cavity geometry without sublingual cavity, which is typical of the subject in Fig. 1a, was first studied, as shown in Fig. 2. The air flows in from the right (corresponding to the palatal constriction in Fig. 1a) to the left (corresponding to the lips). For simplicity, the lip radiation was neglected and a pressure release condition was applied at the outlet. A sinusoidal velocity input was imposed at the inlet to the model.

Fig. 3 shows the pressure amplitude contours at 1000, 2000, and 4000 Hz. At 1000 Hz and lower frequencies, the pressure amplitude contours show one-dimensional wave propagation from the inlet to the outlet in the horizontal direction (Fig. 3a). As the frequency increases, a second wave propagating in the vertical direction begins to evolve and dominate in the bottom part of the front cavity expansion (Fig. 3b). At still higher frequencies, the wave front becomes tilted toward the inlet (Figs. 3c and 3d). This evolution clearly shows that two-dimensional waves are excited in the front cavity at high frequencies (for this particular geometry, the critical frequency is about 1000 Hz). Thus, the front cavity behaves acoustically more like a single tube with a side branch.

The vocal tract acoustic response function was calculated and is shown in Fig. 4. This is compared to predictions from VTAR, a vocal tract acoustic modeling software based on the plane-wave assumption and the cross-sectional area function [3][4]. Two methods were used in determining the area function input to VTAR. In method one, the front cavity was treated as one single tube and sound waves propagate in the horizontal direction. In the second method, the bottom part of the front cavity (with depth of about 1.3 cm, which was deduced from the zero frequency in the acoustic response function in the FEM model) was treated as a side branch. Sound waves are assumed to propagate in the vertical direction in the side branch, and in the horizontal direction in the remaining portion of the front cavity. Area functions for both methods were calculated in the orientation perpendicular to the wave propagation, assuming a depth of 2 cm into the modeling plane. These two area functions were used

as input to VTAR, and the resulting acoustic response functions are shown in Fig. 4 for comparison. As expected, by modeling the front cavity as one single tube, the first method does not generate zeros in the acoustic response function. While it predicts the first resonance in the right range (a difference of about 230 Hz from the FEM prediction), the second resonance prediction is about 1610 Hz away from the FEM prediction.

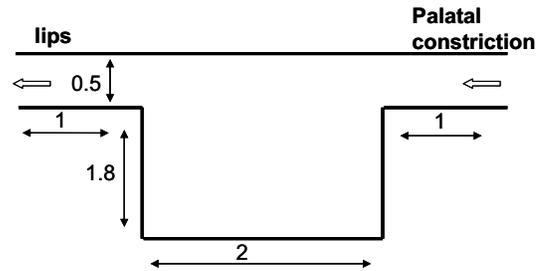


Figure 2: Front cavity without sublingual cavity in simulation 2.1. Dimensions in cm.

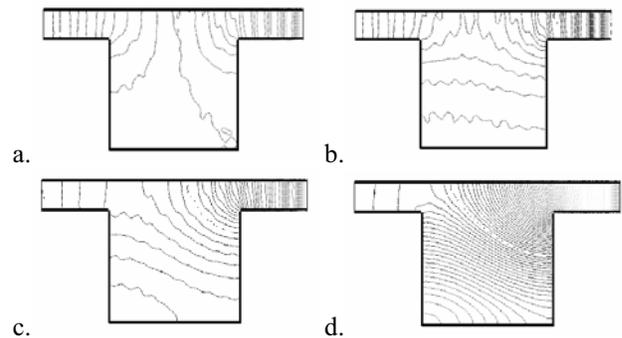


Figure 3: Pressure amplitude contours in simulation 2.1 at: a). 1000 Hz; b). 2000 Hz; c). 4000 Hz; and d). 6000 Hz.

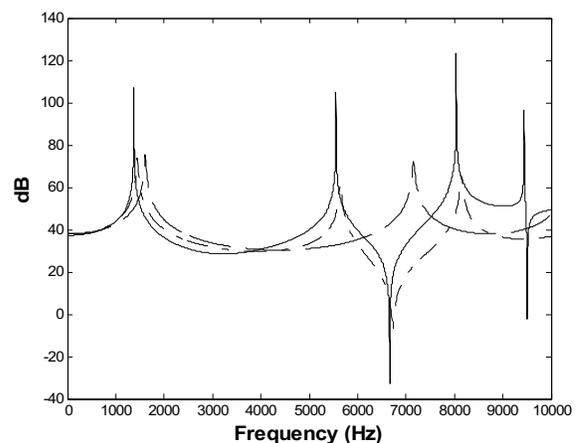


Figure 4: Acoustic response of the front cavity calculated from the FEM model (solid line) and VTAR (dashed line: modeled as a single tube; dash-dotted line: modeled as a single tube with a 1.3 cm side branch).

Note only the front cavity was included in the model. The first resonance would correspond to F3 if the whole vocal tract was modeled, and the second resonance corresponds to, in this particular case, F6 in the complete vocal tract acoustic response function. Therefore both methods give fairly accurate predictions of the acoustic response function up to frequencies in the range of F5. However, the effective length of the side branch is expected to be increased by the presence of a sublingual cavity. This will lower the zero frequency into the range of F4/F5, and make a larger difference in the resulting acoustic response between these two methods. This difference is discussed further in the next section.

2.2. Front cavity with sublingual cavity

Fig. 5 shows a simplified model of the front cavity representing the type shown in Fig. 1b. The dimensions used in the model are estimates from the midsagittal profile picture. A sublingual cavity is present as outlined in dashed lines. As in the previous model, the inlet (on the right side) corresponds roughly to the palatal constriction. The outlet corresponds to the lips and a pressure release condition was imposed neglecting the radiation impedance.

Harmonic analysis of the model was conducted with imposed sinusoidal velocity input at the inlet. The resulting pressure amplitude contours are shown in Fig. 6 for four different frequencies. Similar behavior of the flow as seen in the previous model was observed. At low frequencies, sound waves propagate in the front cavity as a quasi-one-dimensional wave, with minimal air trapped in the sublingual cavity. As frequency increases, two-dimensional waves begin to establish themselves. As expected, the airflow in the sublingual cavity, instead of acting independently as a side branch to the entire front cavity, combines with the bottom portion of the front cavity and behaves as a side branch to the main flow in the upper portion of the front cavity.

The acoustic response of the front cavity was calculated from the FEM model and is shown in Fig. 7. Also shown in the figure is the acoustic response for the model excluding the sublingual cavity, i.e., the geometry shown in Fig. 5 without the portion enclosed by the dashed lines. As shown in the figure, both models have similar first resonance frequencies, but different zero frequencies. In the model without the sublingual cavity, the zero frequency is about 6716 Hz, corresponding to an effective quarter-wavelength resonator length of 1.30 cm, which is the portion of the front cavity with waves propagating upwards. For the model with a sublingual cavity, this effective length is increased by its presence, which has an effective length of approximate 0.6 cm considering its

tapering geometry. This gives a lowered zero frequency of about 4827 Hz in the resulting acoustic response function from the original model. Note the zero will be located at about 14.6 kHz only if the sublingual cavity is modeled as a side branch to the entire front cavity.

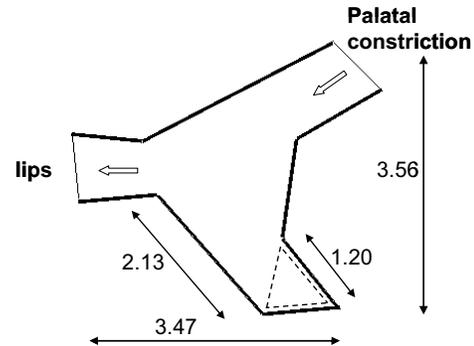


Figure 5: Front cavity without sublingual cavity in simulation 2.1. Dimensions in cm.

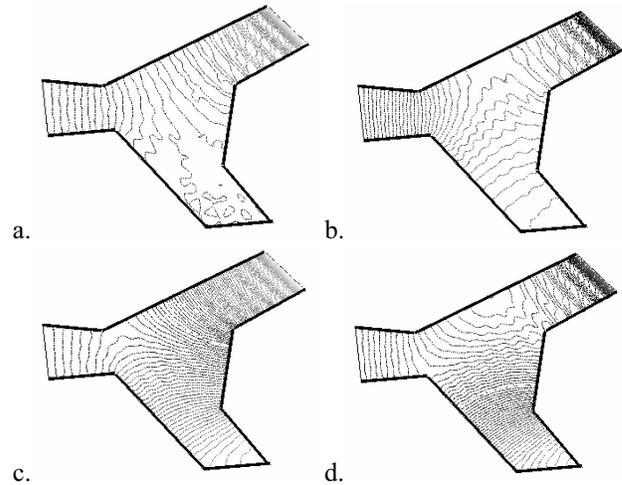


Figure 6: Pressure amplitude contours in simulation 2.2 at: a). 1000 Hz; b). 2000 Hz; c). 4000 Hz; and d). 6000 Hz.

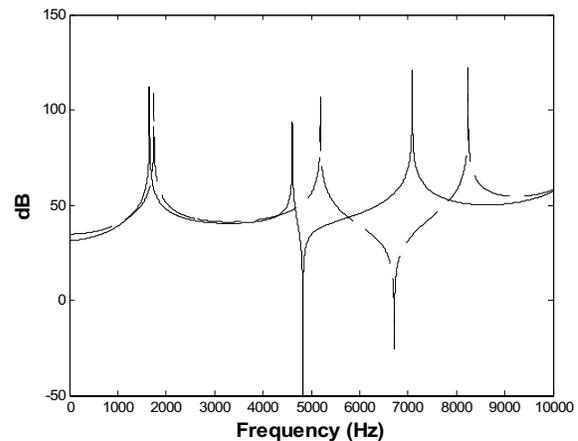


Figure 7: Acoustic response of the front cavity calculated from FEM models with (solid line) and without (dashed line) the sublingual cavity.

3. CONCLUSIONS

The simulations discussed above have shown that the large-volume front cavity is better modeled as a single tube with a side branch, instead of a single tube alone, as in previous work. However, both methods give fairly accurate predictions of the first front cavity resonance, or F3. The difference lies in the second resonance frequency region, and, depending on the specific geometry, this could lie in the range of F4/F5. If only the first three formants are of interest, it suffices to model the entire front cavity as a single resonator and to assume anterior-posterior plane wave propagation. However, the presence of a zero in the spectrum will certainly affect any formant frequency nearby in the frequency axis. In extreme cases, when a deep front cavity is involved or a long sublingual cavity is present, the zero frequency will be even lower, in the neighborhood of 3 kHz [5]. In such cases the more accurate way of modeling outlined in this study is then desirable.

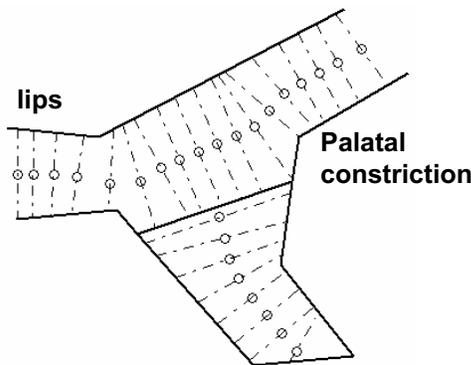


Figure 8: Vocal tract midlines and cross sections used for area function extraction.

This study also has implications for area function extraction from MR images. To separate the front cavity from a sublingual space requires first dividing the front cavity into two portions, and then scanning the two portions in directions perpendicular to wave propagation in both portions. Fig. 8 shows an example for the model used in section 2.2. The particular location to divide the front cavity is highly dependent on the specific geometry involved, and this will be the objective of future work.

4. ACKNOWLEDGMENT

This work was supported by NIH grant 1 R01 DC05250-02.

5. REFERENCES

- [1] C.Y. Espy-Wilson, S.E. Boyce, M. Jackson, S. Narayanan, and A. Alwan, "Acoustic modeling of American English /r/," *J. Acoust. Soc. Am.*, pp. 343-356, 2000.
- [2] A. Alwan, S. Narayanan, and K. Haker, "Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part II. The rhotics," *J. Acoust. Soc. Am.*, pp. 1078-1089, 1997.
- [3] Z. Zhang, and C.Y. Espy-Wilson, "A vocal tract model of American English /l/," *J. Acoust. Soc. Am.*, pp. 1274-1280, 2004.
- [4] X. Zhou, Z. Zhang, and C.Y. Espy-Wilson, "VTAR: A Matlab-based computer program for vocal tract acoustic modeling," *J. Acoust. Soc. Am.*, pp. 2543, 2004.
- [5] K.N. Stevens, *Acoustic Phonetics*, The MIT Press, Cambridge, Massachusetts, 1998.