Towards a 3D articulatory model of velum based on MRI and CT images

Antoine Serrurier
Pierre Badin
Institut de la Communication Parlée, UMR CNRS 5009 – INPG – Université Stendhal, Grenoble, France

This paper describes the processing of *MRI* and *CT images* needed for developing a 3D linear articulatory model of velum. The 3D surface that defines each organ constitutive of the vocal and nasal tracts is extracted from *MRI* and *CT images* recorded on a subject uttering a corpus of artificially sustained French vowels and consonants. First, the 2D contours of the organs have been manually extracted from the corresponding images, expanded into 3D contours, and aligned in a common 3D coordinate system. Then, for each organ, a generic mesh has been chosen and fitted by elastic deformation to each of the 46 3D shapes of the corpus. This has finally resulted in a set of organ surfaces sampled with the same number of 3D vertices for each articulation, which is appropriate for *Principal Component Analysis* or *linear decomposition*. The analysis of these data has uncovered two main uncorrelated articulatory degrees of freedom for the velum's movement. The associated parameters are used to control the model. We have in particular investigated the question of a possible correlation between jaw / tongue and velum's movement and have not find more correlation than the one found in the corpus.

1. Introduction

The problem of nasality is complex and has given rise to a large number of studies, from both perception and production points of view (Rossato *et al.*, 2003; Dang *et al.*, 1994; Feng *et al.,* 1996; Teixeira *et al.*, 2000; Huffman *et al.*, 1993). The nasality feature is related to the velum position: lowering the velum, and thus opening the velopharyngeal port, is a simple gesture that induces strong and complex changes in the vocal tract acoustical behaviour. The realisation of nasality involves (1) an articulatory level that deals with the shape of the

articulators and their articulatory degrees of freedom, and (2) a control level that deals with the coordination of these articulators.

The present article describes our first attempts to acquire 3D data of the complex geometry of the articulators and cavities involved in nasality (velum, nasal passages, velopharyngeal port, paranasal sinuses) for developing a three-dimensional articulatory model based on one specific subject. A number of reasons motivate this work:

- The uvula – an appendix of the velum in the midsagittal region – is often in contact with both the back of the tongue and the pharyngeal wall, creating an occlusion in the midsagittal plane: lateral channels can however remain open and should thus be taken into account by a 3D geometry description.
- The complex muscular structure of this region, in particular the interspersion between muscles from the velum, the nasopharynx and the tongue, and the *Passavant's pad* made of the fusion of the fibres of the *palatopharyngeus* muscle with those of the *pterygopharyngeal* portion of the superior constrictor (Zemlin, 1968) leads to a sphincter like behaviour (Amelot *et al.*, 2003) that controls the velopharyngeal port opening and should be considered in three dimensions.
- A 3D model offers the possibility to provide accurate area functions of the complex nasal passages and nasopharyngeal port, with the associated articulatory control parameters, that are needed for acoustical models and thus for speech synthesis
- In the framework of the development of virtual audiovisual talking heads (cf. e.g. Badin *et al.*, 2003), the 3D visualisation of the velum constitutes an interesting addition.

2. Modelling approach

2.1. Duct vs. organs

Our aim is to develop a 3D model of the various structures and organs involved in the nasopharyngeal tract. A similar approach has been successfully applied to the tongue and to the lips (Badin *et al.*, 2002). The present model will thus constitute a complement to these models, based on the same French subject.

The final result of an articulatory model is the shape of the complete vocal and nasal tracts, needed for the aerodynamic / acoustic stage of the speech production process. We could have developed a model of *tract* or *duct*, as in Badin *et al.* (1998). However, this approach is not well suited to take precisely into account the complex geometry of the various speech articulators. This is

why we have decided to develop an *organ-based* model, i.e. to model each organ separately, and to reconstruct the oral and nasal tracts subsequently.

2.2. A subject-oriented linear modelling approach

Following a method already proven for orofacial articulatory modelling (Badin *et al.*, 2002), the 3D geometry of the various non-rigid organs will be modelled as the weighted sum of a small number of linear components. These components will be extracted by linear analysis from a set of vocal and nasal tract shapes representative of the speech production capabilities of the subject. The analysis is based on both *Principal Component Analysis (PCA)* and *linear regression.* The weights of the sum constitute the *articulatory control parameters* associated with the components: a given set of values of these parameters produces a given single shape of the organs.
Two stages are necessary to develop the model: the construction of the database of shapes from images (described in section 3.), and then the analysis of the corresponding 3D coordinates (an example is detailed in section 4.)

Note that our approach is *subject-oriented*, i.e. that the model will be based on a single (French) subject. This avoids to merge the physiological characteristics and control strategy of different subjects, which can be quite different. For example, the closure of the nasopharyngeal port can be partial for some speakers while total for others. As well, compensation strategies can be different between speakers. Finally, a practical reason is the large amount of data to process. This approach can be naturally criticized as it is speaker dependent and may not generalize extensively. However, in this first approach, the model does not intend to cover all the possible aspects of nasal articulation, but to explain a possible mechanism of articulatory movements.

The corpus consisted of a set of artificially sustained articulations designed as to cover the maximal range of articulations: the French oral and nasal vowels [a ɛ e i y u o ø ɔ œ ɑ̃ ɛ̃ œ̃ ɔ̃], and the consonants [p t k f s ʃ m n ʁ l] in three symmetrical contexts [a i u]. This corpus was supplemented by two specific articulations frequent in speech: the *rest* articulation (in a rest position) and the *prephonatory* articulation (in the preparatory phase preceding phonation). Finally we have a corpus of 46 French phonemes.

As described above, the corpus contains only 46 static articulations with many more oral articulations than nasal ones. The use of a limited number of target articulations to build a model is justified by the investigation of Badin *et al.* (1998) who showed that a limited number of target articulations leads to an

accurate articulatory model. The corpus intends to cover the maximal range of the articulators' positions, and the lack of balance between orals and nasals will not change the ability of the model to reach all possible articulations.

3. Determination of the organ shapes from MR and CT images

In order to be able to develop an articulatory model following the approach described above, it is needed to obtain a 3D surface representation of the organs of the vocal and nasal tracts (jaw, tongue, nasopharynx, oropharynx, velum, paranasal sinuses, nasal passages, nostrils, lips, epiglottis and hard palate) for each articulation of the corpus, based on stacks of MR and CT images.

3.1. *Acquisition and pre-processing of the CT and MR images*

A Computer Tomography scan of the head of the subject was made, to serve as a reference. A stack of 149 axial images with a size of 512×512 pixels, a resolution of 20 pixels/cm, and an inter slice space of 0.13 cm, spanning from the neck to the top of the head, was recorded for the subject at rest (see one example image in Figure 1a). These images allow making the distinction between bones, soft tissues and air, but do not allow discriminating different soft tissues. They will be used to locate bony structures and to determine accurately their shapes for reference (see section 4.2).

Stacks of sagittal MR images were recorded for the French subject sustaining artificially during about 45 sec. each of the 46 articulations of the corpus (Figure 1b illustrates vowel [a]). The subject was instructed to artificially sustain the articulation throughout the whole acquisition time. The consonants were produced in three different symmetrical vocal contexts [VCV], V belonging to
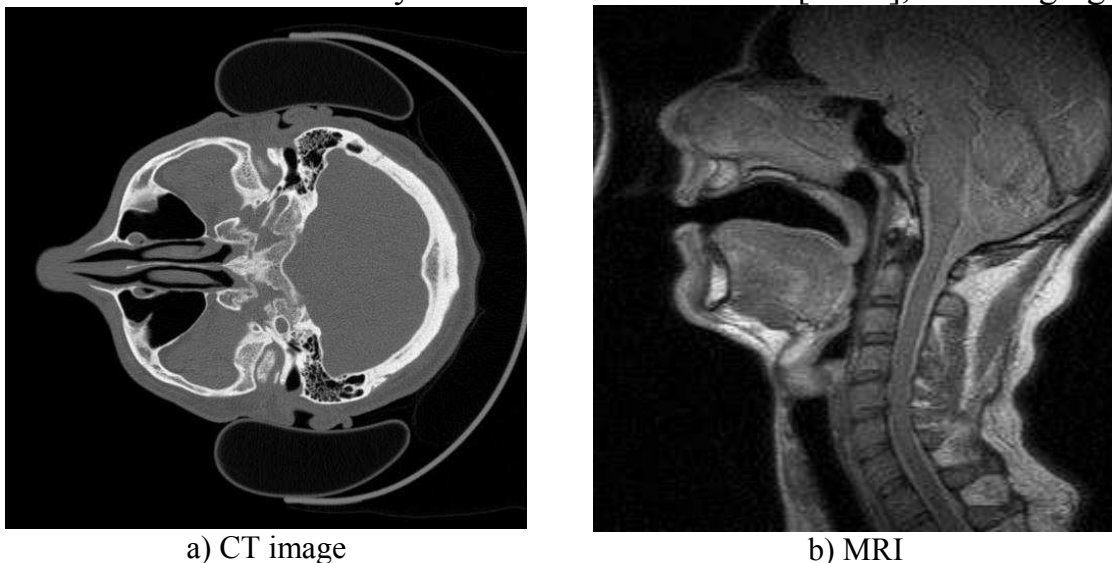

a) CT image	b) MRI

Figure 1: Original images: (a) an axial *CT image* and (b) an *MRI image* for vowel [a].

[a i u]. A set of 25 sagittal images with a size of 256×256 pixels, a resolution of 10 pixels/cm and an inter slice distance of 0.4 cm was obtained for each articulation. These images allow to make the distinction between soft tissues and air, and to discriminate soft tissues, but not to distinguish clearly the bones. Note that for both imaging techniques, the subject was in a supine position, which may alter somehow the natural shape of articulators.

Due to the complexity of the contours of the various organs, to the relatively low resolution of the images, and to the need of an accurate reconstruction of the organs, the extraction of contours has been performed manually, plane by plane. This is a rather accurate process, except for regions where the surface of a structure is tangent to the plane, and thus the tracing difficult and not accurate: this happens for instance when tracing the tongue in sagittal planes far from the midsagittal one, and nearly tangent to the tongue sides. This is why we have supplemented the initial original stacks of images in one single orientation (*axial* for CT images and *sagittal* for MR images) by extra sets of images reconstructed by intersection of the initial stack with planes having a more useful orientation, i.e. being more perpendicular to the organ surface.
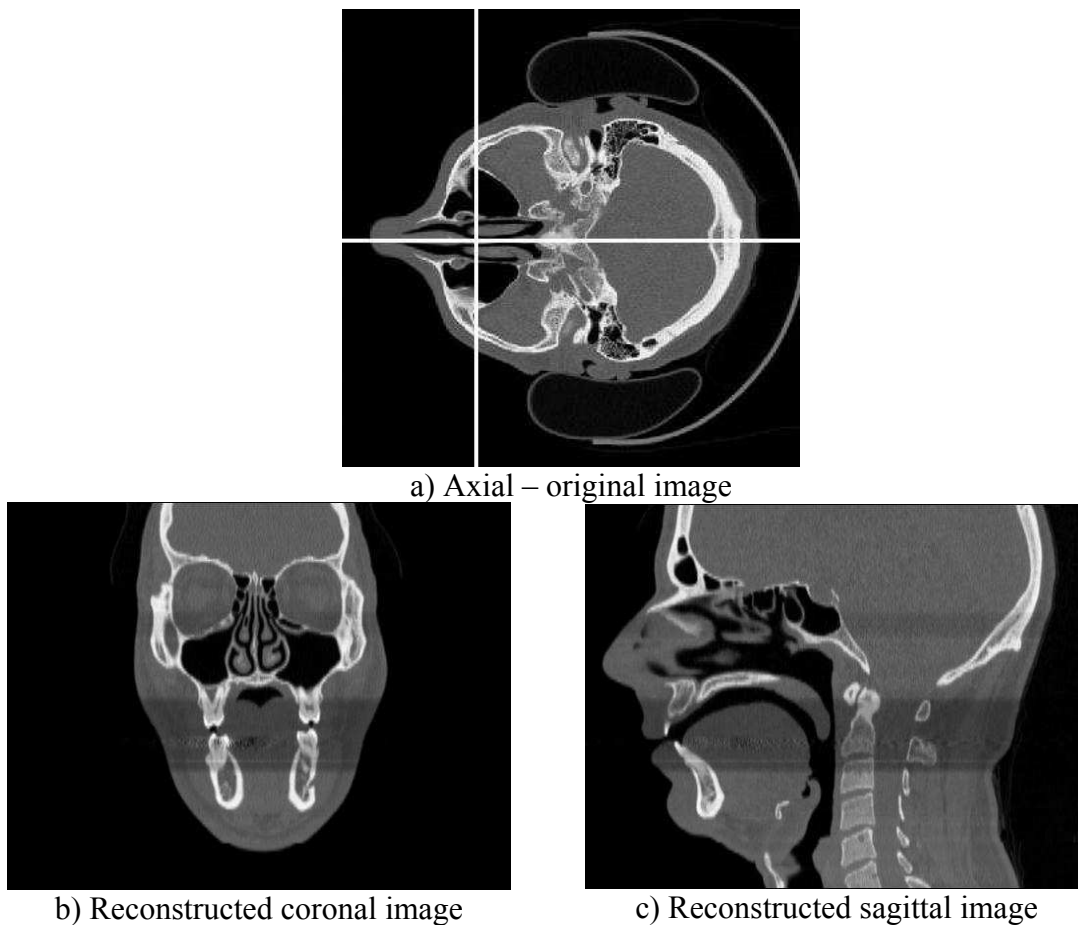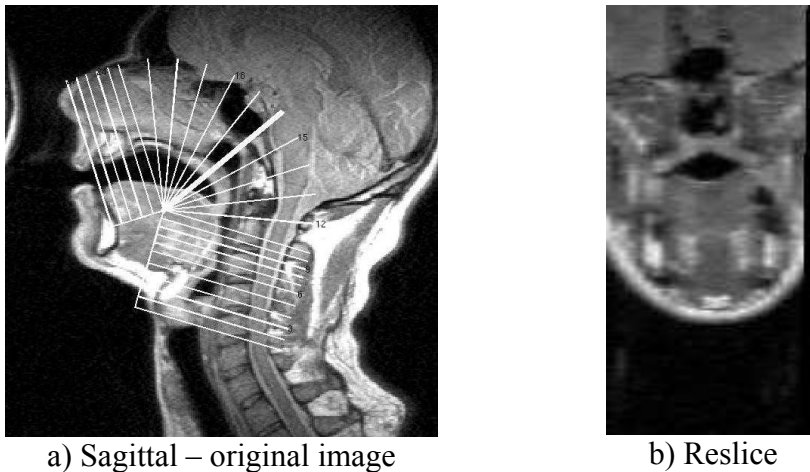
a) Axial – original image

b) Reconstructed coronal image   c) Reconstructed sagittal image
Figure 2: Example of CT images resclicing

The CT images have been thus resliced / interpolated in two stacks of 512 coronal images and 512 sagittal images (each new image has a size of 512×395 pixels, and a resolution of 20 pixels/cm), leading altogether to three stacks of perpendicular CT images that give a high resolution in the three orientations (see Figure 2 for an illustration).

For the MR images, the initial sagittal stack was resliced in images perpendicular to the vocal tract, considering that they will be used to extract organs shapes around the vocal tract (e.g. velum, tongue, etc.). They were thus resliced in 27 planes orthogonal to the midsagittal plane and intersecting it along a semipolar grid, as illustrated in Figure 3. Each new image is arbitrarily given a size of 200×100 pixels and a resolution of 10 pixels/cm. Finally we dispose of two redundant stacks of MR images for each articulation.

a) Sagittal – original image   b) Reslice
Figure 3: Example of MR images reslicing: one perpen-
dicular image for an [a] articulation

3.2.   *Determination of the rigid bony structures*

A number of structures that makes up the vocal tract can be considered as rigid: jaw, hyoid bone, hard palate, nasal passages, nostrils and various paranasal sinuses. The shape of these structures has thus been reconstructed in the following way, from the CT images:

1. Manual edition of each organ, plane by plane, in one of the three stacks, or a combination of them, depending on the form and orientation of the organ, in such a way as to maximise the accuracy for complex organs (e.g. the nasal passages were hand-edited in coronal and axial stacks). Figure 4a shows, in black, the contour of the jaw manually edited from an axial CT image.

2. Expansion of the set of all 2D plane contours into a 3D coordinate system (see Figure 4b for the jaw). These 3D points are then processed through a 3D meshing reconstruction software service provided by the Prisme Research Group at INRIA (http://cgal.inria.fr/Reconstruction) to form a 3D surface meshing based on triangles (Figure 4c).
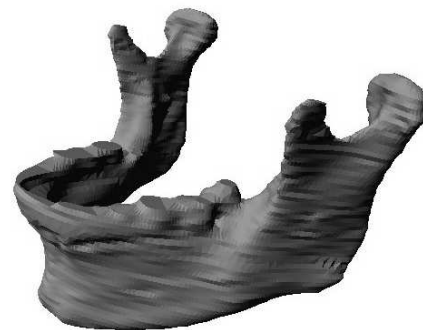
Figure 4 illustrates the 3D reconstruction process of the jaw: manual edition on plane images, and 3D meshing.
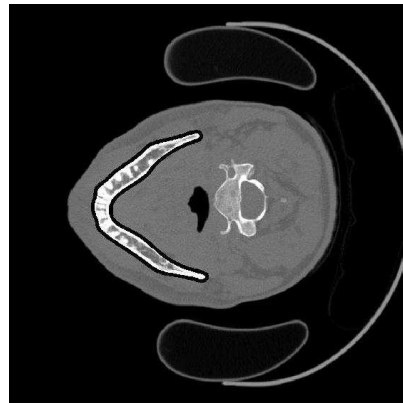
a) Manually edited contour in an axial image
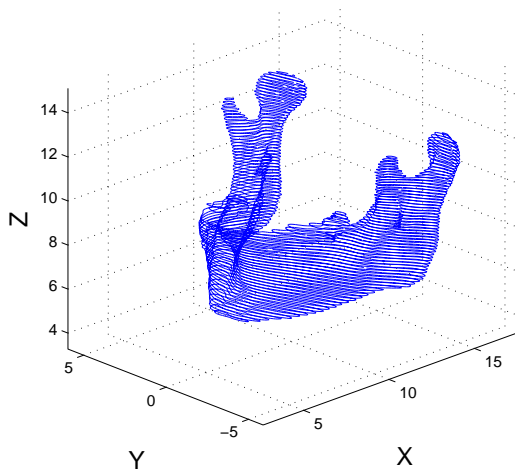
b) Set of 2D contours   c) 3D surface reconstruction

Figure 4: Illustration of the 3D reconstruction of the jaw

3.3.   Alignment of the images on a common reference

Before attempting to determine the shape of the soft structures, an important step is the alignment of the stacks of images for each of the articulations with a common reference. The process is composed of four steps:

1. We define an arbitrary common reference: the absolute 3D reference coordinate system is attached to the skull of the subject. In that way, some organs are *always* fixed in this system (hard palate, nasal passages,

paranasal sinuses, etc.). The reference coordinate system is arbitrary defined as follows (inherited from 2D reference coordinate system of Beautemps *et al.*, 2001):

- The x-axis is oriented from anterior to posterior in the midsagittal plane and approximately in the occlusal plane, the y-axis from left to right, and the z-axis from feet to head
- The point of coordinates (5, 0, 10) is arbitrarily set at the lower edge of the *Upper Incisors in the midsagittal plane*

2. We align the stacks of CT images with this reference by manually placing the hard palate shape in the reference system defined above. The geometrical alignment transformation gives the position of the three stacks of CT images in the common reference. This transformation corresponds to the six degrees of freedom of a solid object and is thus defined by 6 parameters: 3 parameters for the 3D rotation and 3 parameters for the 3D translation; it will be referred to as a (*3D*) *rototranslation*.

3. Considering that the subject may have changed its position between two MR images stacks recording, each stack must be aligned with the common reference by using an appropriate 3D rototranslation. This rototranslation is obtained by aligning the rigid structures (hard palate, nasal passages, paranasal sinuses), extracted from CT images, with each of the MR images stack. The alignment of these rigid structures with a given MR images stack is a semi-automatic process: (1) anchor points of the rigid structures are manually marked with care on some of the MR images of the stack, (2) the minimization of the cumulated distance between these 3D points and the corresponding nearest points on the 3D rigid structures provides the 3D rototranslation needed. The minimization is carried out with the MATLAB function *fminunc* (minimization without constraint). A similar approach was proposed by Takemoto *et al.* (2004), the main differences being that their minimization error was the value of the volume overlap between the reference to align and the target data.

4. The previous procedure is also applied to the jaw and the hyoid bone for each articulation. This allows to determine the relative position of these bones in relation to the fixed rigid structures: by combining this relative 3D rototranslation and the absolute one corresponding to the given stack, the positions of these two structures are known in the reference common to each articulation.

It is then possible to project the intersection of the 3D surface boundaries of the rigid structures with the plane corresponding to the MR images on these images, in order to provide some useful anchor points for the interpretation of the images
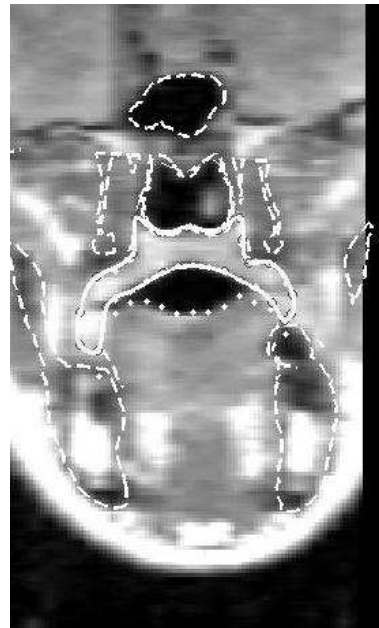
and for the tracing of the soft structure contours (Figures 5a and 5b illustrate these images).

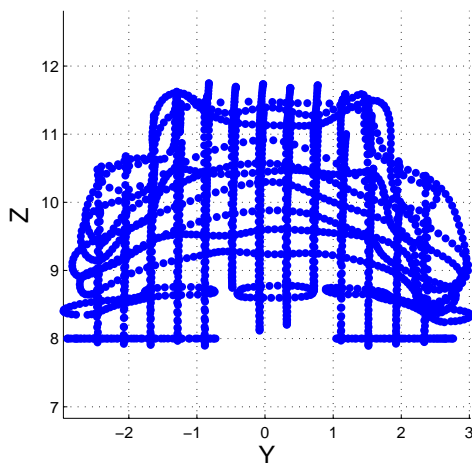### 3.4.   *Determination of the soft structures*

The determination of the soft structures (velum, nasopharynx, oropharynx, tongue, lips) is achieved in much the same way as for the rigid structures, but
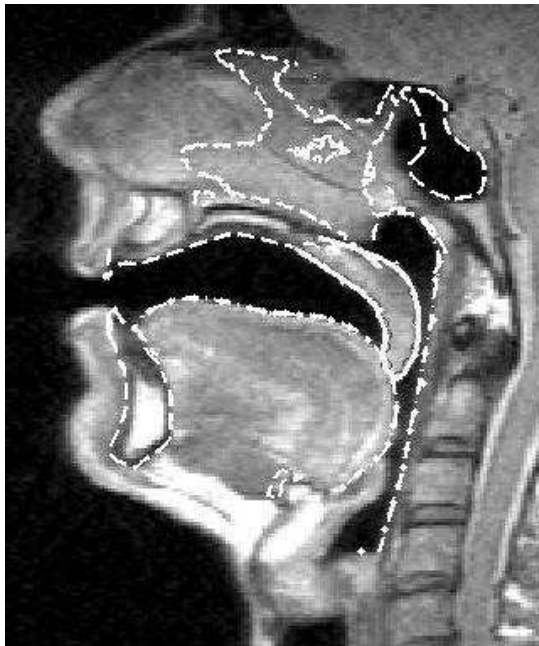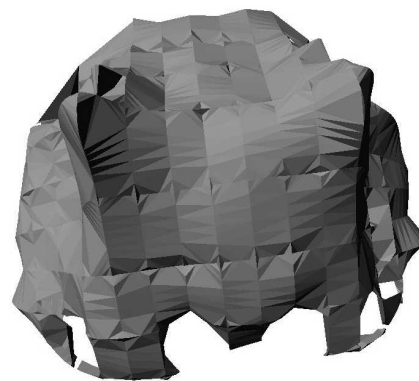

a) Edition in the sagittal plane   b) Edition in a perpendicular plane

c) Set of all 2D contours seen from front   d) 3D surface reconstruction

Figure 5: Example of manually edited contours (a and b), and of 3D reconstruction (c and d) of the velum for an [a] articulation. In a and b, the dashed lines correspond to bony structures superimposed on the images while dotted and dash-dotted lines correspond to soft structures other than the velum previously hand-edited. The solid line correspond to the velum being edited.

from the MR images of each articulation. As explained above, in order to maximise the accuracy, plane contours were edited in both stacks. The contours extracted from the rigid organs were superposed on the MR images in order to provide reference information. Figure 5 illustrates the edition of the velum in both MRI stacks for the vowel [i]. Bony structures and previously hand-edited soft structures other than the velum are superimposed on the images to help the detection of the velum (figures 5a and 5b). The velum contour (solid line on images 5a and 5b) is established as a 2D spline controlled by a limited number of points in the corresponding image plane. The set of all 2D plane contours expanded into the 3D coordinate system forms a 3D description of the given soft organ (figure 5c for a front view). The alignment with the common reference is then possible through the 3D rototranslation of the stack. An example of surface reconstruction through the 3D meshing reconstruction software service previously cited is given for the articulation [a] (figure 5d).

In order to ensure a common geometric representation of all the articulations for a given soft organ, a unique generic 3D surface mesh, made of triangles, is defined for each organ, and is fitted by elastic deformation to each of the 46 3D shapes of the corpus. This procedure provides thus a 3D sampling of each organ surface with the same 3D vertices for each of the 46 articulations of the corpus, which is appropriate for the statistical study and the linear modelling of the organs. The elastic deformation of the generic mesh to fit to each 3D shape extracted for the 46 articulations is computed through the matching software *TestRigid* developed at the TIMC laboratory in Grenoble (Couteau *et al.*, 2000). This reconstruction process finally provides a set of soft and rigid organ surfaces described in terms of triangular meshes having the same number of vertices, in a common reference coordinate system, for each of the 46 articulations of the corpus. This set of data forms the basis for the articulatory modelling of the subject, as will be illustrated in the next section.

4. Modelling of the velum

4.1. Brief anatomic description of the nasopharyngeal tract

The nasopharyngeal tract is made up of several organs: the velum, the nasopharyngeal walls, the nasal passages, the various paranasal sinuses and the nostrils. These organs can be classified into two types: rigid (or quasi rigid) structures (nasal passages, paranasal sinuses, nostrils) and non-rigid structures (velum, nasopharynx).

Rigid structures are geometrically complex. The septum separates the nasal passages. Each cavity is made up of a lot of thin and complex passages whose surface is very large. It is limited at one side by the nasopharyngeal port and by nostrils at the other one. Several paranasal sinuses (maxillary, sphenoid, etc.) are connected to these cavities through very narrow passages and thin membranes.

The velum and the nasopharynx constitute the non-rigid structures. The velum is mainly made of five muscles. It ensures the principal closure gesture of the velopharyngeal port through his major muscle, the *levator veli palatini* muscle, which stretches symmetrically from the medial region of the velum to the right and left Eustachian tubes. In addition, the velum is connected to the neighbouring organs: with the pharynx through the *palatopharyngeus* muscle, and with the tongue through the *palatoglossus* muscle (see Figure 6). The pharynx is principally active through the *superior*, *middle* and *inferior constrictors* muscles in sphincter action.
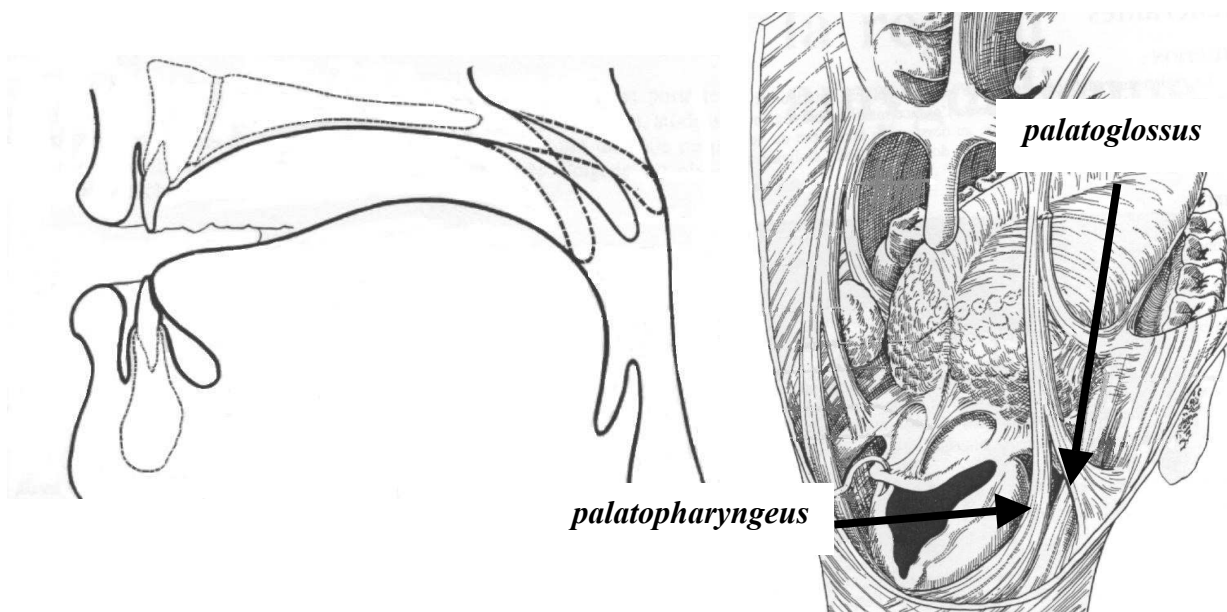
a) midsagittal view for three velum positions   b) posterior view

Figure 6: Anatomy of the velum: midsagittal view (a) and posterior view (b). Muscles linked to the tongue and to the nasopharynx (from Bouchet *et al.*, 1980)

4.2.   A first articulatory model of velum

We present in this section the results of a first attempt to establish an articulatory model of the velum based on the data obtained by the procedure described in the previous sections. The available observed data were the 3D coordinates of the 9794 vertices of the velum, of the 3369 vertices of the pharyngeal wall, and of the 4167 vertices for the tongue surface in the vicinity of the nasopharynx, for a

subset of 22 articulations taken among the 46 articulations of the full corpus: [a ɛ œ ɔ ɑ̃ ɛ̃ œ̃ ɔ̃ mᵃ mⁱ mᵘ pᵃ pⁱ pᵘ nᵃ nⁱ nᵘ tᵃ tⁱ tᵘ rest phrephonation].
Following the method described above, these generic meshes have been matched to the 22 targets extracted from the MRI data with a root mean square error of 0.12 cm for the part of the mesh located between the sagittal planes distant from ±1.5 cm from the midsagittal plane.

Our first attempt was to apply a *direct PCA* to the 22 observed velum shapes. The aim was to explain the maximum of the variance of the data by a minimum number of articulatory control parameters. A PCA was thus applied to these 29382 variables in order to extract a few articulatory control parameters by exploiting the correlations between neighbouring points and due to the physical continuity of the organs. Table 1 shows that almost 88 % of the cumulated variance of all the velum points is explained by only two articulatory control parameters, the first parameter explaining almost 75 % by itself. As explained below, the second parameter complements the closure mechanism of the velum and explains almost 13% of the full-cumulated variance of velum points. Table 1 gives the Root Mean Square error between the velum reconstructed from the parameters determined by PCA and the 22 shapes of velum: the error of reconstruction is less than 1 millimetre when two parameters are used.

Table 1: Explained variance and Root Mean Square error of reconstruction for the direct PCA model approach.

Parameter	Explained variance	Cumulated explained variance	RMS
First PCA parameter	74.9 %	74.9 %	0.11 cm
Second PCA parameter	12.9 %	87.8 %	0.07 cm

The figure 7 illustrates the shape of the velum (centre of the figure), of the tongue (partially represented in the bottom part of the figure) and of the nasopharynx (in the background) for two opposites values of each of the two parameters (±3). The action of the first parameter corresponds to a motion of the velum simultaneously along the vertical and horizontal directions. Considering its orientation and its prime importance for speech (Bell-Berti, 1993), the levator veli palatini muscle can be thought to be much involved in this movement. The second parameter is more related to an horizontal motion which complements the closure of the nasopharyngeal port by a back to front effect.
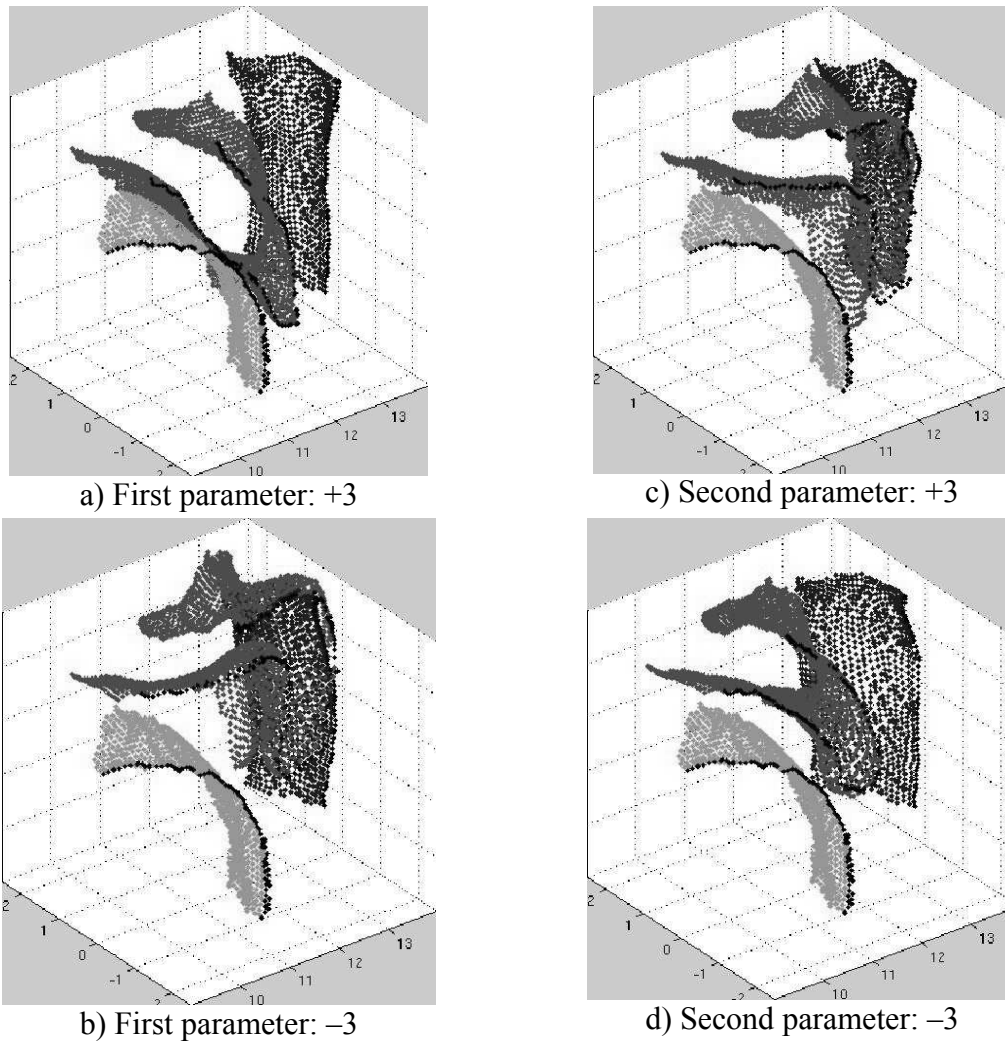
a) First parameter: +3

c) Second parameter: +3

b) First parameter: –3

d) Second parameter: –3

Figure 7: Representation of the velum, nasopharynx and tongue for opposite articulatory control parameters values for the first two components obtained by PCA. For clearness reasons we have represented only the right side of the three organs; the dark points are located in the midsagittal section.

In order to assess the possible correlation between velum and jaw / tongue through the palatoglossus muscle suggested in the literature (Wrench, 1999), we have used jaw height as the first articulatory control parameter of the velum, through a linear regression: we found that jaw height explains 37 % of the full cumulated variance of all the velum points (cf. Table 2). In comparison, jaw height explains 16,7% of the full variance of the tongue for the same subject and a similar corpus (Badin *et al.*, 2002).

Table 2: Variance of the velum points explained by the jaw height parameter.

Parameter	Explained variance
Jaw Height parameter	37.0 %

This explanation of the variance confirms the correlation between the complete shape of the velum and the jaw height in our corpus. The plot of velum height *vs.* jaw height for the 22 articulations of the corpus in figure 8 illustrates this correlation. The correlation coefficient between jaw and uvula heights is 0.37 for this corpus.
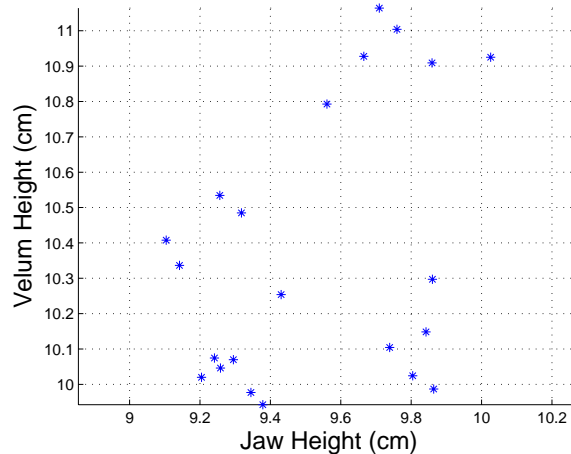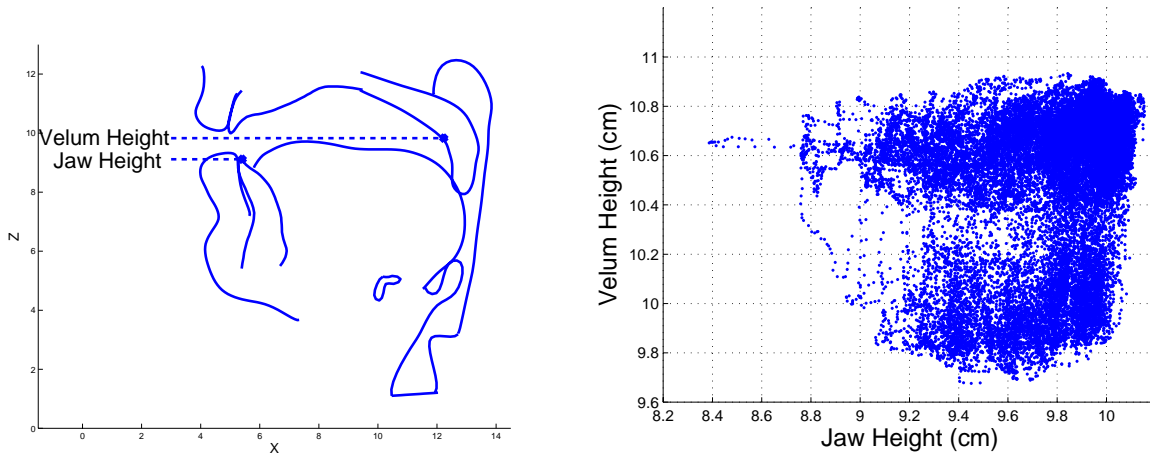

Figure 8: Velum height *vs.* jaw height for the 22 articulations. The associated correlation coefficient is 0.37.

However this correlation may be due to the content of our specific corpus of sustained articulations. In order to detect a possible corpus effect, a further investigation was carried out by means of an ElectroMagnetic midsagittal Articulograph. One of the coils of the articulograph was attached to the velum at about half way between the junction of the hard palate and of the velum and the extremity of the uvula of the subject, in the most mobile region of the velum, so as to provide an estimation of the velum movements, while another coil was attached to the lower incisors, in order to provide an estimation of jaw vertical movement (see figure 9a). A corpus of non-sense words [pVCVp] (V being one of the 14 French vowels and C one of the 16 French consonants) was recorded on the same subject. The figure 9b represents velum coil height *vs.* jaw coil height for the whole corpus (note that the complete trajectories are used). It appears clearly that the jaw and velum height parameters are not correlated. We found however a fairly strong correlation (0.88) between these parameters for the subset of the corpus constituted by the non-nasal vocalic targets; the correlation observed on the MRI corpus could be partly explained by a similar effect. Therefore, there is no reason to build a model with jaw height as first parameter.

a) Positions of the EMA coils   b) Velum height *vs.* jaw height

Figure 9: Positions of the EMA coils and velum height *vs.* jaw height
for the whole EMA corpus

5. Conclusion

The three-dimensional description of the velum shape for a set of 22 sustained articulations has allowed us to develop a first 3D linear articulatory model of the velum. The variance explained by the first two PCA factors for the 3D velum is not much below what we have found on the same corpus for the midsagittal contour of the velum sampled by means of a grid system similar to what Beautemps *et al.* (2001) used for the vocal tract, i.e. respectively 83.6% and 7.8% for the first and second factors. Badin *et al.* (2002) have shown that the whole 3D shape of the lips and of the tongue can be fairly well predicted from their midsagittal contours; this seems the case also for the velum and will be further checked. A 3D velum model driven by parameters measured in the midsagittal plane such as velum height for instance, could thus be developed. The next modelling step will be the determination of the area functions for both oral and nasal tracts as a function of the shapes of velum, tongue and pharynx wall, and thus the determination of the acoustical characteristics of the complete tract. A longer-term objective is finally to extend all this process to the complete 3D vocal and nasal tracts in order to build a complete 3D articulatory model of speech.

Acknowledgements

GMCAO team at TIMC for the use of the matching software *TestRigid* (Yohan Payan, Franz Chouly, Maxime Bérar). Finally we would like to thank Véronique Delvaux and Pascal Perrier for their helpful comments on the first version of the manuscript.

References

Amelot A., Crevier-Buchman L. & Maeda S. (2003) Observations of the velopharyngeal closure mechanism in horizontal and lateral directions from fiberscopic data. *Proc. of the 15th ICPhS*, Barcelona, Spain: 3021-3024.

Badin, P., Bailly, G., Raybaudi, M. & Segebarth, C. (1998). A three-dimensional linear articulatory model based on MRI data. In *Proceedings of the Third ESCA / COCOSDA International Workshop on Speech Synthesis*, Jenolan Caves, Australia, December 1998, 249-254.

Badin, P., Bailly, G., Revéret, L., Baciu, M., Segebarth, C. & Savariaux, C. (2002) Three-dimensional articulatory modeling of tongue, lips and face, based on MRI and video images. *Journal of Phonetics,* 30(3): 533-553.

Badin, P., Bailly, G., Elisei, F. & Odisio, M. (2003) Virtual Talking Heads and audiovisual articulatory synthesis (Invited talk at the symposium "*Articulatory synthesis. Advances and prospects"*). In *Proceedings of the 15th International Congress of Phonetic Sciences* (M.-J. Solé, D. Recasens & J. Romero, editors), Barcelona, Spain, 1: 193-197

Beautemps, D., Badin, P. & Bailly, G. (2001). Linear degrees of freedom in speech production: Analysis of cineradio- and labio-film data and articulatory-acoustic modeling. *Journal of the Acoustical Society of America,* 109(5): 2165-2180.

Bell-Berti, F. (1993). Understanding velic motor control: studies of segmental cotext. In Huffman & Krakow (Eds). *Phonetics and phonology. Nasals, Nazalisation and the Velum*. Vol 5. Academic Press Inc, 63-85.

Bouchet, A., & Cuilleret, J. (1980). Anatomie topographique, descriptive et fonctionnelle (la face, la tête et les organes des sens). *Edition Simep*.

Couteau, B., Payan, Y. & Lavallée, S. (2000). The mesh-matching algorithm: an automatic 3D mesh generator for finite element structures. *Journal of Biomechanics,* 33: 1005-1009.

Dang, J., Honda, K. & Suzuki, H. (1994) Morphological and acoustical analysis of the nasal and the paranasal cavities. *Journal of the Acoustical Society of America,* 96: 2088-2100.

Feng, G. & Castelli, E. (1996) Some acoustic features of nasal and nasalized vowels: a target for vowel nasalisation. *Journal of the Acoustical Society of America,* 99(6): 3694-3706.

Huffman, M. K. & Krakow, R. A. (1993). Phonetics and phonology. Nasals, Nazalisation and the Velum. Vol 5. Academic Press Inc.

Rossato, S., Badin, P. & Bouaouni, F. (2003) Velar movements in French: an articulatory and acoustical analysis of coarticulation. In *Proceedings of the 15th International Congress of Phonetic Sciences* (M.-J. Solé, D. Recasens & J. Romero, editors), Barcelona, Spain: 3141-3144.

Takemoto, H., Kitamura, T., Nishimoto, H. & Honda, K (2004). A method of tooth superimposition on MRI data for accurate measurement of vocal tract shape and dimensions. In *Acoustical Science and Technology*, 25(6):468-474

Teixeira, A., Vaz, F. & Príncipe, J.C. (2000) Nasal vowels following a nasal consonant. In *Proceedings of the 5th Seminar on Speech Production: Models and Data & CREST Workshop on Models of Speech Production: Motor Planning and Articulatory Modelling*, Kloster Seeon, Germany: 285-288.

Wrench, A.A. (1999). An investigation of sagittal velar movement and its correlation with lip, tongue and jaw movement. In *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco, United States: 435-438

Zemlin W.R. (1968). Speech and Hearing Sciences. Anatomy and Physiology. *Prentice-Hall, Inc.*