XXXVI Ibero-Latin American Congress on Computational Methods in Engineering
CILAMCE
Rio de Janeiro, 22-25 Nov 2015

ANALYTICAL AND NUMERICAL MODELING OF VOCAL TRACT IN VOWEL PHONATION

Álvaro Campos Ferreira

Marcus Vinicius Girão de Morais

alvaro.campos.ferreira@gmail.com

mumorais@gmail.com

Universidade de Brasília

Campus Darcy Ribeiro, 70910-900, Distrito Federal, Brasília, Brasil

Abstract. *The acoustical properties of the vocal tract have been the subject of several studies due to the complex nature of the human voice. The complex geometry the vocal tract assumes during the phonation of vowels is obtained from MRI and acoustic models are constructed to determine the resonance frequencies, also known as formants, which characterizes the sound of the vowel. This information can be used by health professionals to identify possible voice pathologies. The present work shows the acoustic behavior of a simplified model of a vocal tract based on the experimental measurements of a vocal tract during the phonation of vowels. Analytical and numerical models are used to predict the acoustic behavior of the vocal tract geometry. Three models are provided: a transfer matrix analytical model, a boundary element method model and a finite element method model. The models have shown good agreement for distinct resonance frequencies and related mode shapes. A comparison between the mode shapes for each resonance frequency is carried out and three dimensional effects are identified. The mode shapes produced by the present work are consistent with the acoustic behavior of open-closed cavities.*

Keywords: *modal analysis, vocal tract, boundary element method, transfer matrix method, acoustics*

1 INTRODUCTION

The acoustics of the vocal tract is an important study area for professionals which use the voice as an instrument to convey information and ideas, such as teachers, orators, singers and theater professionals. A resonant voice is desirable to obtain a high amplitude of the spoken sound with a lowering or even absence of excessive strain in the vocal folds (I. Titze, 2001). This resonant quality is already well known by bel-canto singers and theater professionals, but its production is still based in body sensations of the speaker. The study of the acoustic properties of the resonant voice and vocal tract during phonation are an essencial step for the advance of vocal technology and vocal coaching for professionals which use spoken speech.

Ekholm, Papagiannis, and Chagnon (1998) tries to fill the gap in the communication between voice pedagogues and voice scientists. Subjective ratings were related to objective measurements taken from acoustic analysis of the voice signal. Some acoustic phenomena correlations to critical perceptual parameters were identified, helping to bridge the terminology gap between vocal artists and scientists.

It is known that the acoustic resonance of the vocal tract defines the quality of the vowel produced. Different vocal tract configurations produces different vowels, with different modes and resonance frequencies, often called 'formants' in phonoaudiology. A study of the resonance strategies of 22 trained singers shows that trained singers may use the resonance of the vocal tract to obtain certain desirable characteristics in the voice (Henrich, Smith, & Wolfe, 2011). The two first resonance frequencies of the vocal tract were obtained experimentally. These results indicate that singers can repeatedly tune their resonances with precision, and that there can be considerable differences in the resonance strategies used by individual singers, particularly for voices in the lower ranges.

The study of the resonance of the vocal tract during phonation of high pitches by trained tenors is also the subject of research (I. R. Titze, Mapes, & Story, 1994). Results suggest that tenors maintain their first formant well above the fundamental for all vowels except /u/. The purpose of this seems to be to distribute the acoustic energy between the other harmonics rather than boosting the fundamental. Tuning the first formant to the fundamental, or boosting the fundamental, is a technique used by sopranos but seems to be deliberately avoided by tenors to preserve male quality. An analysis by synthesis of the vocal tract is carried out to determine the resonance frequencies of the vocal tract. A parameter optimization was used to match the spectra of the model to the measured spectra. It appears that a traditional source-filter analysis can explain the power spectra of the high tenor voice.

1.1 VOICE ACOUSTICS

A schematic model of the speech system, which includes the vocal tract is shown in Figure 1, adapted from the work of Cataldo, Sampaio, and Nicolato (2004). The portion between the glotis and mouth will be modelled in the present study.

The study of vocal tract models with approximated geometry is possible using images from standard Magnetic Resonance Imaging (MRI) or similar procedures. The procedure is carried out during phonation of a specific vowel and pitch and the area functions are

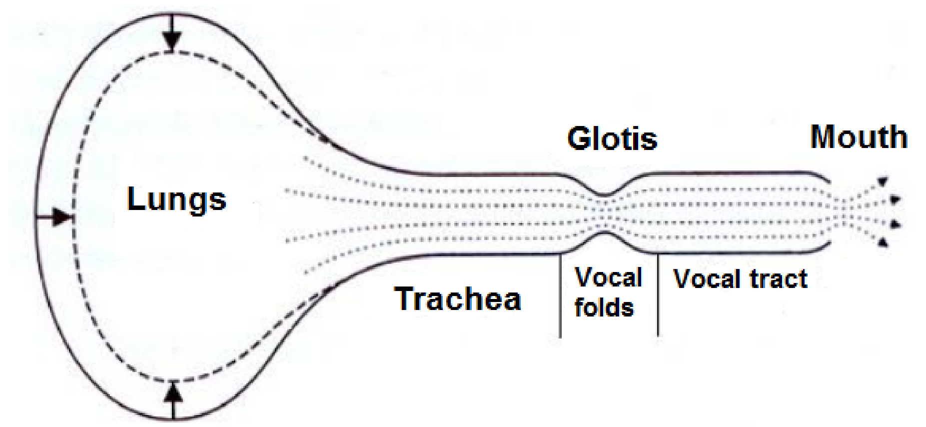


Figure 1: Schematic model of the speech system (Cataldo et al., 2004)

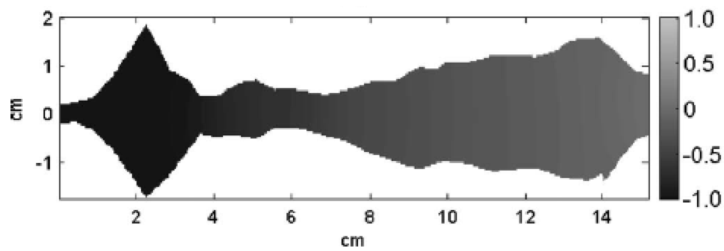
inferred from the MRI images. From that point on, numeric and analytical methods may be used to predict the resonance frequencies and mode shapes for the vocal tract shape. This is the procedure used by Clement et al. (2007) for the prediction of the three first resonance frequency and mode shapes of the vocal tract. This three first resonance frequencies are the most important for the identification of the vowel. The analytic method used in this study for the prediction of the resonance frequencies is unidimensional, therefore any three-dimensional effects are neglected. Only the lenght and area function of the vocal tract are used in the analytic model.

According to recent model investigations, vocal tract resonance is relevant to vocal registers. Echternach, Sundberg, Markl, and Richter (2011) investigates the difference between the vocal tract shape in different types of phonation: modal and falsetto registers. Modal register is more energetic and engages a bigger portion of the vocal folds. Falsetto register is softer and engages a lesser portion of the folds. The subjects produced a sustained pitch (349 Hz) on the vowel /a/ in both registers while in a MRI machine. The area functions from the MRI machine were used to estimate the resonance frequencies of the vocal tract, which were consistent with later measurements of the same subjects in a sound treated room. Vocal tract shape differed between modal and falsetto register. In modal as compared to falsetto the lip opening and the oral cavity were wider and the first formant frequency was higher. In this sense the presented results are in agreement with the claim that the formant frequencies differ between registers. The shape of the vocal tract during these different types of phonation were measured in Tom, Titze, Hoffman, and Story (2001) as well. Differences in vocal tract configuration and formant characteristics are shown.

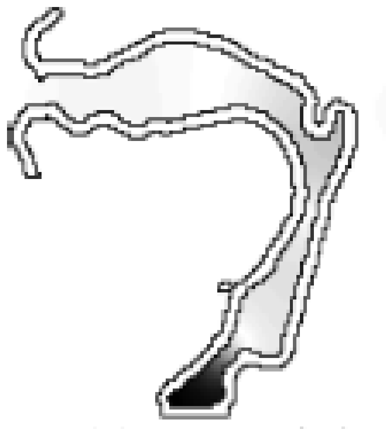
Different exercises are constantly performed by voice coachs and the acoustics of these exercices may be studied, as in Story, Laukkanen, and Titze (2000). The acoustic characteristics of a posterior semi-occlusion of the vocal tract by bilabial fricatives or with small diameter tubes are obtained. Analytical models are build using area functions to obtain the resonance frequencies and mode shapes. The author concludes that this semi-occlusion, either by the bilabial fricatives or small diameter tubes, increase the impedance of the vocal tract and lower the first resonance frequency, which helps maintain the inertance of the vocal tract. This helps the speaker to maintain this desirable inertance of the vocal tract even after the removal of the semi-occlusion.

Exercises using small diameter tubes is tackled by Guzman et al. (2013). A trained singer produces an /a/ vowel in a comfortable pitch inside a computed tomography (CT) machine. After fifteen minutes of silence, the singer exercises his phonation using a resonance tube and, again, using a small diameter straw. Images from the CT shows that the velum closed the nasal passage, the laringeal position was lowered and the hypopharynx area was increased. The singer formant was identified more prominently after either the resonance tube or straw exercises. The author concludes that the use of the resonance tubes during phonation increases the acoustic impedance which prompts the singer to increase phonation efficiency after the exercises.

The acoustic study of the vocal tract during phonation is carried out by Hannukainen, Lukkari, Malinen, and Palo (2007) and Takemoto, Parham, and Mokhtari (2010), using the finite element method and finite difference method, respectively. Approximate geometry of the vocal tract are obtained using MRI images are used to construct the models. Figures 2a and 2b shows the normalized acoustic pressure for the first resonance mode of the vowels /o/ and /a/, respectively.



(a) Vowel /o/ from Hannukainen et al. (2007)



(b) Vowel /a/ from Takemoto et al. (2010)

Figure 2: Normalized acoustic pressure for the first resonance mode of the vowels /o/ and /a/

A study aimed to build a more realistic model of the vocal tract and the sound radiation from the vocal tract into free space using the boundary element method is presented by (Kagawa, Shimoyama, Yamabuchi, Murai, & Takarada, 1992). The resonance frequencies and mode shapes were calculated using the BEM for five japanese vowels under various wall conditions, including a muscular impedance and rigid wall conditions. The effects on these conditions on resonance frequencies were examined. The sound field around the head and body when sound is radiated from the mouth was separetely calculated. Scattering due to the head and body were demonstrated.

2 BOUNDARY ELEMENT METHOD

The formulation used in this work is the direct boundary element method for the Helmholtz equation, described in more details in Dominguez (1993), Wrobel (2001) and Kirkup (2007).

The propagation of acoustic waves through a fluid medium Ω is described by the wave equation. When the motion is assumed to be time-harmonic, the wave equation reduces to the Helmholtz equation and take the form shown in Eq. (1).

$$\nabla^2\phi + k^2\phi = 0 \tag{1}$$

where ϕ is a reduced velocity potential, $k = \omega/c$ is the wave number, c is the speed of sound and ω is the angular frequency. The velocity potential ϕ is related to the acoustic pressure p as shown in Eq. (2).

$$p = -i\rho\omega\phi \tag{2}$$

where ρ is the density of the fluid medium Ω.

The boundary integral equation for the problem can be found by starting from Green's second identity written in Eq. (3).

$$\int_\Omega \left(\phi\nabla^2\phi^* - \phi^*\nabla^2\phi\right) d\Omega = \int_\Gamma \left(\phi\frac{\partial\phi^*}{\partial n} - \phi^*\frac{\partial\phi}{\partial n}\right) d\Gamma \tag{3}$$

where ϕ^* is a weight function and Γ is the boundary of domain Ω. Substituting the Laplacian $\nabla^2\phi$ in Eq. (1) for $-k^2\phi$ from Eq. (3) gives

$$\int_\Omega \phi\left(\nabla^2\phi^* + k^2\phi^*\right) d\Omega = \int_\Gamma \left(\phi\frac{\partial\phi^*}{\partial n} - \phi^*\frac{\partial\phi}{\partial n}\right) d\Gamma \tag{4}$$

If the weight function ϕ^* satisfies Eq. (5), then the function is said to be a fundamental solution of Eq. (1) and it corresponds to the field generated by a unit concentrated harmonic source.

$$\nabla^2\phi^*(X', x) + k^2\phi^*(X', x) = -\delta(X', x) \tag{5}$$

where X' is the source point, x is the field point and δ is the Dirac delta function. Introducing the property of Eq. (5) in Eq. (4) produces the boundary integral representation of the problem, shown in Eq. (6).

$$\phi(X') = \int_\Gamma \frac{\partial\phi(x)}{\partial n}\phi^*(X', x)d\Gamma - \int_\Gamma \phi(x)\frac{\partial\phi^*(X', x)}{\partial n}d\Gamma \tag{6}$$

The fundamental solution and the normal derivative of the fundamental solution for three-dimensional problems are shown in Eqs. (7) and (8), respectively.

$$\phi^* = -\frac{e^{ikr}}{4\pi r} \tag{7}$$

$$\frac{\partial\phi^*}{\partial n} = -\frac{e^{ikr}}{4\pi r}\left(ik - \frac{1}{r}\right)\frac{\partial r}{\partial n} \tag{8}$$

where $r = |X' - x|$ is the distance between the source point X' and the field point x. Taking X' to the boundary Γ in Eq. (6), in view of the behaviour of the fundamental solution when $X' \to x'$, $x' \in \Gamma$ produces Eq. (9).

$$c(x')\phi(x') = \int_{\Gamma} \frac{\partial \phi(x)}{\partial n} \phi^*(x', x) d\Gamma - \int_{\Gamma} \phi(x) \frac{\partial \phi^*(x', x)}{\partial n} d\Gamma \tag{9}$$

where $c(x') = 1/2$ for smooth boundary on x'.

The BEM is a numerical method of solution of boundary integral equations, based on a discretization procedure. Two types of approximations are required for the application of the method: the first is geometrical and involves a subdivision of boundary Γ in N small segments, such that $\Gamma \approx \sum_{i=1}^{N} \Gamma_i$; the second is an approximation of the variation of the velocity potential and its normal derivative within each element. The simplest possible approximation is a constant one, which assumes that ϕ and $\partial\phi/\partial n$ are constant within each element and equal to their value at the midpoint. Introducing this approximations, one obtains Eq. (10).

$$c_j \phi_j = \sum_{i=1}^{N} \frac{\partial \phi_i}{\partial n} \int_{\Gamma_i} \phi^*(x', x) d\Gamma - \sum_{i=1}^{N} \phi_i \int_{\Gamma_i} \frac{\partial \phi^*(x', x)}{\partial n} d\Gamma \tag{10}$$

where ϕ_i and $\partial\phi_i/\partial n$ are the values of ϕ and $\partial\phi/\partial n$ at the node located in the midpoint of element i. In the case of constant elements, the number of nodes is equal to the number of elements.

Calling

$$H_{ij} = c + \sum_{i=1}^{N} \left(\int_{\Gamma_i} \frac{\partial \phi^*}{\partial n} d\Gamma \right), c = \begin{cases} \frac{1}{2} \; if \; i = j \\ 0 \; if \; i \neq j \end{cases} \tag{11}$$

and

$$G_{ij} = \sum_{i=1}^{N} \left(\int_{\Gamma_i} \phi^* d\Gamma \right) \tag{12}$$

Equation (10) takes the following form

$$\sum_{i=1}^{N} H_{ij} \phi_i = \sum_{i=1}^{N} G_{ij} \frac{\partial \phi_i}{\partial n} \tag{13}$$

Using a collocation technique to all nodal points along the boundary gives a system of equations which can be written in matrix form as shown in Eq. (14).

$$[H]\{\phi\} = [G]\{q\} \tag{14}$$

where $\{\phi\}$ and $\{q\}$ are vectors containing the nodal values of the potential and its normal derivative; $[H]$ and $[G]$ are square $N \times N$ matrices of influence coefficients. The diagonal term $[H_{ii}]$ is integrated analytically and its value corresponds to $[H_{ii}] = 1/2$. The term $[G_{ii}]$ is integrated numerically, with the difference that the unitary normal vector is substituted by a null vector.

Once the boundary conditions of the problem are applied, the problem stated in Eq. (14) can be rearranged in the form shown in Eq. (15).

$$[A]\{x\} = \{b\} \tag{15}$$

where $[A]$ is a matrix containing all knowns, $\{x\}$ is a vector containing all unknowns of the problem and $\{b\}$ is the 'load' vector. This system is solved using a standard direct solver. A reverse arrangement is used to obtain the original system $[H]\{p\} = [G]\{q\}$.

Once the values of the unknowns are obtained in the boundary, the value of the velocity potential in any internal point is obtained using Eq. (16).

$$\phi_j = \sum_{i=1}^{N} \frac{\partial \phi_i}{\partial n} \int_{\Gamma_i} \phi^*(x', x) d\Gamma - \sum_{i=1}^{N} \phi_i \int_{\Gamma_i} \frac{\partial \phi^*(x', x)}{\partial n} d\Gamma \tag{16}$$

where ϕ_j is the value of the velocity potential for internal point j.

3 RESULTS

Using the information for the geometry of the vocal tract from Clement et al. (2007), a BEM model was set up to predict the resonance frequencies and modes of two different vowels: /a/ and /u/. Three-dimensional meshes were created using the commercial software GiD 11.0.7.

The total length of the vocal tract was estimated from the MRI information shown in Table 1. Each cross-section area corresponds to one centimeter distance from the last. The areas were approximated by square sections. The length of each side is the square root of the cross-section area.

Table 1: Cross-section area of the vocal tract for different vowels $[cm^2]$

Vowel cross-section	1	2	3	4	5	6	7	8	9
/a/	1.8	1.8	2.8	1.5	0.8	1.3	1.5	1.7	2.8
/u/	1.4	3.4	6.6	8.7	9.8	9.2	7.6	5.8	4.4
Vowel cross-section	10	11	12	13	14	15	16	17	
/a/	4.5	7.1	9.3	13.5	15.5	7.8			
/u/	2.3	0.7	1.5	3.9	8.6	6.9	1.5	1.3	

BEM models were created for vowels /a/ and /u/. These models were created using constant elements. The mesh for vowel /a/ consists of 226 constant elements and the mesh for vowel /u/ consists of 619 constant elements. The meshes for vowels /a/ and /u/ can be seen in Figures 3 and 4 respectively.

Using the models created, a frequency sweep was performed using the BEM program creating a frequency response function (FRF) for each vowel. This allowed the determination of the resonance frequencies. Table 2 compares the resonance frequencies obtained using BEM, the transfer matrix method (MT) and a finite element method (FEM) model created using commercial software ANSYS. The FEM model for vowel /a/ contained 18271 nodes and 15000 tetrahedral fluid elements and the model for vowel /u/ contained 20691 nodes and 17000 tetrahedral fluid elements.

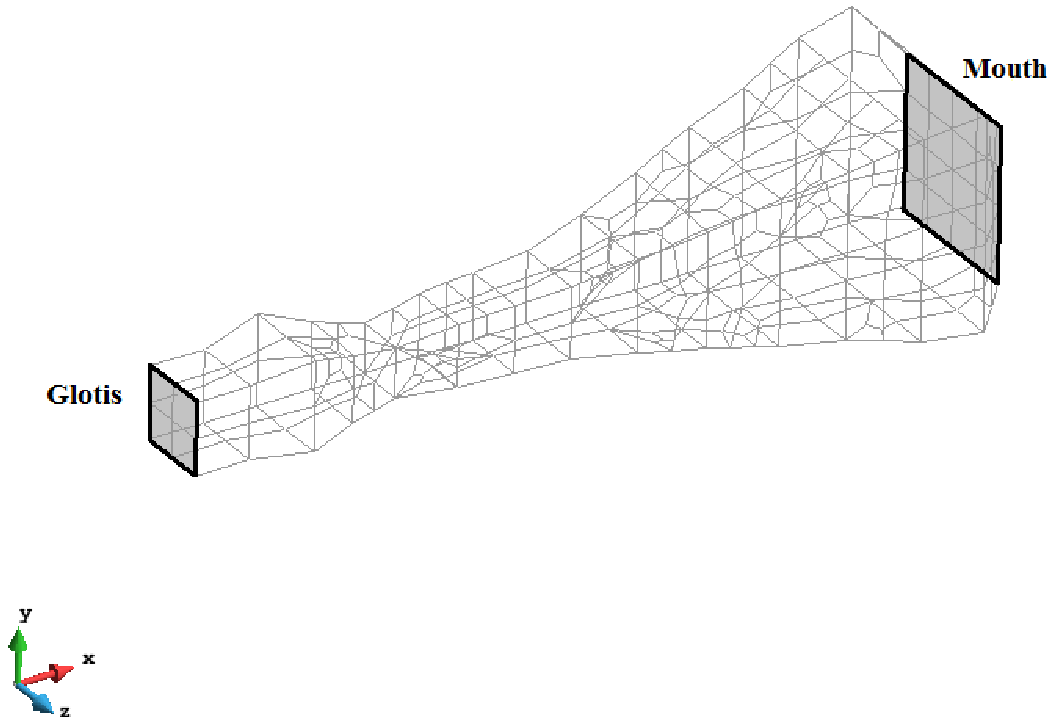


Figure 3: Mesh created for vowel /a/

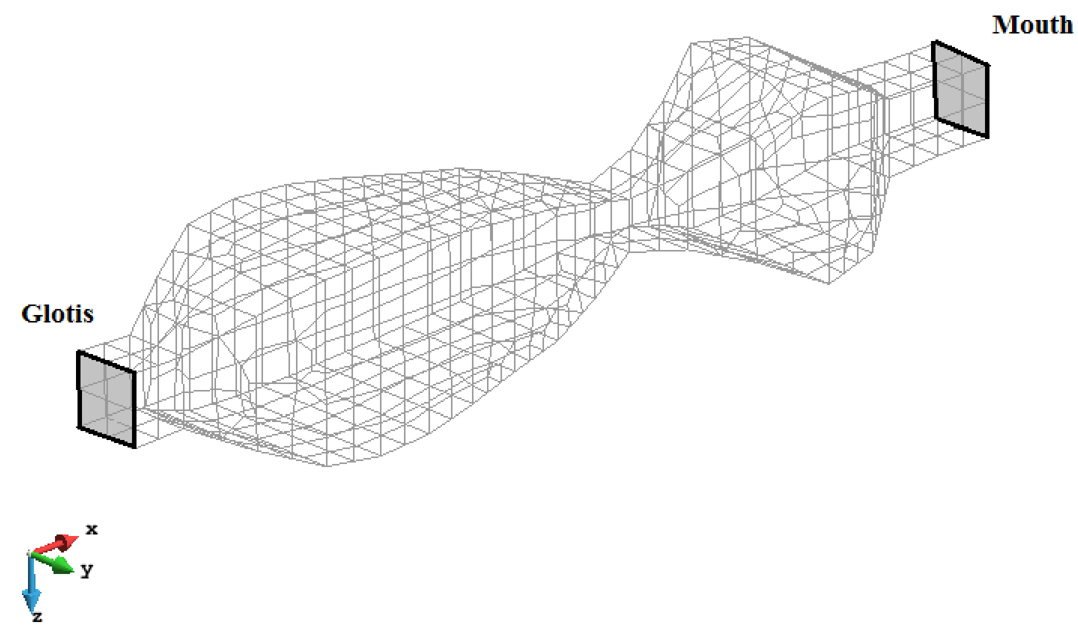


Figure 4: Mesh created for vowel /u/

Table 2: First, second and third resonance frequencies for the vocal tract producing different vowels

	Clement et al. (2007)	MT	BEM	FEM
First mode [Hz]				
Vowel /a/	774	774	804	804
Vowel /u/	340	323	320	317
Second mode [Hz]				
Vowel /a/	1300	1580	1767	1460
Vowel /u/	1114	1182	1145	1139
Third mode [Hz]				
Vowel /a/	2724	2852	2814	2808
Vowel /u/	2538	2600	2678	2678

The FRF plot for vowel /a/ is shown in Figure 5. The resonance frequencies can be identified in the figure as the three peaks. As these peaks don't correspond exactly to the resonance frequency, but are only an approximation, the modes obtained are operational modes and not resonance modes. These operational modes will approximate the resonance modes as the operational frequency approximates the resonance frequency.

Using equation (16), it is possible to determine the pressure variation inside the of the boundary of each vowel. This helps to create a comparison plot of the different models created in this work. Figures 6, 7 and 8 show the comparison between the BEM, FEM and MT models. The graph shows the normalized pressure in a central line throughout the vocal tract for vowel /a/. It is worth noting that the mesh for vowel /a/ is considerably rough, containing only 226 constant elements. It is possible to identify the three-dimensional effects in the BEM and FEM models which are not present in the MT model. The transfer matrix solution disconsiders this effect as the transition between the cross-sections areas is discontinuous.

Figure 9 shows the FRF plot for vowel /u/ obtained using the BEM. In this figure it is possible to identify the first three resonance frequencies.

Using the same procedure as with vowel /a/, a comparison is made between the different solutions for vowel /u/. Figures 10, 11 and 12 show this comparison for the first three resonance frequencies, respectively.

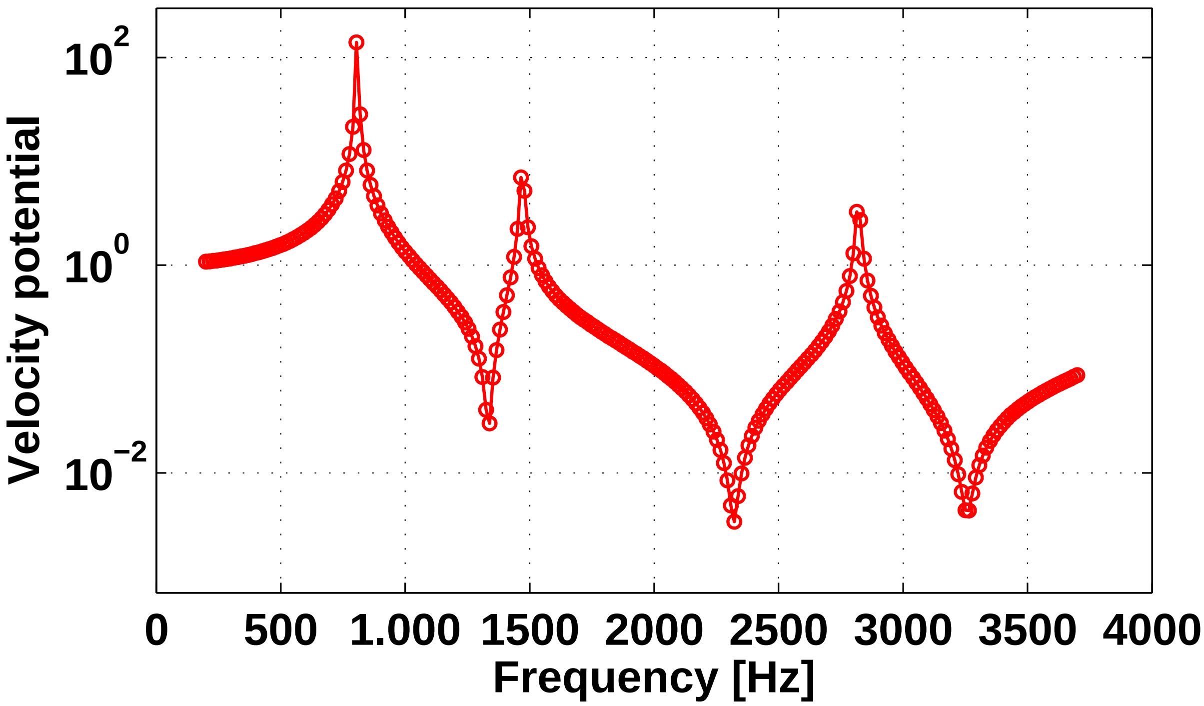


Figure 5: Frequency response function for vowel /a/ showing the first three resonance frequencies

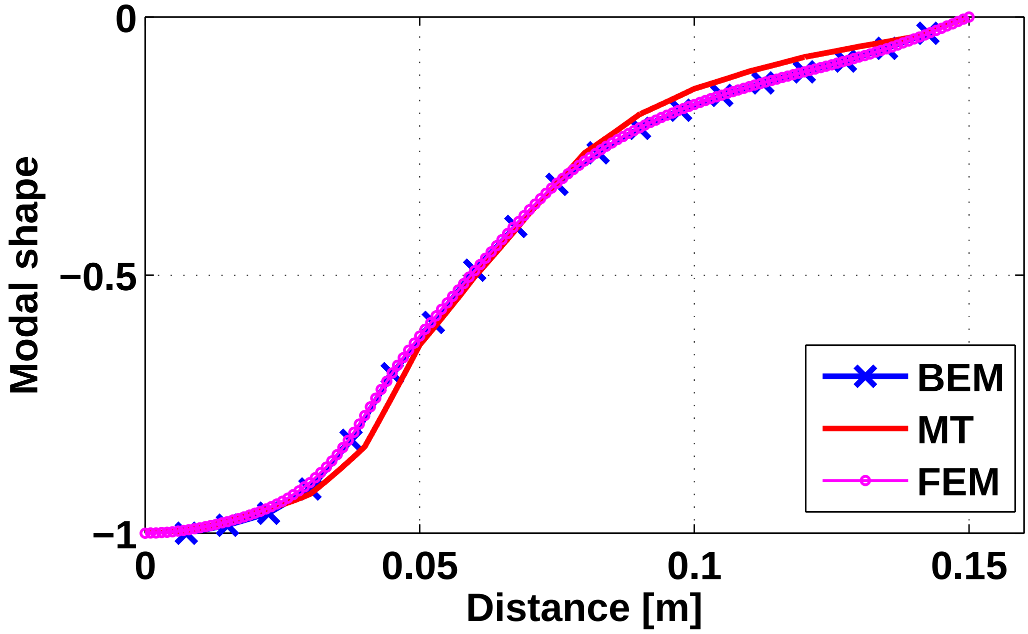


Figure 6: First modal shape comparison between different solutions for vowel /a/

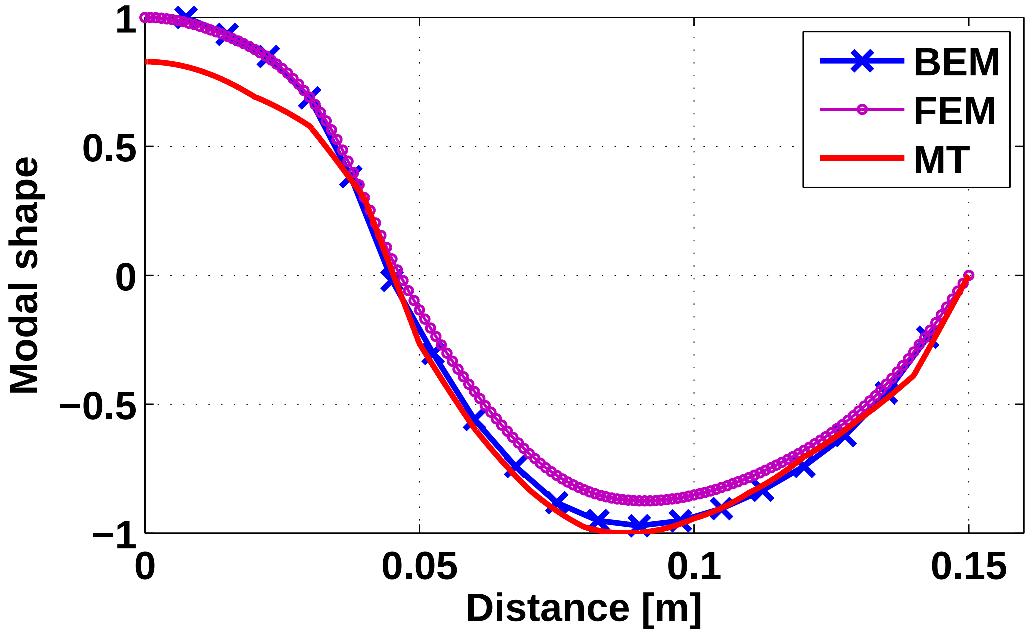


Figure 7: Second modal shape comparison between different solutions for vowel /a/

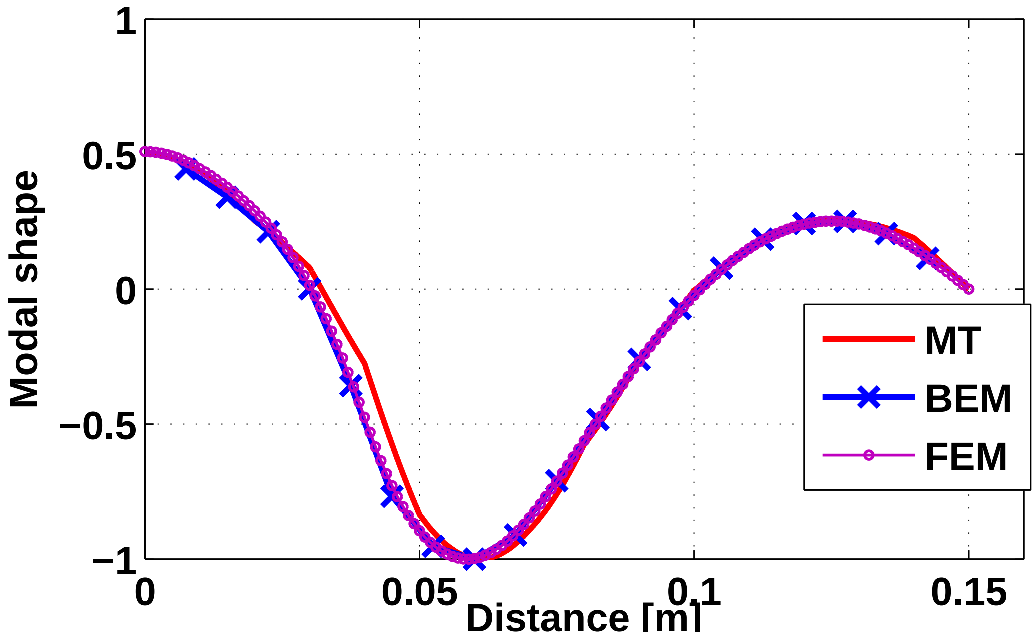


Figure 8: Third modal shape comparison between different solutions for vowel /a/

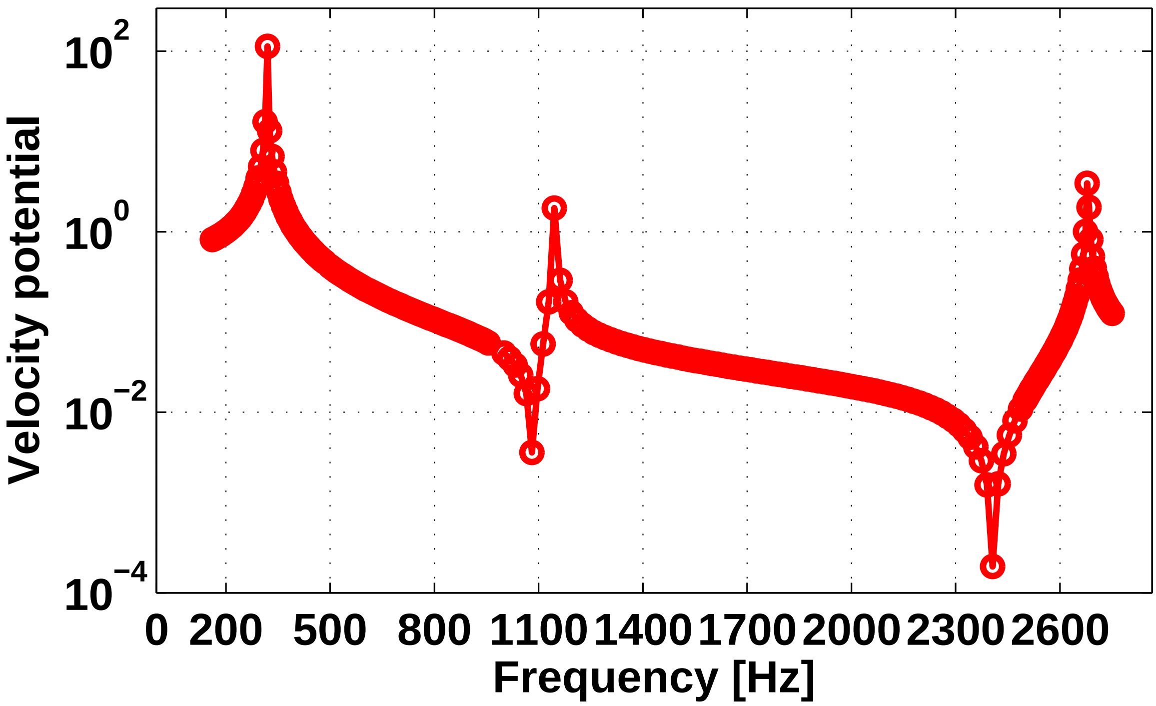


Figure 9: Frequency response function for vowel /u/ showing the first three resonance frequencies

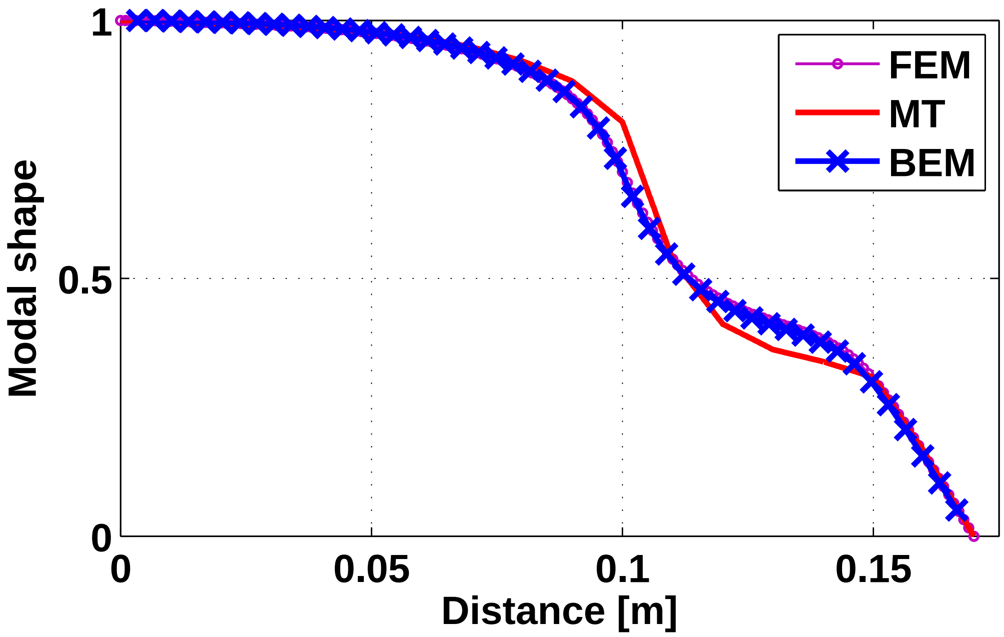


Figure 10: First modal shape comparison between different solutions for vowel /u/

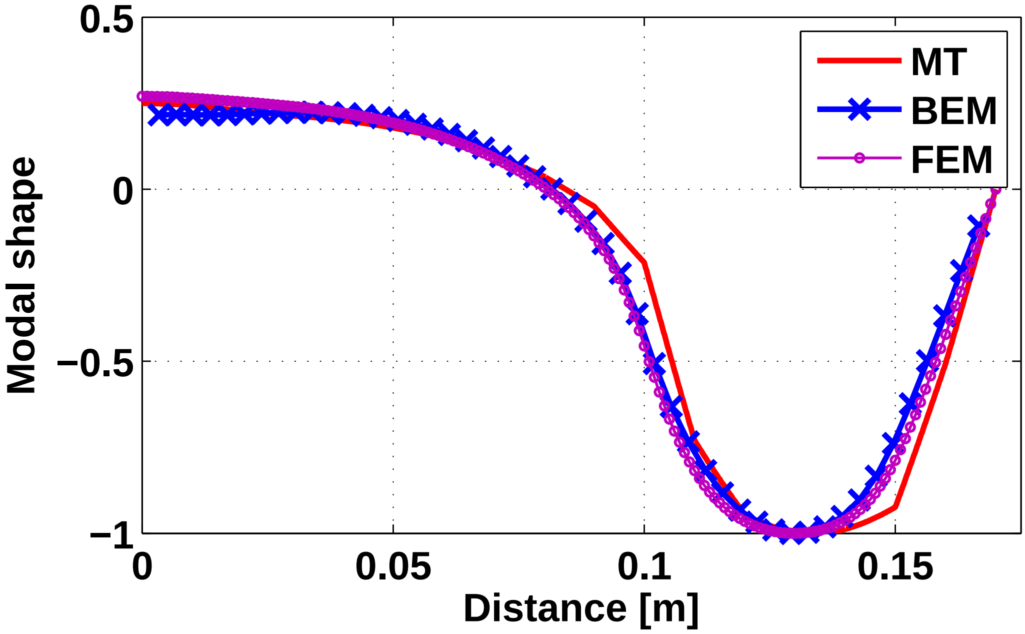


Figure 11: Second modal shape comparison between different solutions for vowel /u/

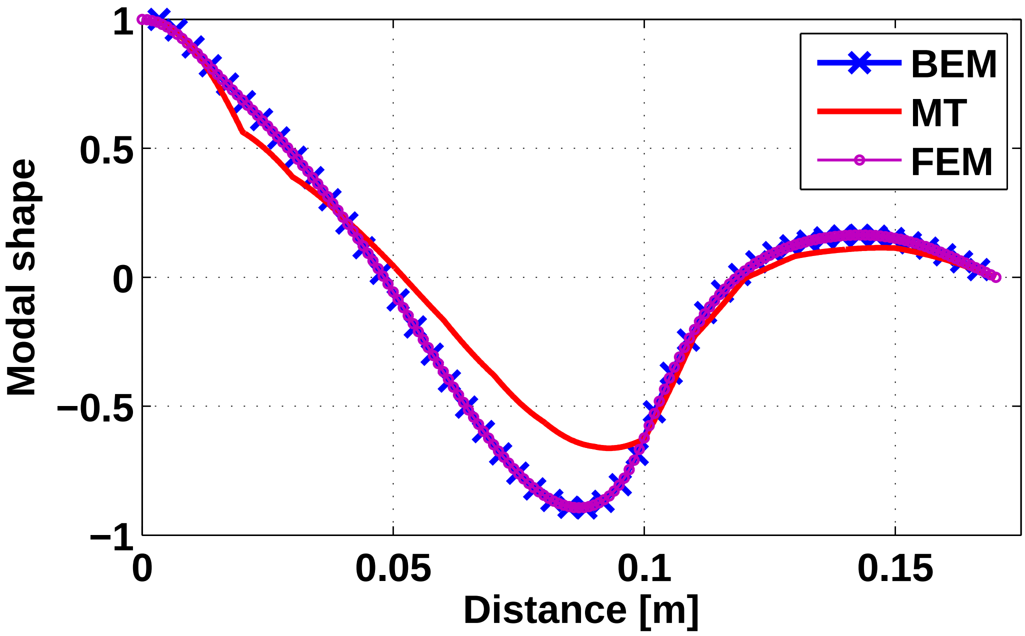


Figure 12: Third modal shape comparison between different solutions for vowel /u/

4 CONCLUSION

A vocal tract model based on anatomical data producing two different vowels was created using the boundary element method (BEM), finite element method (FEM) and the analytic solution of the transfer matrix (MT). The analysis was made in the frequency domain. The first three resonance frequencies and corresponding mode shapes were determined for each model.

A comparison was carried out and three dimensional effects were identified. BEM and FEM models showed good agreement. Even though BEM model was obtained with a rough mesh than FEM model. MT solution was also in agreement with the numeric models. But MT method lacked three dimensional effects which can be observed by report to modal shape for BEM and FEM models.

This is an introductory work in voice acoustics modeling using numeric methods. There are still many unanswered questions such as the relationship between the dimensions of vocal tract and resonance frequencies. This work provides a benchmark example and the necessary tools to tackle new applications of vocal tract acoustics modelling.

Future work may include voice synthesis and vocal pathology identification based on numeric models of the vocal tract.

ACKNOWLEDGEMENTS

The authors would like to thank the CarAcousII project (FAP-DF nº193.000.015/2012) for the financial support. The first author would also like to thank CNPq for the research scholarship granted.

References

Cataldo, E., Sampaio, R., & Nicolato, L. (2004). Uma discussao sobre modelos mecanicos de laringe para sintese de vogais. *ENGEVISTA*, *6(1)*, 47-57.

Clement, P., Hans, S., Hartl, D. M., Maeda, S., Vaissiere, J., & Brasnu, D. (2007). Vocal tract area function for vowels using three-dimensional magnetic resonance imaging. a preliminary study. *Journal of Voice*, *21(5)*, 522 - 530.

Dominguez, J. (1993). *Boundary elements in dynamics.* Av. Reina Mercedes, s/n, 41012 Seville, Spain: Escuela Superior de Ingenieros Industriales Universidad de Sevilla.

Echternach, M., Sundberg, J., Markl, T. B. M., & Richter, B. (2011). Vocal tract area functions and formant frequencies in opera tenors modal and falsetto registers. *journal of the acoustical society of america*, *129(6)*, -.

Ekholm, E., Papagiannis, G. C., & Chagnon, F. P. (1998). Relating objective measurements to expert evaluation voice quality in western classical singing" critical perceptual parameters. *Journal of Voice*, *12(2)*, -.

Guzman, M., Laukkanen, A.-M., Krupa, P., Horacek, J., Svec, J. G., & Geneid, A. (2013). Vocal tract and glottal function during and after vocal exercising with resonance tube and straw. *Journal of Voice*, *27(4)*, 523.e19 - 523.e34.

Hannukainen, A., Lukkari, T., Malinen, J., & Palo, P. (2007). Vowel formants from the wave equation. *Journal of the Acoustical Society of America*, *122(1)*, EL 1 - 7.

Henrich, N., Smith, J., & Wolfe, J. (2011). Vocal tract resonances in singing: Strategies used by sopranos, altos, tenors, and baritones. *journal of the acoustical society of america*, *129(2)*, -.

Kagawa, Y., Shimoyama, R., Yamabuchi, T., Murai, T., & Takarada, K. (1992). Boundary element models of the vocal tract and radiation field and their response characteristics. *Journal of Sound and Vibration*, *157(3)*, 385-403.

Kirkup, S. M. (2007). *The boundary element method in acoustics.* Integrated Sound Software.

Story, B. H., Laukkanen, A.-M., & Titze, I. R. (2000). Acoustic impedance of an artificially lengthened and constricted vocal tract. *Journal of Voice*, *14(4)*, 455-469.

Takemoto, H., Parham, & Mokhtari. (2010). Acoustic analysis of the vocal tract during vowel production by finite-difference time-domain method. *Journal of the Acoustical Society of America*, *128(6)*, 3724-3738.

Titze, I. (2001). Acoustic interpretation of resonant voice. *Journal of Voice*, *15(4)*, 519-528.

Titze, I. R., Mapes, S., & Story, B. (1994). Acoustics of the tenor high voice. *journal of the acoustical society of america*, *95(2)*, -.

Tom, K., Titze, I. R., Hoffman, E. A., & Story, B. H. (2001). Three-dimensional vocal tract imaging and formant structure: Varying vocal register, pitch, and loudness. *journal of the acoustical society of america*, *109(2)*, -.

Wrobel, L. C. (2001). *The boundary element method - applications in thermo-fluids and acoustics.* John Wiley & Sons.