FROM SAGITTAL CUT TO AREA FUNCTION:
AN RMI INVESTIGATION

Alain Soquet, Véronique Lecuit, Thierry Metens, and Didier Demolin
Université Libre de Bruxelles, CP 110,
50 av. F.–D. Roosevelt
B–1050 Brussels, Belgium
E-mail: asoquet@ulb.ac.be

ABSTRACT
This paper presents a comparative study of transformations used to compute the area of a cross sections of the vocal tract from the mid-sagittal diameter of the vocal tract. MRI techniques have been used to obtain cross sections of the vocal tract in a study of French oral vowels uttered by two subjects. The measured cross sectional areas are compared to the cross sectional areas computed by the different transformations.

1. INTRODUCTION
The generation of area functions from measurements of the sagittal section is an important step in the study of the relation between vocal tract geometry and speech acoustics. Many authors have proposed transformations performing this particular task, in this study we used the transformations designed by Fant [1], Maeda [4], Perrier et al.[5], and Sundberg [6]. In previous work, we carried out a comparative study of these transformations [3]. The data we used was a set of sagittal cuts obtained by means of X-rays. The outcome of this study is that the choice of a particular transformation has important influence both on vocal tract area function and on acoustic cues.

However, this first study had a major drawback: the comparison of the different transformations could only take place in the acoustic space: indeed no information on the actual area function was available. To overcome this problem, we gathered information about the area of several sections of the vocal tract by means of an MRI investigation. We then compared the area measured on the MRI data to the estimated area computed by means of these transformations on the basis of the measured mid-sagittal distance.

2. MRI DATA
This study focused on the set of ten French oral vowels, uttered by a male and a female subject.

For each vowel, we performed 14 MRI scans of 4 mm thickness distributed along the vocal tract and perpendicular to the axis of the tract. On these scans, it is possible to measure both the mid-sagittal distance and the cross sectional area. Hence, the transformations can then be applied on the mid-sagittal dimension and the result can be compared to the actual cross sectional area. Figure 1 shows the position of the 14 scans on the mid-sagittal profile for the vowel [u] uttered by the female subject along with a sample of three cross sections. The area and the mid-sagittal distance can easily be measured on these cross sections.

3. SAGITTAL CUT TO AREA FUNCTION TRANSFORMATIONS
Transformations to go from the mid-sagittal dimension to the cross sectional area are based upon the original transformation designed by Heinz and Stevens in 1965 [2], which is: \( A = \alpha d^\beta \), where \( d \) is the mid-sagittal dimension, \( A \) is the cross sectional area, and \( \alpha \) and \( \beta \) are coefficients specified by the authors.

These transformations rely upon an accurate distinction of the different regions of the vocal tract. Indeed, the \( \alpha \) and \( \beta \) are defined for a particular region. Hence, we divided the vocal tract into nine regions: middle floor of larynx, larynx, low pharynx, mid pharynx, oro-pharynx, velum, hard palate, alveolar zone, and labial zone. Figure 2 represents these different regions on a mid-sagittal contour of the vocal tract. This division into nine regions allows accurate implementation of the transformations.

Figure 1: Position of the 14 scans on the mid-sagittal profile for the vowel [u] uttered by the female subject along with a sample of three cross sections.
3.1. Transformations evaluation

For each section of the vocal tract the performances of the transformations were compared. Figure 3 and figure 4 show scatter plots displaying the area measured on MRI data versus the area computed by one transformation for the ten vowels uttered by the male subject and female subject respectively. Each symbol represents a region of the vocal tract, starting from the larynx, (region 2) up to the lips (region 9). A line of slope equal to one is added as a reference. Many observations from our first study were confirmed. There are significant differences between the transformations. Some transformations keep overestimating the size of large sections, while others give rise to flattened area functions.

Aside from Perrier, most authors use the same coefficients for extended regions of the vocal tract. Figure 5 and figure 6 show scatter plots displaying the sagittal distance versus the cross section measured on IRM data for the male and female subject respectively. These plots gather data for three adjacent regions, Maeda’s transformation is used in regions 1 to 3, Fant’s in regions 3 to 5, and Sundberg’s in regions 6 to 8. These plots demonstrate that these transformations do not properly render the geometry of the vocal tract of our subjects.

We pursued our analysis by testing if we could optimize the transformation parameters so as to match each subject of the corpus. A power curve fit was applied to the measured data. Given the particular shape of the vocal tract, it is necessary to define new coefficients for each subject and each of the nine regions. Table 1 displays the coefficients derived from the MRI data for both subjects, the correlation coefficients from the curve fit are also given. As can be seen on this results, transformations can only take into account the main trends of a relation between the area and the mid-sagittal distance, even on clean data and with a large number of different regions.

4. CONCLUSION

The data provided by MRI proved to be of considerable interest in this study. The availability of information about the shape of the tract allowed us to compare the transformations upon the basis of a reliable reference. These transformations were designed through the study of a few subjects and with imaging techniques less elaborate than the MRI techniques used today. The use of these transformations on other subjects than the original ones is inappropriate and may lead to errors.

5. REFERENCES

3. V. Lecuit, “Sagittal Cut to Area Function Transformations: A Comparative Study”, Mémoire, Université Libre de Bruxelles. (Contact: vlecuit@ulb.ac.be).

Table 1: Coefficients derived from the MRI data for both subjects

<table>
<thead>
<tr>
<th>Zone</th>
<th>Male subject</th>
<th></th>
<th></th>
<th>Female subject</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>α</td>
<td>β</td>
<td>R</td>
<td>α</td>
<td>β</td>
<td>R</td>
</tr>
<tr>
<td>Larynx floor</td>
<td>1.563</td>
<td>0.040</td>
<td>0.031</td>
<td>0.461</td>
<td>1.112</td>
<td>0.784</td>
</tr>
<tr>
<td>Larynx</td>
<td>1.515</td>
<td>1.690</td>
<td>0.829</td>
<td>0.776</td>
<td>0.702</td>
<td>0.521</td>
</tr>
<tr>
<td>Low pharynx</td>
<td>1.824</td>
<td>1.381</td>
<td>0.957</td>
<td>1.857</td>
<td>1.224</td>
<td>0.974</td>
</tr>
<tr>
<td>Mid pharynx</td>
<td>1.347</td>
<td>1.626</td>
<td>0.973</td>
<td>1.739</td>
<td>1.225</td>
<td>0.982</td>
</tr>
<tr>
<td>Oro-pharynx</td>
<td>0.731</td>
<td>1.806</td>
<td>0.989</td>
<td>2.001</td>
<td>0.806</td>
<td>0.913</td>
</tr>
<tr>
<td>Velum</td>
<td>1.393</td>
<td>1.079</td>
<td>0.874</td>
<td>1.839</td>
<td>0.924</td>
<td>0.873</td>
</tr>
<tr>
<td>Hard Palate</td>
<td>1.334</td>
<td>1.514</td>
<td>0.969</td>
<td>1.816</td>
<td>1.431</td>
<td>0.964</td>
</tr>
<tr>
<td>Alveolar</td>
<td>1.875</td>
<td>1.193</td>
<td>0.955</td>
<td>2.667</td>
<td>1.465</td>
<td>0.799</td>
</tr>
<tr>
<td>Lips</td>
<td>4.697</td>
<td>2.481</td>
<td>0.900</td>
<td>2.416</td>
<td>1.661</td>
<td>0.835</td>
</tr>
</tbody>
</table>
Figure 3: Scatter plots displaying the area measured on MRI data versus the area computed by one transformation for the ten vowels uttered by the male subject. Each symbol represents a region of the vocal tract, starting from the larynx, region 2, to end up with the lips, region 9. A line of slope equal to one is added as a reference.

Figure 4: Scatter plots displaying the area measured on MRI data versus the area computed by one transformation for the ten vowels uttered by the female subject. Each symbol represents a region of the vocal tract, starting from the larynx, region 2, to end up with the lips, region 9. A line of slope equal to one is added as a reference.
Figure 5: Scatter plots displaying the sagittal distance versus the cross sectional area measured on MRI data from the male subject. These plots gather data for three adjacent regions, the line represents a particular transformation. Maeda’s transformation is used in regions 1 to 3, Fant’s in regions 3 to 5, and Sundberg’s in regions 6 to 8.

Figure 6: Scatter plots displaying the sagittal distance versus the cross sectional area measured on MRI data from the female subject. These plots gather data for three adjacent regions, the line represents a particular transformation. Maeda’s transformation is used in regions 1 to 3, Fant’s in regions 3 to 5, and Sundberg’s in regions 6 to 8.