

THREE-DIMENSIONAL MEASUREMENT OF THE VOCAL TRACT BY MRI

Didier Demolin, Thierry Metens and Alain Soquet

Université Libre de Bruxelles
Service de Linguistique Générale
Unité de Résonance Magnétique de l'Hôpital Erasme
Institut de Phonétique et de langues vivantes

email: ddemoli@ulb.ac.be

ABSTRACT

Recent studies have shown that MRI techniques are reliable to measure mid-sagittal cuts of the vocal tract. Three-dimensional images of the vocal tract have been proposed e.g. by Foldvick et al. [6], [7] but despite this advance, these images do not offer any accurate measurement of sagittal cuts or area functions. Advance in MRI performances allows to collect data much faster and with a better accuracy. In this paper we propose to use this new technology to make measurement of sagittal, coronal, coronal oblique and transversal cuts. The data collected allow accurate measurement of area at different points along the vocal tract. The different cuts obtained by this method are then used to make measurements of the vocal tract shape. The study focuses on French oral vowels.

1. INTRODUCTION

Accurate measurement of the vocal tract shape during phonation is a fundamental part in the study of speech production. In recent years magnetic resonance techniques were used to measure mid-sagittal images of the vocal tract and to obtain area functions from these data, Demolin et al. [3], [4], [5], Moore [11]. This technique has also allowed to obtain 3 dimensional VT shapes, Rokkaku et al. [12], Matsamura and Sugira [10], Baer et al. [1], [2], Matsumura [9], Yang and Kasuya [13]. However, if a true 3D model based on MRI data is theoretically possible, there are still several important problems that remain to be solved to reach this goal. In this paper we are presenting data obtained in the coronal, coronal oblique and transversal plans, while speakers were pronouncing one vowel without interruption. The 14 areas obtained in this way can then be placed on the flux line of the vocal tract. This flux line is determined from a mid-sagittal cut obtained by a method determined in Lecuit [8].

2. METHOD

2.1. MRI parameters

The imaging was performed on a 1.5 T MRI system with a quadrature Head-Neck coil (Philips Gyroscan NT ACS, Best The Netherlands). The images consisted of 3 stacks of 14 slices (5 coronal, 3 coronal oblique, 6 transversal.)

This proton density weighted TSE acquisition took 13.8 seconds, with the following parameters: TR=1716 ms, TE=9 ms, gradient profile low-high, ETL=11; Partial Fourier encoding: 60%. Field of

view: 250 x 200 mm, Matrix 161 x 256. The slice thickness was 4mm and a gap of 5mm was left between adjacent slices of each stack. A double presaturation band was applied to avoid foldover artifacts. The acquisition was launched immediately after the beginning of the phonation. The acquisition was planned on a sagittal survey image realized during a similar phonation, with similar acquisition parameters.

2.2. Subjects and composition of the data

This study was carried out on French vowels spoken by four different subjects (two female and two male) living in Brussels. Subjects were asked to pronounce and to sustain the 10 French oral vowels [i, e, ε, ə, y, ø, œ, ɔ, o, u], for 14s. In the present study nasal vowels were not taken into account.

Before the main exam of the study, using the same MRI parameters, a stack of 11 cuts was taken in the sagittal plan. The medio-sagittal cut was used to posit the cuts of the 3 different stacks examined in the recording session. The 6 transversal cuts cover the middle floor of the larynx, the larynx vestibule and the inferior part the pharynx (low pharynx for the male subjects, low and mid-pharynx for the female subjects). The 3 coronal oblique cuts cover the oro-pharynx and the velum regions. Finally, the 5 coronal cuts cover the lips, alveolar and hard-palate regions. Before recording, the three stacks, have been repositioned for each vowel. The latter point is important since the position of the articulators varies for each vowel uttered.

3. ANALYSIS

The examination of the areas obtained for every vowel in the experiment show some interesting facts. Even if it is difficult to compare male and female subjects, mainly because of vocal tract length differences, some regularities are observable between the 4 subjects of the study. First we have to mention that the accurate localization of the vocal folds from a mid-sagittal cut is not possible and therefore the measure of the length of the vocal tract is approximate. However, the first cut (SL 1 on figure 2 and 3) was posited in such a way to be situated in the middle floor of the larynx. Figure 2 shows that SL 1 is almost situated on the vocal folds, while figure 3 shows that SL 1 is situated slightly above the vocal folds. SL 2 and SL 3 of the transversal stack are generally situated in the larynx vestibule and show areas just above the glottis. The inferior part of the pharynx (low and mid pharynx) found on SL 4 to 6 shows that the pharynx is more open for high vowels. A comparison between vowels shows the following order for pharyngeal width:

i>e>ɛ>a for front unrounded vowels; y>ø>œ for front vowels and u>o>ɔ for back vowels. There is one exception in the data where SL 4 of [ø] is bigger than SL 4 of [y]. In figure 1, SL 4 and 5 show the lower and upper parts of the epiglottis which is not stuck to the tongue root.

SL 7 to 9 are situated in the oro-pharynx and velum regions. This is the constriction zone for the back vowels. The areas found in this region show that the tongue exhibit a depression in the sagittal plan. The depth of this depression is very important for the front vowels and less important for the back vowels, but it is always present. In the case of the high, close, front vowels this sagittal depression can account for at least one 4th of the total area. Consequently these data show that the shape of the tongue in the bend of the vocal tract must be taken into account for the calculus of sagittal to area function.

SL 10, 11, 12 are situated in the hard-palate region. This is the constriction zone for the front vowels. SL 13 is situated in the alveolar region and SL 14 shows areas for the lips. A comparison between rounded vowels shows that, as expected, lips areas are smaller for high vowels than for more open vowels. The order is the following for front rounded vowels: y<ø<œ and u<o<ɔ for back vowels but only for the male subject. The female subject shows a different order in lip rounding i.e. ø<y<œ and o<u<ɔ.

All the areas obtained in the experiment were measured by the following method. First, the outline of the cross section is digitized with a graphic pad, and then, the area is computed by means of numerical integration. Results are shown in table 1 and 2

Table 1: Areas of the 14 sections (SL) for the male subject given in cm².

SL	[i]	[e]	[ɛ]	[a]	[y]	[ø]	[œ]	[ɔ]	[o]	[u]
1	1.80	1.38	1.49	1.13	2.33	1.79	1.38	1.49	4.59	1.96
2	1.67	1.40	1.52	1.04	5.68	4.71	4.44	4.65	4.79	5.37
3	5.37	4.55	4.24	3.66	5.31	3.24	3.85	2.27	3.16	4.53
4	5.59	3.53	3.20	1.75	4.47	2.11	1.58	1.11	1.75	3.50
5	4.61	2.64	1.31	0.67	4.12	1.54	0.76	1.12	1.46	3.31
6	5.02	2.45	1.52	0.46	4.47	1.73	0.96	0.65	0.80	1.87
7	6.32	4.22	3.47	2.00	5.51	3.49	2.99	2.84	2.71	1.92
8	3.94	2.51	2.62	2.23	3.75	2.84	2.46	2.66	2.60	1.01
9	3.04	2.60	2.78	2.07	3.58	3.25	3.35	4.56	4.22	1.43
10	0.60	1.15	3.01	4.56	1.32	3.20	4.40	8.65	6.56	2.79
11	0.45	1.46	3.23	4.88	1.69	3.13	4.18	7.42	6.63	4.34
12	0.70	1.75	2.81	3.87	1.05	2.57	3.21	4.80	4.58	9.68
13	0.58	0.76	1.78	2.58	2.77	2.27	2.55	2.85	3.50	1.96
14	1.72	1.82	3.10	3.29	0.54	0.57	1.09	2.57	0.49	0.39

4. DISCUSSION

From the data obtained by both the MRI and the measurement of the areas, it is theoretically possible to try to reconstruct the vocal tract shape. However, several important problems remain to be taken into account.

Table 2: Areas of the 14 sections (SL) for the female subject given in cm².

SL	[i]	[e]	[ɛ]	[a]	[y]	[ø]	[œ]	[ɔ]	[o]	[u]
1	0.36	1.13	0.91	0.93	1.17	0.60	0.53	0.43	0.49	0.39
2	0.47	0.42	0.42	0.49	0.68	0.59	0.47	2.05	0.57	0.51
3	2.87	0.47	0.44	0.46	0.85	2.50	2.48	2.05	2.23	2.77
4	4.23	2.17	1.83	1.18	2.89	3.27	1.96	2.94	3.04	3.15
5	4.27	2.54	1.41	0.86	4.15	2.99	2.05	1.51	2.46	5.02
6	5.06	2.90	2.19	0.92	4.20	2.40	2.00	0.99	1.77	5.00
7	4.80	4.10	3.08	1.53	5.98	1.48	2.92	1.65	1.21	2.58
8	4.45	2.75	2.24	0.71	4.62	1.81	3.13	1.06	1.35	2.17
9	3.20	2.92	3.03	1.51	5.01	1.15	4.67	3.19	0.97	1.16
10	1.15	1.02	1.64	1.28	1.86	0.50	3.56	3.32	0.44	0.29
11	0.28	0.71	1.38	2.61	0.64	1.89	4.04	8.24	2.70	1.51
12	0.30	1.12	2.08	4.05	0.72	3.40	6.00	8.15	5.34	3.54
13	0.27	1.19	2.68	3.16	0.50	2.45	4.01	2.88	2.48	2.87
14	4.16	2.88	3.33	2.54	0.70	0.51	1.41	2.03	0.59	0.70

First, there is a limitation in the number of cuts that can be taken in the stacks because the information is degraded or lost at the intersection between two different cuts. This phenomenon is illustrated at figure 1, where darkened lines represent the intersections between the coronal cuts with the coronal oblique cut (SL 9) in the anterior part of the face.

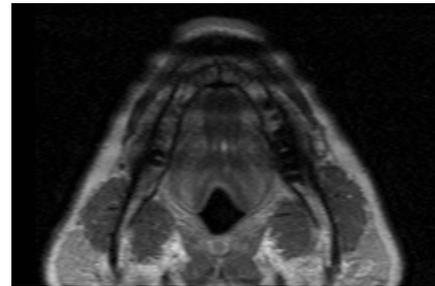


Figure 1: Intersections of coronal cuts (SL 10-14) with coronal oblique cut (SL 9) showed by the darkened lines (vowel [i] pronounced by the male subject).

Another problem is the determination of the air-tissue boundaries. Matsumura et al. [10] have proposed a computer algorithm to trace the boundaries of the vocal tract and the nasal cavity but some problems remains to be solved. For example, in our data we have noticed that there is a problem to make the difference between air and teeth in the hard palate and alveolar regions (SL 11-13 on figures 2 and 3) since teeth appear in dark because of the lack of mobile hydrogen. The visualization of the dental crown has still to be done but no reliable method has been found yet.

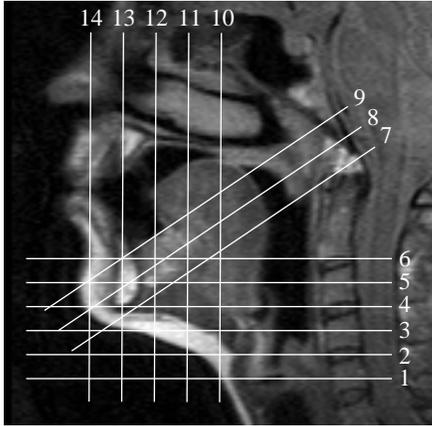


Figure 2: Sagittal cut of the vowel [u] (female subject).

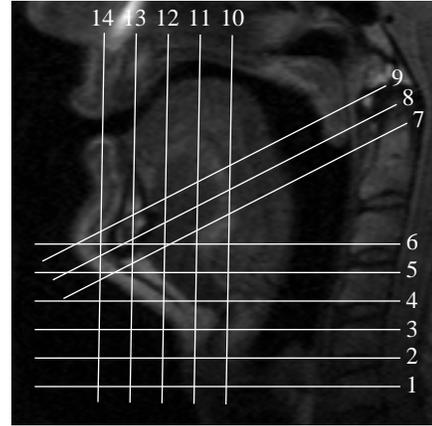


Figure 3: Sagittal cut of the vowel [i] (male subject).

Table 3: Cuts 1 to 14 of vowel [u] (female subject).

SL 14		SL 7	
SL 13		SL 6	
SL 12		SL 5	
SL 11		SL 4	
SL 10		SL 3	
SL 9		SL 2	
SL 8		SL 1	

Table 4: Cuts 1 to 14 of vowel [i] (male subject).

SL 14		SL 7	
SL 13		SL 6	
SL 12		SL 5	
SL 11		SL 4	
SL 10		SL 3	
SL 9		SL 2	
SL 8		SL 1	

Nevertheless, 3D reconstruction of the vocal tract is not the only aim of MRI. Indeed, analysis of the data allows direct and accurate visualization of the position and shape of the different articulators. Such studies have been done for the velum (Demolin et al. [5]), and for the position and shape of the tongue during the production of vowels.

5. CONCLUSION

In this paper we have made measurements of the vocal tract from areas obtained by MRI in the coronal, coronal oblique and transversal plans. The areas obtained, while speakers were uttering French vowels, have allowed us to measure the shape of the vocal tract at different places from the glottis to the lips. These data offer some promising perspectives for the making of true 3D measures of the vocal tract, but before this some basic problems in the interpretation of MRI data remain to be solved. Such problems involve the visualization of the teeth and the identification of air-tissue boundaries.

6. REFERENCES

1. Baer T., J. C. Core, L. C. Gracco and P. W. Nye. Analysis of vocal tract shape and dimensions using magnetic resonance imaging: vowels. *J. Acoust. Soc. Am.* 90, 2, 799-828, 1991.
2. Baer T., J. C. Gore, S. Boyce and P. W. Nye. Application of MRI to the analysis of speech production. *Magnetic Resonance Imaging*, 5, 1-7, 1987.
3. Demolin D. and Segebarth, C. Analyse de production de voyelles de quelques langues du Soudan central par IRM." *Actes des XIXèmes Journées d'Etude de la Parole*, Université Libre de Bruxelles. 37-42, 1992.
4. Demolin D., J-M Hombert, V. Lecuit, C. Segebarth and A. Soquet. An MRI study of French vowels. *Eurospeech*, Madrid. 2235-2238, 1995.
5. Demolin D., M. George, V. Lecuit, T. Metens and A. Soquet. Détermination par IRM de l'ouverture du velum des les voyelles nasales du français. *Actes des XXIèmes JEP*. Avignon, 1996.
6. Foldvick A.K., O. Husby, J. Kvaerness, I.C. Norli and P.A. Rinck. MRI (magnetic resonance imaging) for filming articulatory movements. *Proceedings of the 12th ICPhS*. Aix-en-Provence. 34-36, 1991.
7. Foldvick, A.K., U. Kristiansen and J. Kværness. Three-dimensional ultrasound and magnetic resonance imaging: a new dimension in phonetic research. *Proceedings ICPhS 95 Stockholm.*, Vol. 4, 46-49, 1995.
8. Lecuit, V. "Sagittal Cut to Area Function Transformations: A Comparative Study", Mémoire, Université Libre de Bruxelles. (Contact: vlecuit@ulb.ac.be).
9. Matsumura M. Measurement of three dimensional shapes of Vocal Tract and nasal cavity using magnetic resonance imaging technique. *Proc. ICSLP 92*. 779-782, 1992.

10. Matsumura M. and A. Sugira. Modeling of 3-dimensional Vocal tract Shapes Obtained by Magnetic Resonance Imaging for Speech Synthesis. *Proc. ICSLP*. 425-428, 1990.
11. Moore C. A. The correspondance of vocal tract resonance with volumes obtained from magnetic resonance imaging. *Journal of Speech and Hearing Research*, 35. 1009-1023, 1992.
12. Rokkaku M. S. Imaizumi, S. Niimi and S. Kiritani. Measurement of the three dimensional shape of the vocal tract on the magnetic resonance imaging technique. *Ann. Bull. RILP 20*. 47-54, 1986.
13. Yang C.S. and H. Kasuya. Accurate Measurement of Vocal Tract Shapes from Magnetic Resonance Images of Child, Female and Male subjects. *Proc. ICSLP 94*. 623-626, 1994.

7. MORE DATA

SL	Female subject (vowel [u])	Male subject (vowel [i])
1	[IMAGE A766G01.GIF]	[IMAGE A766G15.GIF]
2	[IMAGE A766G02.GIF]	[IMAGE A766G16.GIF]
3	[IMAGE A766G03.GIF]	[IMAGE A766G17.GIF]
4	[IMAGE A766G04.GIF]	[IMAGE A766G18.GIF]
5	[IMAGE A766G05.GIF]	[IMAGE A766G19.GIF]
6	[IMAGE A766G06.GIF]	[IMAGE A766G20.GIF]
7	[IMAGE A766G07.GIF]	[IMAGE A766G21.GIF]
8	[IMAGE A766G08.GIF]	[IMAGE A766G22.GIF]
9	[IMAGE A766G09.GIF]	[IMAGE A766G23.GIF]
10	[IMAGE A766G10.GIF]	[IMAGE A766G24.GIF]
11	[IMAGE A766G11.GIF]	[IMAGE A766G25.GIF]
12	[IMAGE A766G12.GIF]	[IMAGE A766G26.GIF]
13	[IMAGE A766G13.GIF]	[IMAGE A766G27.GIF]
14	[IMAGE A766G14.GIF]	[IMAGE A766G28.GIF]