



HAL
open science

Modélisation tridimensionnelle des organes de la parole à partir d'images IRM pour la production de nasales - Caractérisation articulatoire-acoustique des mouvements du voile du palais.

Antoine Serrurier

► **To cite this version:**

Antoine Serrurier. Modélisation tridimensionnelle des organes de la parole à partir d'images IRM pour la production de nasales - Caractérisation articulatoire-acoustique des mouvements du voile du palais.. domain_stic.othe. Institut National Polytechnique de Grenoble - INPG, 2006. Français. NNT : . tel-00156977

HAL Id: tel-00156977

<https://theses.hal.science/tel-00156977>

Submitted on 25 Jun 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE

N° : □□□□□□□□□□

THÈSE

pour obtenir le grade de

DOCTEUR de l'INPG

Spécialité : SIGNAL, IMAGE, PAROLE, TÉLÉCOMS

préparée au laboratoire

Institut de la Communication Parlée, UMR CNRS 5009

dans le cadre de l'Ecole Doctorale

« Électronique, Électrotechnique, Automatique et Traitement du Signal »

présentée et soutenue publiquement par

Antoine Serrurier

le 8 décembre 2006

Titre :

**Modélisation tridimensionnelle des organes de la parole
à partir d'images IRM pour la production de nasales
Caractérisation articulatoire-acoustique des mouvements du voile du palais**

Directeur de thèse :

Pierre Badin

JURY

M.	Gang Feng,	Président
M.	Philip Hoole,	Rapporteur
M.	Yves Laprie,	Rapporteur
M.	Pierre Badin,	Directeur de thèse
M.	Shinji Maeda,	Examinateur
M.	Gérard Bailly,	Examinateur

INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE

THÈSE

pour obtenir le grade de

DOCTEUR de l'INPG

Spécialité : SIGNAL, IMAGE, PAROLE, TÉLÉCOMS

préparée au laboratoire

Institut de la Communication Parlée, UMR CNRS 5009

dans le cadre de l'Ecole Doctorale

« Électronique, Électrotechnique, Automatique et Traitement du Signal »

présentée et soutenue publiquement par

Antoine Serrurier

le 8 décembre 2006

Titre :

**Modélisation tridimensionnelle des organes de la parole
à partir d'images IRM pour la production de nasales
Caractérisation articulatoire-acoustique des mouvements du voile du palais**

Directeur de thèse :

Pierre Badin

JURY

M.	Gang Feng,	Président
M.	Philip Hoole,	Rapporteur
M.	Yves Laprie,	Rapporteur
M.	Pierre Badin,	Directeur de thèse
M.	Shinji Maeda,	Examineur
M.	Gérard Bailly,	Examineur

*À mon père,
scientifique convaincu.*

Remerciements

Décidé à travailler dans le domaine de la parole, me voilà débarqué à l'automne 2003 du pays du whisky et de la pluie, l'Écosse, à l'Institut de la Communication Parlée, au cœur des Alpes. L'accueil y fut chaleureux, et je souhaite en remercier tout particulièrement son directeur, Jean-Luc Schwartz, pour les rapports simples et amicaux qu'il a tout de suite instaurés et qui m'ont permis de me sentir très à l'aise pendant ces trois années à l'ICP.

Vingt-trois ans ont filé depuis la création de l'Institut de la Communication Parlée en 1983 jusqu'à son intégration dans le nouveau laboratoire GIPSA en 2007. Vingt-trois ans que le laboratoire a mis à profit pour élargir les connaissances autour de la parole, grâce à la réunion étonnante, mais non moins enrichissante, de phonéticiens et d'ingénieurs. La boucle est aujourd'hui bouclée entre l'une des premières thèses soutenues à l'ICP par Pierre Badin et la mienne, l'une des toutes dernières (l'avant-dernière exactement, pour quelques petites heures!). Pierre, merci pour m'avoir transmis ce savoir multidisciplinaire, tant d'un point de vue articulatoire que phonétique ou acoustique, mais merci surtout pour m'avoir montré le fascinant entrelacement de toutes ces disciplines pour mieux comprendre ce qu'est réellement la parole. Merci également pour ces nombreuses discussions (et pas seulement scientifiques) que j'ai pu partager avec toi, et pour ta rigueur, doublée de la détermination nécessaire pour la mettre en application. Enfin, je ne te serai jamais assez reconnaissant pour la disponibilité sans faille dont tu sais faire preuve, à rendre jaloux des générations de doctorants (et désolé à ta famille pour toutes ces fois où tu es rentré bien tard...)! Et puisque des remerciements de thèse ne seraient pas dignes sans la traditionnelle formule "sans toi cette thèse ne serait rien", merci encore pour ton voile du palais, sans lequel cette thèse ne serait vraiment rien...

Je souhaiterais remercier les membres du jury qui ont bien voulu se pencher sur ce mémoire pour y apporter leurs commentaires. Merci tout particulièrement à Gang Feng qui a accepté d'en être le président et dont l'intérêt pour les nasales ne s'est jamais démenti. Merci également aux rapporteurs Phil Hoole et Yves Laprie pour leur analyse riche et minutieuse de tous les éléments présentés dans ce mémoire. Je souhaiterais remercier sincèrement Shinji Maeda pour avoir examiné ce travail et qui fut à l'origine (sans le savoir!) de mon engouement pour la parole en étant le premier nom retenu, avec celui de Fant, lors d'un cours sur la parole organisé à Lannion au cours de mon DEA. Enfin, je souhaiterais remercier Gérard Bailly, qui a eu la difficile tâche d'endosser les rôles de chef d'équipe et d'examineur pour ce jury, pour ses suggestions sur ce travail et sa vision sur le long terme.

L'ICP, c'est aussi un couloir (que les campusiens m'excusent pour leur œuf!) et un coin

café. Cet étonnant couloir prend toute sa dimension grâce aux multiples personnalités étonnantes qu'on peut y croiser et avec qui il est toujours enrichissant de converser, de science comme du reste. Je tiens donc remercier tous les chercheurs avec qui j'ai pu partager ce couloir et qui resteront dans ma mémoire : Jean-Luc Schwartz, Matthieu Chabanas, Frédéric Elisei, Frédéric Berthommier, Nathalie Henrich, Benoît Ponsard, Xavier Pelorson, Denis Beautemps, Pierre Escudier. Un grand merci tout particulier à Pascal Perrier (mon parrain de thèse !), toujours à l'écoute lorsqu'on lui demande des conseils. Merci également à tous les thésards et postdocs que j'ai pu côtoyer et qui ont contribué à la bonne ambiance de ce couloir : Stephan, Lucie, Matthias, Noureddine, Julien, Yaroslav, Jana, Stéphanie, David, Jihène, Marc, Anahita, Claire, Mohammad, Viet Anh, Francesca et tous les stagiaires dont je ne préfère pas me lancer dans la liste ! Je souhaiterais remercier plus particulièrement tous ces jeunes chercheurs que j'ai eu la chance de connaître plus personnellement : Amélie et Marion, pour les merveilleux séjours à Lisbonne et Puy-Saint-Vincent, Nicolas, pour ses toilettes sèches et la Vérue, dont je suis fan de la première heure, Paupau, pour ses origines suisses, Guillaume, pour son bureau si ordonné qu'on croirait que personne n'y travaille, et enfin Xavier, pour la responsabilité du café et son arrivée pleine de vie à l'ICP...

L'ICP serait bien boiteux si l'on n'y associait pas le campus, avec son accueil si chaleureux toutes les (trop rares) fois où l'on s'y rend. Merci donc à Christophe, Coriandre et Alain pour leur disponibilité et leur bonne humeur permanente ! Merci également à Solange Rossato, Christian Abry, Louis-Jean Boë et Marie Cathiard pour l'intérêt qu'ils ont porté à mon travail ainsi que leurs conseils et réponses à mes questions. Et merci surtout à la bande de thésard(e)s qui forment pour moi une bonne partie de l'identité du campus : Claire, Aude, Fanny, Virginie et Nicolas.

Merci également à tous les membres de l'équipe technique et administrative qui m'ont aidé tout au long de ces trois ans à résoudre toutes les tracasseries de la vie quotidienne au labo : ce cher Support (bien qu'il ne réponde pas toujours aux mails...), Christian, Nimo, Monique, Nadine et Madame Gaude. Hélène, merci pour avoir partagé avec moi ce bureau. Merci pour ta bonne humeur, ta vitesse de pointe dans le couloir qui augmente avec les heures qui passent, merci pour l'ordinateur que je t'ai lâchement volé pendant ton absence (désolé...) et pour les quelques centimètres de bureau que j'ai également poussé pendant ton absence (hem...). En résumé merci surtout pour cette bonne ambiance, ces fous rires et tous ces petits papotages que j'ai pu partager avec toi pendant 3 ans. Je n'oublierai pas de sitôt ce joyeux bureau ! Une question me brûle cependant encore les lèvres : si j'ai bien compris que c'est Madeleine qui t'amena, il me reste encore à comprendre où...

Et pour refermer cette page ICP, un petit tour par le fameux bureau 523, celui d'en face... Je dois beaucoup à toutes les personnes de ce bureau et apparentés. Pour toutes les fois où la porte s'est refermée sur des ragots à n'en plus finir. Pour sa boîte verte. Pour tous les tableaux pourris. Pour tous les délires inoubliables. Virginie, la grande, merci pour ton accueil à l'ICP et pour ton entrain quotidien. Laurent, merci pour ton caractère pince-sans-rire. À Bertrand, pour sa fin de thèse "à chaud" qui restera dans les mémoires : merci pour tous les bons moments de détente. Une confidence : je crois que tu as raison, Fant est plutôt mon arrière-grand-père finalement.

Annemie, la petite, plus qu'une voisine de bureau tu auras été un soutien sans faille, jusqu'aux dernières heures de l'écriture de ce manuscrit tu auras été présente à l'ICP, facilitant grandement ces retours tardifs et ces nuits de rédaction : pour ta présence, merci ! Et enfin Julie, cousine, collègue, amie, sœur de thèse (ou devrais-je dire jumelle pour avoir accouché le même jour ?), quelle joie de t'avoir connue à l'occasion de ces trois années ! C'est un bout de famille solidement rattaché, et pour longtemps !

À tous mes proches qui ont dû me supporter durant cette période et qui ont contribué à leur manière à cette réussite. Tout d'abord à toi Elisa. Je ne peux quantifier, mais n'oublierai jamais tout ce que tu as pu m'apporter...merci d'être venue jusque là, merci pour tout ce que tu as fait, merci. Merci à toute la bande pseudo-hispanique de Brest : le beau Mario, Cydrille, Gégé, Ana et Cyril pour ses trop rares séjours ensemble ! J'espère que je trouverai dorénavant un peu plus de temps à vous consacrer. Et puis merci à toute une autre bande : celle de greuuuh !! Merci mille fois à tous les membres de ce groupe si hétérogène et délirant pour ces merveilleux moments collectifs réconfortants et ces orgies incommensurables ! Vous resterez sûrement les plus vifs souvenirs de ces années grenobloises : Victor, pour ton tempérament de sur-actif à l'origine de la formation de ce groupe, Andrew, pour ton caractère so british, Vass et Thierry, les piliers grenoblois, pour être toujours présents et motivés (et pour la relecture improvisée de ces remerciements !), Marie, pour tes audaces, Glòria, pour être si catalane, Luis et Hélène, pour vos discussions si endiablées, Deul le, pour les épices de ta cuisine, Mario et Sophie, pour ce mélange si multiculturel, Amandine, pour être toi..., JB et Catherine, pour former un couple si félin, Cécile P.....pour ton rire bien sûr !, et Marion, pour tes questions si rafraîchissantes ! Enfin, une holà toute particulière à tous ceux de ce groupe qui ont dû me supporter, moi et mes principes, en passant par la coLoc : Luis, pour ces polémiques si intenses qu'on a pu partager, Pascal, pour ton insouciance qui fait toujours rêver, Cécile, pour ton sens de la retenue, de la réflexion et de la modération avant d'agir, Marcus, pour ta présence si discrète. Enfin, last but not least, un merci tout spécial aux deux piliers de la coLoc présents pendant la rédaction de ce manuscrit qui ont assuré l'intendance (ah bon, pas que pendant la rédaction ?), ont écoutés à n'en plus finir mes histoires, mes soucis et mes ragots rapportés du labo (autour d'un bon p'tit apéro) bref à qui je dois une partie de ce travail : David et Ruben. David, merci pour ton esprit rationnel et insouciant, pour ta jeunesse éternelle (malgré ce que tu en penses) qui te permet de faire des conneries sans te soucier des conséquences ! Ruben, source infinie de savoir scientifique (même quand tu n'en sais rien !), merci infiniment pour ta complicité, tout simplement...

Enfin, je souhaiterais remercier ma famille pour son soutien sans cesse renouvelé durant ces trois années. Merci à Anne-Sophie, Arnaud et Céline (n'est-ce pas, ma grande !), que les distances éloignent mais avec qui l'univers familial recréé à chaque retrouvaille est magique. Et pour terminer cette longue liste, le dernier MERCI s'adresse à ma mère. Maman, merci pour ta patience, toi qui n'a pu que trop rarement constater l'efficacité du TGV qui relie Grenoble à Paris en 3h (si si !), merci pour ton soutien permanent, merci pour toute la compréhension dont tu sais faire preuve à l'égard de tes enfants. Merci pour ton dévouement.

Table des matières

Introduction	13
1 La production des nasales : définition, description, littérature	17
1.1 La nasalité en français	18
1.1.1 Rappels sur la production de la parole	18
1.1.2 Définition de la nasalité	20
1.1.3 Voyelles et consonnes nasales du français	20
1.2 Description anatomique du conduit nasal	21
1.2.1 Les structures rigides	22
1.2.2 Les structures déformables	23
1.3 Modélisation articulatoire-acoustique des nasales	27
1.3.1 Les approches de la modélisation articulatoire	28
1.3.2 Modélisation articulatoire-acoustique des nasales	30
1.3.2.1 Les données	30
1.3.2.2 Les modèles articulatoires du voile du palais et du port vélopharyngé	33
1.3.2.3 Les études articulatoire-acoustiques	34
1.4 Approche modélisatrice	35
2 Données articulatoires	39
2.1 Techniques d'exploration du conduit vocal	40
2.2 Les données articulatoires	43
2.2.1 Le corpus	43
2.2.2 Les moulages dentaires	44
2.2.3 Les images tomodensitométriques	44
2.2.4 Les images IRM	45
2.2.4.1 Les images IRM sagittales	45
2.2.4.2 Les images IRM coronales	46
2.2.5 Les recoupes des images	46
2.2.6 Les enregistrements EMA	51
2.3 Des images aux formes tridimensionnelles	52
2.3.1 Jongler avec les coordonnées	52
2.3.2 Tracer un contour	54
2.3.3 Obtenir une surface 3D	56

2.3.4	Manipuler les maillages	58
2.4	Alignement dans un repère commun	60
2.4.1	Définition d'un repère absolu	60
2.4.2	Alignement du palais dur dans le repère absolu	61
2.4.3	Alignement semi-automatique d'un maillage de structure rigide sur un bloc d'images	62
2.4.4	Alignement des images IRM	65
2.4.5	Alignement de la mâchoire	67
2.4.6	Alignement des enregistrements EMA	68
2.5	Les formes 3D des structures déformables	68
2.5.1	Principe de détermination des structures déformables	69
2.5.2	Le tracé : choix et évidences	69
2.5.3	Les maillages génériques	77
2.5.4	Déformations des maillages génériques	78
2.5.5	Une approche similaire en deux dimensions	81
2.5.6	Comparaison avec les données articulatoires dynamiques	83
2.5.7	Le cas des fosses nasales	84
2.6	Conclusion	86
3	Modélisation articulaire	89
3.1	Les cavités nasales	90
3.2	Principe de modélisation des structures déformables	95
3.3	Influence de l'inclinaison de la tête	99
3.4	Modèles préliminaires	103
3.4.1	Modèle médiosagittal	103
3.4.2	Modèle tridimensionnel	104
3.5	Modélisation linéaire tridimensionnelle	105
3.5.1	Le voile du palais	106
3.5.2	La paroi du pharynx	116
3.5.3	Comparaison aux données dynamiques	120
3.6	Modélisation linéaire médiosagittale	121
3.6.1	Le voile du palais	121
3.6.2	La paroi du pharynx	123
3.6.3	Comparaison aux données dynamiques	125
3.7	Du 2D vers le 3D	126
3.7.1	Des enregistrements EMA vers la forme 3D	126
3.7.2	Des contours médiosagittaux vers la forme 3D	129
3.8	Le rôle de la langue	131
3.8.1	Un lien moteur ?	131
3.8.2	Un lien mécanique ?	133
3.9	Conclusion	138

4 Fonctions d'aire	139
4.1 Le calcul des fonctions d'aire : principe général	140
4.2 Les fosses nasales	143
4.3 Le port vélopharyngé	145
4.3.1 Effets de VL sur la fonction d'aire	149
4.3.2 Effets de VS sur la fonction d'aire	152
4.4 Influence de la position du voile du palais sur le conduit oral	154
4.5 La covariation des aires de couplage orale et nasale	157
4.6 Conclusion	161
5 Modélisation acoustique	163
5.1 Modèle acoustique	164
5.1.1 Principes généraux	164
5.1.2 Évaluation acoustique du modèle articulatoire	168
5.2 La nasalité dans le triangle vocalique	170
5.2.1 La cible nasopharyngée	170
5.2.2 Les voyelles nasales oralisées	171
5.3 Les effets du couplage acoustique	173
5.3.1 Les résonances du conduit nasal	173
5.3.2 La transition du conduit oral vers le conduit nasopharyngé	174
5.3.3 Les effets acoustiques des paramètres VL et VS	180
5.4 L'influence de la modification de la fonction d'aire orale	186
5.5 Comparaison avec des enregistrements acoustiques	188
5.6 Conclusion	192
Conclusion et perspectives	195
A Modélisation linéaire	199
A.1 Analyse en Composantes Principales	200
A.2 Régression Linéaire	202
Publications	203
Références bibliographiques	205

Introduction



« Si le nez de Cléopâtre eût été plus court, la face du monde en eût été changée ». Comme l'analyse Pascal dans ses *Pensées* (1670) et nous le rappellent *Astérix et Obélix* dans leurs aventures qui constituent une référence sur l'étude de l'influence de la nasalité sur les rapports humains à l'époque romaine¹, la détermination de la taille de l'appendice nasal est une question importante qui mérite le détour. En termes de production de parole, alors que la grande majorité des sons émis sont oraux, il s'avère que « certains sons sortent par le nez », comme le fit remarquer *Fabricius ab Aquapendente* en 1603. Un tel son est alors dit *nasal*. Quel mécanisme articulatoire du conduit vocal est à l'origine d'un son nasal et quels en sont les conséquences acoustiques ? Cette thèse se fixe pour objectif général de contribuer à l'avancée des connaissances sur cette question.

Ce travail s'inscrit dans le cadre général des études sur la production de la parole et du développement de têtes parlantes audiovisuelles virtuelles menés à l'Institut de la Communication Parlée (voir par exemple [Badin *et al.*, 2003]). Le signal de parole audiovisuelle peut être en effet considéré comme l'ensemble des conséquences audible et visible des mouvements des organes articulatoires de la parole. Les modèles tridimensionnels de ces articulateurs constituent donc les éléments fondamentaux des têtes parlantes. C'est donc dans ce contexte que nous sommes intéressés à la production des nasales. Le geste articulatoire à l'origine de la nasalité est aujourd'hui connu : il correspond principalement à l'abaissement du voile du palais, qui assure la connexion du conduit nasal au conduit oral. Ce geste relativement simple a des conséquences acoustiques assez complexes. Si ces caractéristiques acoustiques ont été largement discutées dans la littérature, leur origine articulatoire n'est pas encore complètement clarifiée. En outre, le geste articulatoire lui-même reste aujourd'hui encore mal étudié. L'objectif principal de notre travail a donc été de faire progresser la connaissance des mécanismes articulatoires liés à la nasalité, en s'appuyant sur l'analyse de données articulatoires statiques et dynamiques sur un sujet humain, et en développant un modèle articulatoire tridimensionnel du conduit nasal basé sur ces données.

La **modélisation** constitue tout d'abord un moyen d'analyser, de comprendre, et de reproduire un mécanisme. Contrôlé par un nombre limité de paramètres, un modèle reproduit avec un certain degré de réalisme le mécanisme qu'il cherche à représenter. Un bon modèle résulte du choix optimal entre la précision de la représentation des données mesurées et le nombre de

¹Astérix et Cléopâtre (René Goscinny et Albert Uderzo), 1965

paramètres de contrôle utilisés. L'absence dans la littérature de modèle réaliste de mouvement du voile du palais a motivé notre travail. Même si le mouvement global du voile du palais est relativement simple, la complexité de la géométrie du conduit vocal dans le nasopharynx, au confluent des cavités pharyngée, buccale et nasales justifie amplement une approche tridimensionnelle.

Le degré élevé de réalisme souhaité et le souci de comprendre le mécanisme lui-même nous a naturellement conduits à opter pour une modélisation fonctionnelle basée sur les **données**. Cette approche consiste à faire émerger, par diverses analyses en composantes linéaires, les principaux mouvements des organes à partir des données articulatoires. Cette approche offre à la fois l'avantage de pouvoir élaborer un modèle qui représente avec précision les données – et qui peut donc être évalué de manière très directe par confrontation avec ces données –, et de ne pas impliquer d'hypothèses *a priori* sur le mécanisme étudié. Cette approche hérite de l'expérience acquise à l'Institut de la Communication Parlée ([Badin *et al.*, 2002]) dans ce domaine.

Enfin, la caractérisation des relations articulatoire-acoustiques dans le cas de la nasalité reste un objectif important de ce travail, qui s'inscrit dans une longue tradition à l'Institut de la Communication Parlée : [Mrayati, 1976], [Guérin et Mrayati, 1977], [Lherm, 1984], [Merlier, 1984], [Feng, 1986], [Castelli, 1989] et [Rossato, 2000]. Ainsi, cette thèse a pour ambition de contribuer, grâce au nouveau modèle articulatoire, à résoudre le difficile problème de la compréhension des relations entre caractéristiques acoustiques des voyelles et consonnes nasales et forme et position des organes impliqués dans la nasalité.

Organisation du mémoire

Ce travail vise à concevoir un modèle articulatoire tridimensionnel du conduit nasal basé sur les données et à contribuer, à l'aide de ce modèle, à la caractérisation articulatoire-acoustique des mouvements du voile du palais. Le mémoire se décompose pour cela en cinq chapitres.

Dans le **premier chapitre** nous rappelons les éléments indispensables à la compréhension de la nasalité. Nous exposons brièvement les notions de base en production de la parole et définissons la notion de nasalité. Une courte description anatomique des différentes structures impliquées dans le mécanisme de production des nasales et une rapide revue de l'état de l'art sur le sujet nous conduisent à poser les bases de l'approche de modélisation que nous choisissons.

Le **second chapitre** est centré sur l'acquisition et le prétraitement des données articulatoires qui constituent les fondements de notre travail. Ce chapitre explique notamment la détermination à partir d'images anatomiques volumiques des formes tridimensionnelles des organes dont la constitution d'une base de données servira à construire le modèle articulatoire.

Le **troisième chapitre** présente le cœur de cette thèse : le modèle articulatoire. Une première partie est dédiée à la description géométrique de la partie rigide du conduit nasal. Le reste du

chapitre s'articule autour de la modélisation de la partie déformable du conduit nasal, véritable enjeu de ce travail. Après avoir posé les bases du principe d'extraction des mouvements à partir des données et exposé brièvement les modèles préliminaires développés à l'Institut de la Communication Parlée basés sur cette approche, nous proposons un modèle tridimensionnel. Dans la suite du chapitre nous mettons en lumière les relations entre les contours médiosagittaux et les formes tridimensionnelles des organes étudiés et nous nous intéressons aux interactions possibles avec la langue.

La suite du mémoire est centrée sur la détermination des caractéristiques acoustiques du conduit vocal dans le cas de la nasalité. Ces caractéristiques sont déduites de la fonction d'aire du conduit nasal obtenue à partir du modèle tridimensionnel développé au chapitre précédent. Le **quatrième chapitre** est centré sur la détermination de la fonction d'aire des parties rigide et déformable du conduit nasal à partir du modèle articulatoire tridimensionnel. L'influence des mouvements du voile du palais sur les fonctions d'aire orale et nasale est étudiée.

Dans le **dernier chapitre** nous étudions les évolutions spectrales de la fonction de transfert acoustique en fonction des mouvements des organes du modèle articulatoire. La dernière partie présente la comparaison des fonctions de transfert obtenues à l'aide du modèle avec des enregistrements acoustiques sur le même sujet.

Chapitre 1

La production des nasales : définition, description, littérature

Les nasales ont toujours été la source de nombreuses interrogations, tant du point de vue production que perception ([Ferguson *et al.*, 1975], [Huffman et Krakow, 1993]). Les nombreux modèles articulatoires du conduit vocal développés depuis une cinquantaine d'année ont cependant laissé de côté le difficile problème des nasales. Ce vide traduit la difficulté à obtenir des données précises et fiables du conduit nasal, relativement peu accessible comparé au reste du conduit vocal. Les nouvelles méthodes d'acquisition de données développées ces vingt dernières années permettent aujourd'hui de reconsidérer ce problème. Après quelques rappels sur la nasalité et l'anatomie du conduit nasal, ce chapitre dresse un portrait de la modélisation des nasales jusqu'à aujourd'hui et justifie l'approche adoptée dans ce travail.

1.1 La nasalité en français

1.1.1 Rappels sur la production de la parole

Ce paragraphe se limite à des rappels sur les principes généraux de la production de la parole. Il est possible pour plus de détails de se référer aux différents ouvrages de la littérature dans ce domaine (par exemple [Flanagan, 1972], [Minifie *et al.*, 1973], [Hardcastle, 1976] ou [Kent, 1997]).

L'appareil phonatoire peut se décomposer en trois parties : le système subglottique, composé des poumons et de la trachée artère, le larynx, qui contient les cordes vocales, et le conduit vocal, qui rassemble l'ensemble des articulateurs et cavités supraglottiques (voir la figure 1.1). Les sons de parole sont normalement produits lors de la phase d'expiration grâce à un flux d'air

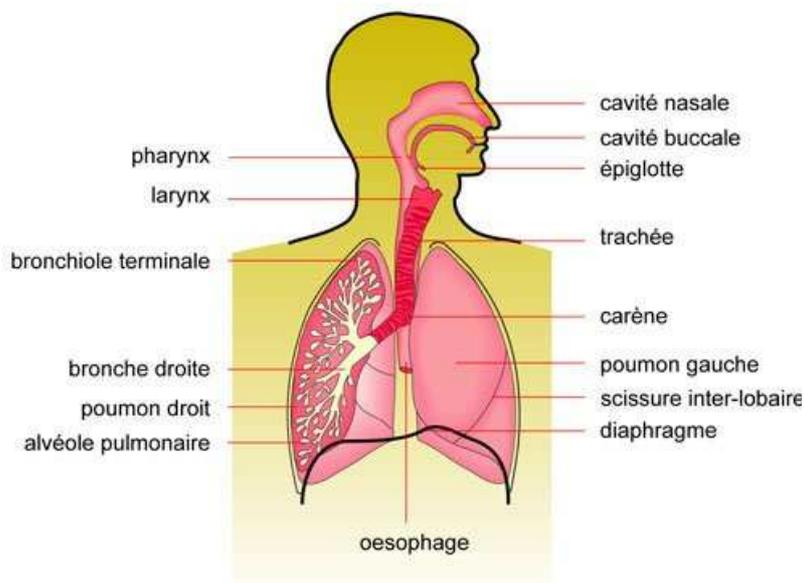


FIG. 1.1 – Appareil phonatoire (www.infovisual.info).

contrôlé, en provenance des poumons et passant par la trachée artère. Le flux entre alors dans le conduit cartilagineux du larynx pourvu d'un système musculaire complexe. Situé dans la région

moyenne du cou, le larynx peut se déplacer vers le haut ou vers le bas. De ce fait, la longueur de la cavité pharyngienne (située juste au-dessus) peut se trouver modifiée. A l'intérieur du larynx se situent les cordes vocales, deux membranes symétriques particulièrement importantes pour la phonation. Au passage de l'air, elles peuvent se mettre à vibrer de manière pseudo-périodique, fournissant ainsi une onde de débit ressemblant à des dents de scie, à une fréquence appelée fréquence fondamentale. Les sons obtenus sont alors dits voisés. L'espace entre les cordes vocales est appelée glotte.

L'air laryngé passe dans le conduit vocal qui contient plusieurs cavités supra-glottiques reliées entre elles (voir figure 1.2) :

- Le pharynx : conduit situé à l'arrière de la langue entre l'entrée de la cavité buccale et du conduit nasal d'une part et le larynx d'autre part.
- La bouche : cavité située entre le pharynx et la sortie aux lèvres, limitée en bas par la langue et en haut par le palais.
- Le conduit nasal : situé entre le pharynx auquel il est relié par le port vélopharyngé, dont l'ouverture ou la fermeture est principalement contrôlée par le voile du palais, et les sorties aux narines. Ce conduit est essentiellement constitué de deux cavités, les fosses nasales. Celles-ci partent du cavum, cavité située au-dessus du voile du palais, pour rejoindre les deux narines en sortie. De nombreuses cavités résonantes, les divers sinus paranasaux, sont reliées aux fosses nasales par des passages fins.

L'ensemble des cavités supraglottiques agit comme un système de résonateurs acoustiques de formes et de volumes qui peuvent varier au cours du temps grâce à la mobilité des articulateurs associés, et qui modulent ainsi le spectre des sources d'excitation acoustiques de la sortie du larynx.

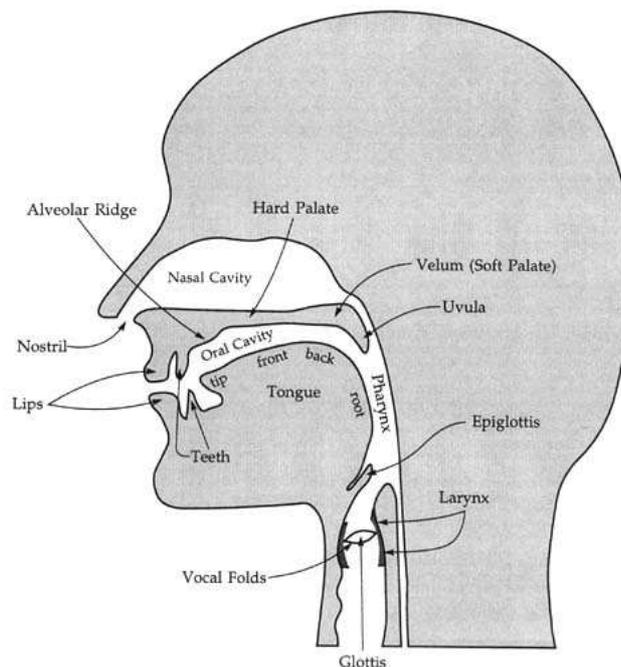


FIG. 1.2 – Coupe médiosagittale du conduit vocal (www.ling.upenn.edu).

1.1.2 Définition de la nasalité

[Crystal, 1997] définit la nasalité d'un point de vue phonétique comme l'ensemble des sons produits avec le voile du palais abaissé de manière à autoriser le passage d'un flux d'air audible à travers le nez. Comme nous le verrons plus loin, le mouvement d'abaissement du voile du palais lors de la production de voyelles et de consonnes nasales s'accompagne en français d'un certain nombre d'ajustements des autres articulateurs du conduit vocal. Les voyelles et les consonnes nasales se distinguent alors en français de leurs correspondantes orales par le trait phonologique [+ nasal], par opposition au trait [- nasal].

Les indices acoustiques des différents phonèmes qui permettent d'identifier le trait [+ nasal] sont appelés les corrélats acoustiques de la nasalité. Bien qu'il existe quelques corrélats dans le domaine temporel (sur la longueur des voyelles par exemple, voir [Delattre, 1968]), la recherche de ceux-ci dans le domaine fréquentiel a constitué un champ très important des travaux sur la nasalité. On peut citer quelques contributions majeures dans ce domaine : [Delattre, 1954], [House et Stevens, 1956], [Fant, 1960], [Fujimura et Lindqvist, 1971], [Lindqvist-Gauffin et Sundberg, 1972], [Mrayati, 1976], [Guérin et Mrayati, 1977], [Hawkins et Stevens, 1985], [Maeda, 1993] ou [Feng et Castelli, 1996]. Si le geste articulatoire paraît simple, les conséquences acoustiques sont nombreuses et diverses. [Maeda, 1993] souligne d'ailleurs la différence des corrélats en fonction des voyelles. Malgré cette grande diversité intra- et inter-locuteurs, quelques tendances générales ressortent :

- Le spectre dans les basses fréquences se trouve modifié par rapport à celui des phonèmes oraux correspondants.
- Un premier formant nasal apparaît entre 250 Hz et 400 Hz.
- L'apparition de ce formant s'accompagne de l'apparition d'un zéro dans la zone du premier formant oral, qui subit ainsi plusieurs modifications : baisse de son amplitude, élargissement de sa bande passante.
- Le spectre dans les hautes fréquences subit des modifications avec l'apparition de nouveaux formants et des variations d'amplitude, de fréquence et de largeur de bande des formants oraux.
- Ces changements dans les plus hautes fréquences, apparemment moins significatifs, engendrent une distribution spectrale d'énergie plus diffuse, avec des pics moins intenses.

La difficulté à trouver les corrélats acoustiques de la nasalité souligne une relation articulatoire-acoustique complexe : les effets de l'abaissement du voile du palais s'accompagnent de nombreuses variations du spectre, en partie dépendantes de la configuration orale du conduit vocal.

1.1.3 Voyelles et consonnes nasales du français

Le français, comme près de 99% des langues du monde, possède des consonnes nasales ([Stefanuto et Vallée, 1999]). Toutes occlusives, elles sont au nombre de trois, dont deux très dominantes, la bilabiale /m/ (*mot*) et l'alvéolaire /n/ (*nez*) ; une consonne vélaire, /ŋ/ (*gnon*), est moins fréquente que les deux précédentes. Notons que certains phonologues analysent ce dernier

son comme une simple variante de /nj/ (d'après [Carton, 1974]).

À l'inverse des consonnes, seulement 22% des langues du monde possèdent des voyelles nasales (voir par exemple [Vallée, 1994]). Le phénomène de nasalisation pour les voyelles du français apparaît vers le XI^{ème} siècle pour aboutir vers le XVII^{ème} siècle aux *voyelles nasales*, terme utilisé pour la première fois pour le français par l'Abbé de Dangeau en 1694¹, telles que nous les connaissons aujourd'hui en France. Elles sont au nombre de quatre : / $\tilde{\text{œ}}$ / (*un*), / $\tilde{\text{o}}$ / (*bon*), / $\tilde{\text{ɛ}}$ / (*vin*) et / $\tilde{\text{ɑ}}$ / (*blanc*). Notons que la distinction entre les voyelles / $\tilde{\text{œ}}$ / et / $\tilde{\text{ɛ}}$ / n'est pas faite par tous les locuteurs français et qu'il est donc courant de ne considérer que les trois voyelles / $\tilde{\text{ɑ}}$, $\tilde{\text{o}}$, $\tilde{\text{ɛ}}$ / comme voyelles nasales du français ([Carton, 1974]).

La nasalisation des voyelles orales / ɑ - a , ɔ , ɛ , œ / par l'ouverture du port vélopharyngé pour former les voyelles nasales / $\tilde{\text{ɑ}}$, $\tilde{\text{o}}$, $\tilde{\text{ɛ}}$, $\tilde{\text{œ}}$ / s'accompagne d'un certain nombre d'ajustement des autres articulateurs du conduit vocal (voir par exemple [Zerling, 1984], [Maeda, 1993] ou [Delvaux *et al.*, 2002a]) :

/ ɑ / - / $\tilde{\text{ɑ}}$ / : le recul de la langue est plus marqué pour le / $\tilde{\text{ɑ}}$ / tandis que les lèvres sont plus arrondies et plus avancées. L'articulation oralisée du / $\tilde{\text{ɑ}}$ / se rapproche en ce sens du / ɑ / et du / ɔ /.

/ ɛ / - / $\tilde{\text{ɛ}}$ / : le recul de la langue et son abaissement dans la région centrale sont assez marqués pour la voyelle nasale.

/ œ / - / $\tilde{\text{œ}}$ / : les contrastes entre ces deux voyelles ressemblent à ceux entre les voyelles / ɛ / et / $\tilde{\text{ɛ}}$ / à un degré moins marqué. Une certaine protrusion des lèvres peut être observée.

/ ɔ / - / $\tilde{\text{o}}$ / : la langue recule dans le cas du / $\tilde{\text{o}}$ / pour se rapprocher du / o / tandis que les lèvres sont plus arrondies et plus avancées. Cette voyelle nasale serait plus labialisée que le / ɔ / et le / $\tilde{\text{ɑ}}$ /.

On observe donc clairement qu'une voyelle nasale ne se différencie pas de sa *contrepartie* orale par un simple abaissement du voile du palais mais également par un certain nombre de réajustements des autres articulateurs. L'abaissement du voile du palais s'accompagne de manière simplifiée d'un recul de la langue et d'une protrusion des lèvres. Si [Zerling, 1984] considère ces réajustements comme des modifications articulatoires distinctes du trait de nasalité, [Maeda, 1993] pense qu'ils renforcent la perception de la nasalité.

1.2 Description anatomique du conduit nasal

Le conduit nasal peut être décomposé en deux parties : une partie déformable, du point de séparation du conduit oral jusqu'au cavum, dont le passage s'appelle le port vélopharyngé, et une partie rigide comprenant les fosses nasales, les narines et les divers sinus paranasaux. Ces deux parties sont étudiées indépendamment dans cette section.

¹dans *Essais de grammaire contenus en trois lettres d'un académicien à un autre académicien*

Cette partie a été réalisée à l'aide des atlas anatomiques de [Netter, 1997], [Csillag, 1999] et [Sobotta, 2001] et de nombreux ouvrages descriptifs tels que [Minifie *et al.*, 1973], [Hardcastle, 1976], [Le Huche et Allali, 1984], [Bouchet et Cuilleret, 1991], [Bell-Berti, 1993], [Kent, 1997]. Nous remercions également le professeur Jacques Lebeau (CHU Grenoble) et Franz Chouly pour les discussions et leurs conseils bibliographiques.

Il paraît avant tout nécessaire de définir quelques termes d'orientation utilisés en médecine et qui seront repris tout au long de ce document :

sagittal : désigne une orientation qui divise le corps humain en une partie droite et gauche. Le plan *médiosagittal* est le plan particulier qui divise le corps humain en deux parties droite et gauche équivalentes ; tous les plans sagittaux lui sont parallèles.

coronal : désigne une orientation qui divise le corps humain en une partie avant (celle contenant le visage) et une partie arrière (contenant le dos).

axial : désigne une orientation qui divise le corps humain en une partie basse (celle des pieds) et une partie haute (celle de la tête).

transverse : désigne une orientation qui coupe un élément de forme allongée, un tuyau par exemple, de manière perpendiculaire à sa ligne directrice.

1.2.1 Les structures rigides

La partie rigide du conduit nasal s'étend du cavum jusqu'aux narines. Le cavum, aussi appelé rhinopharynx ou epipharynx, est une cavité située à l'arrière du conduit nasal, au-dessus du voile du palais et sous le sinus sphénoïdal. Son volume dépend de la position du voile du palais.

Le cavum se sépare ensuite en deux cavités, les choanes, qui s'étendent ensuite jusqu'au nez pour former les fosses nasales. Elles sont limitées en bas par la voûte palatine qui les sépare de la cavité buccale, en avant par les narines et en haut par l'os nasal dans la partie antérieure et par l'os ethmoïde dans la partie postérieure, qui limite également les fosses nasales du maxillaire dans les régions latérales. Les deux cavités sont séparées par une cloison nasale relativement plane, le septum. La forme particulière de chaque cavité est liée aux trois structures cartilagineuses et osseuses que sont les cornets. Les cornets supérieur et moyen sont portés par l'os ethmoïde tandis que le cornet inférieur constitue une structure à part. Ces trois éléments en forme de turbines qui s'étendent tout le long de la cavité façonnent la forme tourmentée de celle-ci : les trois espaces qui s'enroulent par le dessous autour des trois cornets s'appellent les méats supérieur, moyen et inférieur (voir la figure 1.3). Les cornets et la paroi du septum sont recouverts d'une muqueuse. La partie terminale des cavités nasales est composée d'un orifice cartilagineux, le nez. Les deux conduits, les narines, débouchent au-dessus des lèvres dans un plan presque axial.

Les sinus paranasaux sont constitués des sinus maxillaires, frontaux, ethmoïdaux et sphénoïdal (voir la figure 1.3). Ceux-ci sont d'importantes cavités logées dans les structures osseuses et situées autour des fosses nasales, auxquelles elles sont reliées par de petits passages étroits appelés ostia. Le sinus sphénoïdal, unique et médian, est situé dans l'os sphénoïdal, au-dessus du

cavum. Il est relié à la partie postérieure des fosses nasales. Les sinus ethmoïdaux situés dans la partie haute des fosses nasales sont composés de nombreuses petites cellules reliées aux fosses au niveau du méat moyen. Les sinus maxillaires sont les plus importants en volume. Ils sont situés de part et d'autre des fosses nasales, sous les cavités orbitaires, au-dessus des racines des molaires et des prémolaires. Ils sont également reliés aux fosses nasales au niveau du méat moyen. Enfin les sinus frontaux sont situés au-dessus du toit orbitaire et sont encore reliés aux fosses nasales au niveau du méat moyen.

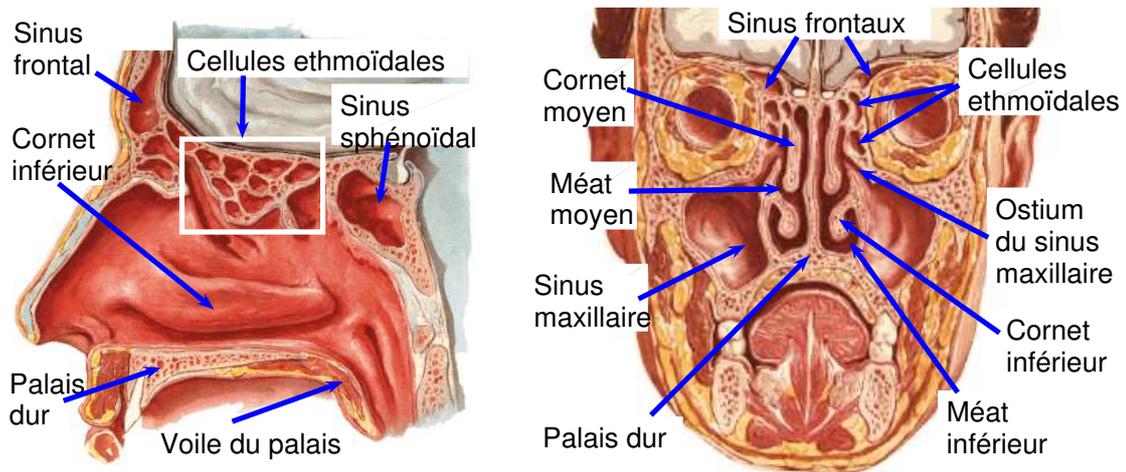


FIG. 1.3 – Coupe sagittale (gauche) et coronale (droite) de la cavité nasale (d'après [Netter, 1997]).

1.2.2 Les structures déformables

Le conduit nasal se connecte au conduit oral dans la région pharyngée. Le passage vertical de la cavité pharyngée au cavum est une région déformable du conduit nasal et s'appelle le port vélopharyngé. Il est bordé par le voile du palais dans sa partie antérieure et par les parois latérales et arrières du pharynx.

Le voile du palais est le principal articulateur du port vélopharyngé. Comme le montre la figure 1.4, ce tissu musculaire prolonge le palais dur vers le fond de la cavité buccale et se sépare en trois branches : deux branches latérales qui s'étirent chacune vers le haut et vers le bas et une branche centrale, la luette, qui vient former un appendice au fond de la cavité buccale et marquer la limite avec la cavité pharyngée. Il est constitué de cinq muscles, deux latéraux qui s'étirent dans la partie supérieure, deux autres latéraux qui s'étirent dans la partie inférieure, et un central :

Le muscle *Levator veli palatini* : c'est le muscle principal du voile du palais, situé dans la partie supérieure. Il prend son origine en éventail dans la partie supérieure médiane de l'aponévrose du voile (voir figure 1.5) et s'étire symétriquement de chaque côté le long de

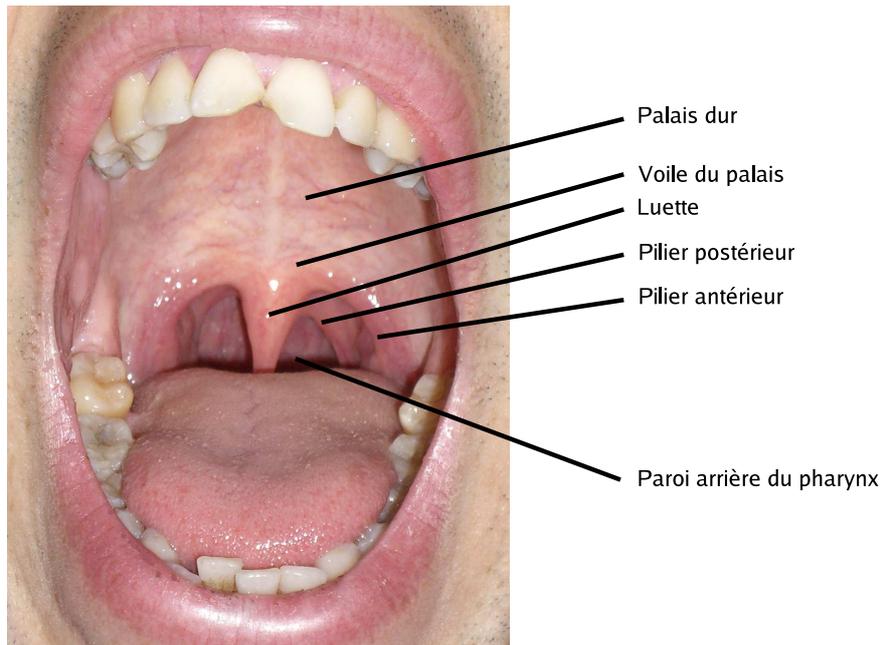


FIG. 1.4 – Vue antérieure de la bouche de l’auteur de cette thèse

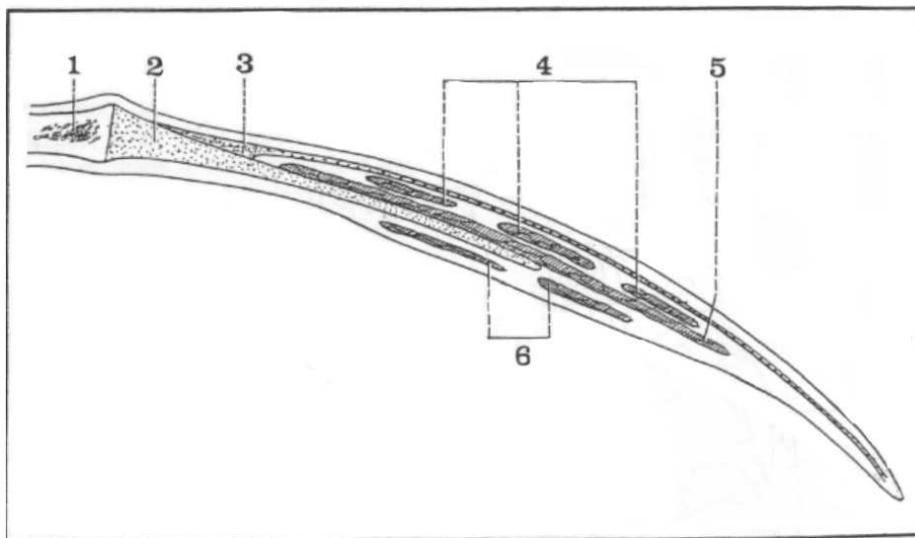


FIG. 1.5 – Coupe médiosagittale schématique du voile du palais (d’après [Bouchet et Cuilleret, 1991]) ; 1 : Palais osseux ; 2 : Aponévrose du voile ; 3 : Azygos de la lurette ; 4 : Levator veli palatini ; 5 : Palatopharyngien ; 6 : Palatoglosse.

la paroi pharyngée jusqu’à la partie interne cartilagineuse de la trompe d’Eustache (voir figure 1.7). Sa fonction essentielle est d’élever le voile du palais (voir figure 1.6a). On lui attribue également l’ouverture de la trompe d’Eustache.

Le muscle *Tensor veli palatini* : c’est le second muscle supérieur du voile du palais. Il prend son origine dans l’aponévrose du voile du palais (voir figure 1.5) et s’étire latéralement et

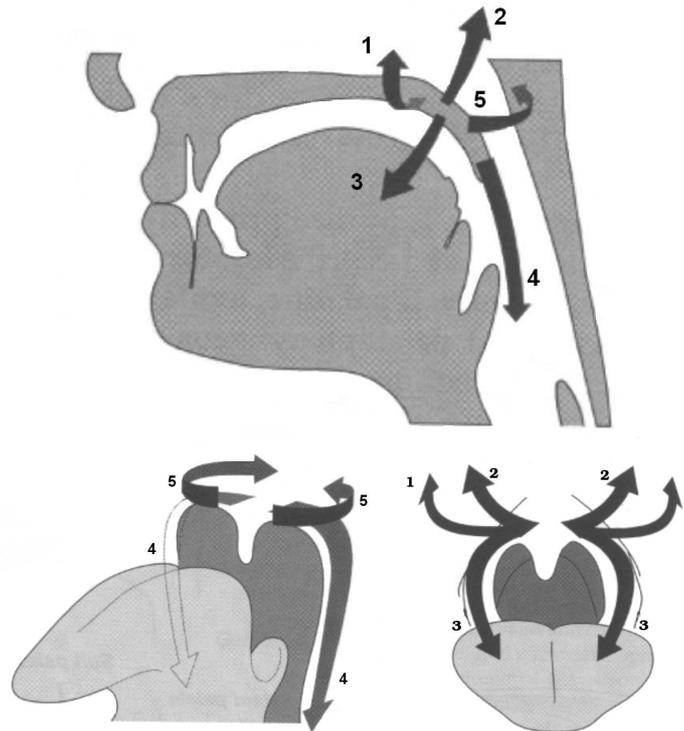


FIG. 1.6 – Vues médiosagittale (haut), antérieure oblique (bas gauche) et antérieure (bas droite) du conduit vocal et directions schématiques de l'action des principaux muscles impliqués dans le mouvement du voile du palais et du port vélopharyngé (d'après [Kent, 1997]) ; 1 : Tensor veli palatini ; 2 : Levator veli palatini ; 3 : Palatoglosse ; 4 : Palatopharyngien ; 5 : Constricteur supérieur du pharynx.

symétriquement jusqu'à la base latérale du crâne en passant par un tendon de l'hamulus qui agit comme une poulie pour assurer une tension latérale du voile du palais (voir figure 1.7a). Son action d'étirement latéral du voile du palais est assez limité en parole (voir figure 1.6).

Le muscle Palatoglosse : c'est un muscle commun avec la langue, donc dans la région inférieure du voile du palais. Prenant son origine dans la partie médiane inférieure sous l'aponévrose du voile (voir figure 1.5), il rejoint les bases latérales de la langue en longeant les bords de la cavité orale de manière cylindrique, formant ainsi les piliers antérieurs (voir figure 1.4), avant de s'unir aux fibres du styloglosse (voir figure 1.7b). Ce muscle qui abaisse le voile du palais semble être le principal muscle antagoniste au levator veli palatini (voir figure 1.6), bien que son rôle en parole reste controversé (voir par exemple [Bell-Berti, 1993]).

Le muscle Palatopharyngien : c'est un muscle commun avec la paroi du pharynx dans la région inférieure du voile du palais. Il prend son origine principale dans la partie médiane supérieure au-dessus de l'aponévrose du voile (voir figure 1.5) et rejoint la paroi du pharynx (voir figure 1.7b) en formant les deux piliers postérieurs sur les bords de la cavité orale (voir figure 1.4). Son rôle en parole semble assez limité ; il peut contribuer en rapprochant ses piliers de ceux du palatoglosse à la constriction de l'isthme du gosier. Il agit principalement dans le cas de la déglutition.

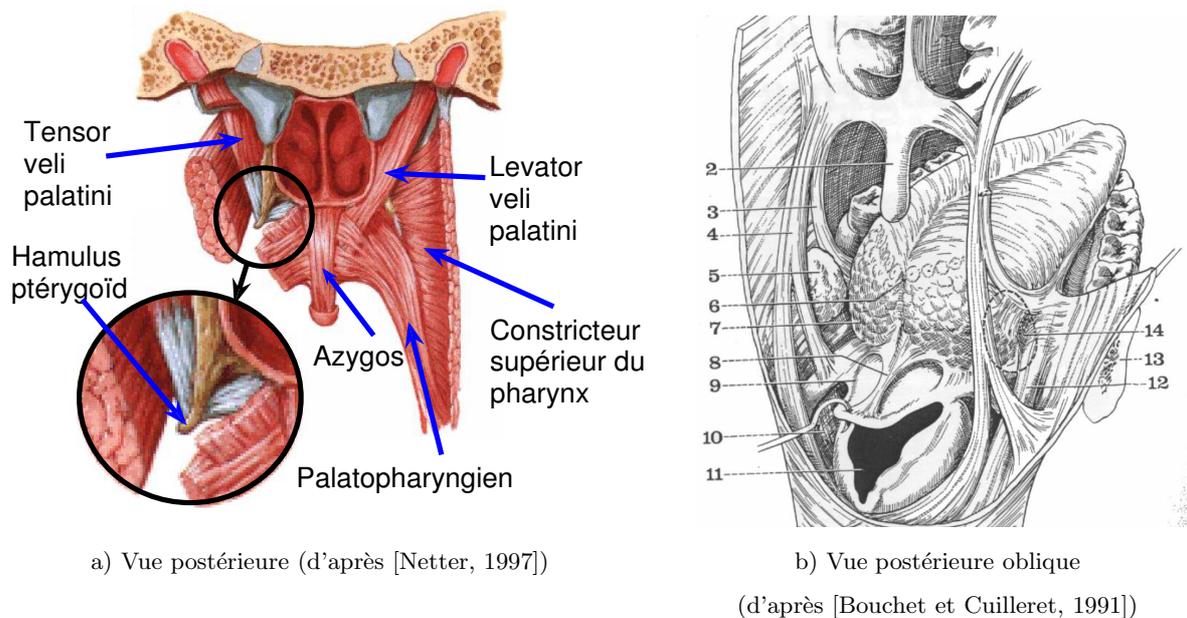


FIG. 1.7 – Vues postérieures schématisques du conduit vocal au niveau du voile du palais. Sur la figure b : 2 : Luvette ; 3 : Palatoglosse ; 4 : Palatopharyngien ; 12 : Constricteur moyen du pharynx.

Le muscle Azygos : c'est le seul muscle entièrement interne du voile du palais. Il s'agit d'un rassemblement de nombreuses petites fibres qui s'étendent de la partie supérieure de l'aponévrose proche du palais dur jusqu'à la pointe de la luvette (voir figure 1.5). Son rôle semble très limité et pourrait légèrement élever la luvette et raccourcir le voile du palais.

Les parois du pharynx sont principalement actives par l'intermédiaire du muscle constricteur supérieur qui entoure le port vélopharyngé dans sa partie supérieure. Les muscles constricteurs moyen et inférieur complètent l'action du constricteur supérieur mais agissent principalement dans le cas de la déglutition. Les muscles stylopharyngien et salpingopharyngien présents dans les paroi latérales du pharynx ne jouent presque aucun rôle en parole.

Le rôle actif ou passif joué par chaque muscle impliqué dans le mécanisme de fermeture du port vélopharyngé en parole a largement été étudié dans la littérature et a mené à de nombreuses interprétations : voir par exemple [Fritzell, 1969], [Lubker *et al.*, 1970], [Dickson et Maue-Dickson, 1972], [Bell-Berti, 1976], [Baken, 1987], [Dixit *et al.*, 1987], [Moon *et al.*, 1994], [Kollia *et al.*, 1995] ou [Wrench, 1999]. [Bell-Berti, 1993] souligne néanmoins que : « *the levator [veli] palatini muscle is widely accepted as the muscle primarily responsible for closing the velopharyngeal port by exerting an upward and backward pull on the velum* ». La fermeture du port vélopharyngé est donc atteinte principalement par la contraction de ce muscle qui provoque une remontée et un recul du voile du palais et le force à entrer en contact avec les parois du pharynx et à assurer la fermeture. Ces parois peuvent également jouer un rôle pour compléter la fermeture du port vélopharyngé en agissant par un effet « sphincter » (voir par exemple [Skolnick *et al.*, 1973],

[Parush et Ostry, 1986], [Finkelstein *et al.*, 1995], [Kent, 1997] ou [Poppelreuter *et al.*, 2000]). La littérature référence quatre types de fermeture du port vélopharyngé, en fonction de l'action des muscles constricteurs des parois du pharynx selon les locuteurs, comme le montre le schéma de la figure 1.8 proposé par [Kent, 1997]. À partir de données fibroscopiques, [Amelot *et al.*, 2003] constatent effectivement plusieurs stratégies de fermeture en fonction du locuteur. On note que dans le quatrième type de fermeture (type *d* sur la figure 1.8), la paroi arrière du pharynx forme une proéminence appelée Bourrelet de Passavant (mise en évidence par [Passavant, 1869]) pour venir en contact avec le voile du palais et participer à la fermeture du port vélopharyngé ([Zemlin, 1968]). Un résumé schématique de l'action des muscles principaux du port vélopharyngé est visible à la figure 1.6.

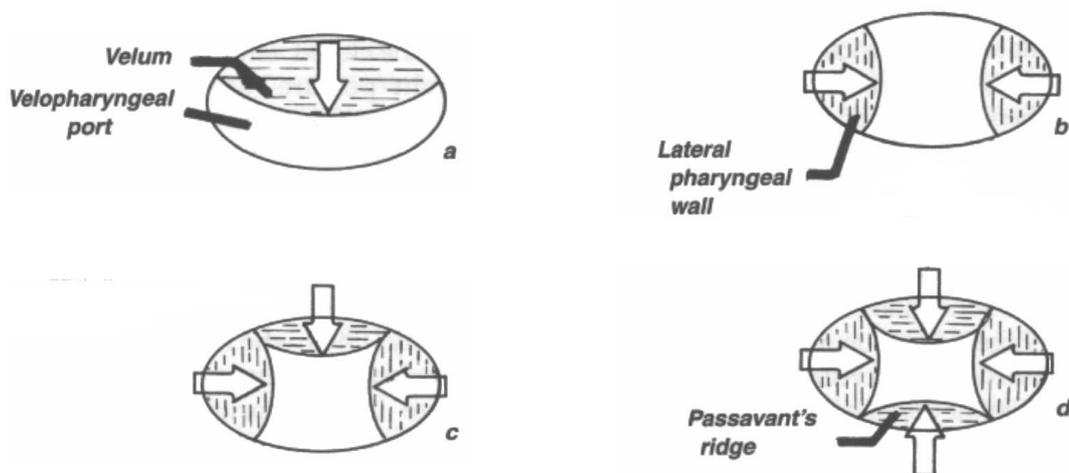


FIG. 1.8 – Schémas des quatre types de fermeture du port vélopharyngé d'après [Kent, 1997] : a) coronale, b) sagittale, c) circulaire, d) circulaire avec Bourrelet de Passavant.

1.3 Modélisation articulatoire-acoustique des nasales

La modélisation articulatoire-acoustique des nasales constitue un enjeu important pour la compréhension de la nasalité. En fonction du degré de représentation de la réalité, plusieurs approches de modélisation articulatoire sont possibles. Cette section détaille ces différentes approches. Le développement d'un modèle articulatoire permet par la suite d'étudier les conséquences acoustiques des gestes des organes du conduit vocal. Au vu de la difficulté à produire un modèle articulatoire complet de nasales, de nombreuses études se sont limitées à l'étude acoustique des variations articulatoires produites directement à l'aide de fonctions d'aire. Nous présentons alors une revue des différents travaux sur les relations articulatoire-acoustiques dans le cas de la nasalité basés sur cette approche.

1.3.1 Les approches de la modélisation articulatoire

La modélisation articulatoire consiste à représenter avec un degré de réalisme plus ou moins élevé la forme du conduit vocal en vue de générer des sons en imitant le processus de production humain qui considère le conduit vocal comme un résonateur acoustique. Le modèle permet de contrôler par un nombre limité de paramètres la forme et la position des articulateurs, et donc la géométrie du conduit vocal. La première machine parlante mécanique aboutie est généralement attribuée à von Kempelen (1791), où le conduit vocal est modélisé par un « résonateur en caoutchouc dont la forme est manipulée par la main gauche de l'opérateur » ([Badin, 2002]). Une machine plus élaborée est produite grâce à Faber en 1846 ([Badin, 2002]) : le conduit vocal est élaboré autour de ses organes principaux ; on se rapproche alors d'un modèle articulatoire. La notion de fonction d'aire, introduite par [Chiba et Kajiyama, 1941], marque le début du développement de nombreux modèles articulatoires ayant pour objectif de simuler la propagation des ondes acoustiques dans les basses fréquences. La fonction d'aire est définie comme l'aire transverse du conduit vocal en fonction d'une coordonnées sur une ligne médiane du conduit. La propagation du son dans le conduit vocal, représenté par sa fonction d'aire, est considérée comme analogue à la propagation électrique dans une ligne de transmission ([Fant, 1960]). Le premier modèle complet de ce type permettant de produire des sons de parole peut être attribué à [Dunn, 1950] (d'après [Badin, 2002]). La seconde moitié du XX^{ième} siècle est marqué par l'émergence de nombreux modèles articulatoires avec l'élaboration d'une fonction d'aire pour objectif (voir par exemple [Boë *et al.*, 1994], [Badin, 2002] et [Bailly *et al.*, 2007] pour des revues détaillées). Un des objectifs de la modélisation articulatoire est donc d'obtenir la forme géométrique du conduit vocal pour calculer les fonctions d'aire associées et coupler ces données à un modèle acoustique.

Notons dès à présent que la description géométrique de la forme du conduit vocal passe nécessairement par sa connaissance tridimensionnelle. Un grand nombre d'études se limitent néanmoins à une description bidimensionnelle dans le plan médiosagittal pour lequel les mesures et les analyses articulatoires apparaissent nettement plus accessibles. Un certain nombre de transformations *ad hoc* existent pour passer d'une distance transverse dans le plan médiosagittal à l'aire transverse associée (voir par exemple [Heinz et Stevens, 1965], [Sundberg *et al.*, 1987], [Baer *et al.*, 1991], [Perrier *et al.*, 1992] ou [Beautemps *et al.*, 2001]). Néanmoins le conduit vocal dans le plan médiosagittal peut ne pas refléter son comportement latéral, comme c'est par exemple le cas pour une articulation de la consonne /l/ pour laquelle une occlusion se produit dans le plan médiosagittal et deux conduits latéraux se forment de part et d'autre de la langue. Nous préférons donc une approche tridimensionnelle qui nous permettra de nous affranchir de l'approximation d'une transformation du plan médiosagittal vers le conduit tridimensionnel.

On peut distinguer trois types de modèles articulatoires (d'après [Bailly *et al.*, 2007]) :

Les modèles géométriques : la forme du conduit vocal est obtenue par une description géométrique et les déformations associées sont définies *a priori* ; les formes obtenues sont ensuite ajustées *a posteriori* sur les données anatomiques du conduit vocal. La relative simplicité de cette modélisation en a fait sa force : les premiers modèles articulatoires furent

de ce type. On peut citer [Coker, 1967] comme premier modèle articulatoire 2D complet. D'autres modèles basés sur ce principe seront développés par la suite : par exemple [Liljencrants, 1971], [Mermelstein, 1973] ou [Rubin *et al.*, 1981]. Peu de modèles 3D de ce type existent. Sur la base de données ultrasons de [Stone et Lundberg, 1996], on peut par exemple citer le travail de [Cohen *et al.*, 1998] sur la langue.

Les modèles biomécaniques : les mouvements et les formes du conduit vocal résultent de l'action de muscles ; notons que les déformations associées sont limitées par des contraintes mécaniques de contact entre les articulateurs ; les fibres musculaires sont les éléments de base d'une telle modélisation. Les formes obtenues sont également ajustées ensuite *a posteriori* sur les données anatomiques du conduit vocal. En 2D, depuis les travaux de [Perkell, 1974] qui ont posé les bases de cette approche, on peut citer le modèle de langue de [Payan et Perrier, 1997]. L'approche biomécanique se prête néanmoins plutôt à la modélisation tridimensionnelle. Un certain nombre de modèles tridimensionnels ont été développés : [Kiritani *et al.*, 1976], [Hashimoto et Suga, 1986], [Wilhelms-Tricarico, 1995], [Gérard *et al.*, 2003] ou [Dang et Honda, 2004].

Les modèles fonctionnels : les formes et les mouvements du conduit vocal résultent de l'analyse statistique de corpus de données de formes du conduit vocal associées à des tâches spécifiques. Contrairement aux modèles géométriques et biomécaniques qui doivent être ajustés sur les données *a posteriori*, ce type de modèle est fondé sur les données articulatoires mesurées sur un ou plusieurs sujets et les mouvements sont issus de l'analyse statistique de ces données. Ces modèles présentent l'avantage d'être plus abordables que les modèles biomécaniques tout en conservant un degré de réalisme beaucoup plus élevé que les modèles géométriques. Enfin, les mouvements issus de ces modèles fonctionnels émergent directement des données. Notons cependant que contrairement aux modèles biomécaniques, ces mouvements sont par construction liés aux tâches spécifiques pour lesquelles les données ont été enregistrées, ce qui souligne le rôle crucial du choix du corpus d'enregistrement. Comme pour les deux approches précédentes, les premiers modèles furent médiosagittaux, avec un premier modèle de [Lindblom et Sundberg, 1971]. [Harshman *et al.*, 1977] proposent une méthode (l'analyse *PARAFAC*) qui permet de prendre en compte des données enregistrées sur plusieurs sujets. [Hoole, 1999] se base sur cette approche pour construire un modèle de langue. [Maeda, 1978] propose une analyse en composantes principales pour faire émerger des données les principaux mouvements des articulateurs : les corrélations entre les mesures articulatoires faites sur les contours des organes sont exploitées pour obtenir les paramètres de contrôle du modèle, très souvent de l'ordre de un ou deux pour chaque organe, un peu plus pour la langue. Cette approche est améliorée l'année suivante ([Maeda, 1979]) pour introduire dans l'analyse une connaissance *a priori* du mouvement des articulateurs en vue d'obtenir une décomposition en mouvements plus réalistes. Cette méthode est à la base de l'élaboration de plusieurs modèles articulatoires (citons en 2D [Maeda, 1990] ou [Beautemps *et al.*, 2001]). Il existe peu de modèles articulatoires fonctionnels tri-

dimensionnels. Quelques modèles de conduit, c'est-à-dire en considérant le conduit vocal comme un tuyau dont on cherche à modéliser la forme et non pas celle des organes qui le bordent, basés sur une analyse en composantes principales, ont été développés par [Yehia et Tiede, 1997], [Badin *et al.*, 1998] et [Engwall, 1999]. Plus récemment, un modèle tridimensionnel de langue, toujours basé sur une analyse en composantes principales, a été développé par [Engwall, 2000] et [Badin *et al.*, 2002]. Enfin, un modèle de langue plus précis et plus complet a été développé sur des principes similaires à ceux utilisés dans le présent travail ([Badin et Serrurier, 2006]).

La modélisation articulatoire permet d'obtenir des descriptions géométriques du conduit vocal en 2D ou en 3D. La fonction d'aire du conduit vocal peut alors être calculée soit directement à partir du 3D, soit par l'intermédiaire de fonctions de passage du 2D en 3D. La complexité du processus de modélisation articulatoire a conduit un certain nombre d'études à s'intéresser directement à la modélisation des fonctions d'aire sans modéliser le conduit vocal complet (par exemple [House et Stevens, 1956], [Maeda, 1982] ou [Feng et Castelli, 1996]). Les fonctions d'aire obtenues sont ensuite utilisées pour calculer les résonances du conduit par un modèle acoustique ([Fant, 1960]). Cette méthode permet d'étudier les relations entre les formants et les formes du conduit vocal et lorsque c'est possible d'affilier les cavités et les formants ([Majid, 1986], [Badin *et al.*, 1990]). L'utilisation de modèles articulatoires permet alors de caractériser acoustiquement chaque paramètre de contrôle.

Notre objectif est d'étudier les relations articulatoire-acoustiques du port vélopharyngé, et plus particulièrement du voile du palais, à l'aide d'un modèle articulatoire. Avant de détailler l'approche adoptée dans ce travail, une revue de différentes études sur ce sujet est proposée au paragraphe suivant.

1.3.2 Modélisation articulatoire-acoustique des nasales

1.3.2.1 Les données

Pas de modèle sans données : c'est la règle fondamentale pour l'établissement de modèles articulatoires et/ou de fonctions d'aire du conduit nasal. [Fritzell, 1969] nous offre un historique très complet de l'étude du port vélopharyngé depuis la Grèce ancienne à nos jours. Après les fondements de l'étude de la parole posés par Hippocrate et Aristote, Galen (129-199) mentionne pour la première fois le voile du palais, et plus précisément la luvette, comme utile en parole, « contribuant au volume et à la beauté de la voix ». Après le Moyen-Âge, l'étude de la voix et du voile du palais fait sa réapparition lors de la renaissance italienne : Leonard de Vinci (1452-1519) dessine pour la première fois le « palato mobile », Fallappio (1561) distingue le palais dur du palais mou et Fabricius ab Aquapendente souligne en 1603 que certains sons passent par le nez. C'est au XVII^{ème} siècle que l'ouverture et la fermeture du port vélopharyngé apparaît contrôlée par le voile du palais, notamment grâce à Amman (1700). Von Kempelen donne d'ailleurs au voile du palais un rôle important dans sa machine parlante (1791) et souligne que la fermeture nasale est atteinte « par apposition du voile du palais sur la paroi arrière du pharynx ». La première description du fonctionnement du port vélopharyngé lors de la respi-

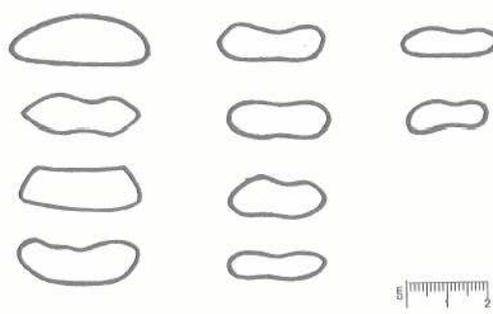


FIG. 1.9 – Sections transverses de couplage du port vélopharyngé pour 10 sujets (d’après [Björk *et al.*, 1961] reproduites par [Reenen, 1982]). Le voile du palais est dans une position de repos.

ration et de la déglutition est réalisée par Hilton (1836) grâce à un sujet possédant un « défaut facial ». Il indique que le voile du palais reste en position haute en parole pour « diriger l’air à travers la bouche ». Cette période marque le début d’un grand nombre d’études sur des patients possédant un défaut maxillo-facial. La seconde moitié du XIX^{ième} siècle voit le développement de la phonétique expérimentale grâce à laquelle la compréhension du fonctionnement du port vélopharyngé progresse. Citons par exemple les travaux de Czermak (1869) qui détermine plusieurs hauteurs du voile du palais en fonction de la voyelle considérée et ceux de [Passavant, 1869] sur la paroi du pharynx. La découverte des rayons-X à la fin du XIX^{ième} siècle permet de faire un lien avec l’époque moderne. Les premières, et uniques à notre connaissance, coupes transverses obtenues par rayons-X viennent de [Björk, 1961], résultats rapportés par [Björk *et al.*, 1961] (voir la figure 1.9). [Björk *et al.*, 1961] étudient également le lien entre l’ouverture du port vélopharyngé dans le plan médiosagittal et l’aire de couplage transverse : pour une distance sagittale supérieure à 0.2 cm, la relation est linéaire. Il faudra ensuite attendre jusqu’à très récemment l’émergence de l’Imagerie par Résonance Magnétique (IRM) pour obtenir une étude plus systématique de l’ouverture du port vélopharyngé ([Demolin *et al.*, 1996a], [Demolin *et al.*, 1998] et [Demolin *et al.*, 2003]). Grâce à des enregistrements d’IRM transverses, [Demolin *et al.*, 2003] mesurent l’ouverture du port vélopharyngé pour quatre sujets prononçant les quatre voyelles du français. La figure 1.10 illustre une image médiosagittale pour un sujet (sujet 1) et les mesures obtenues pour un autre (sujet 2).

La figure 1.10 montre les mesures obtenues pour le sujet 2.

Les études de la géométrie et de la fonction d’aire des fosses nasales restent en nombre assez limité. Deux études qui ont marqué l’histoire de la modélisation des nasales remontent autour de 1960. La première étude connue est celle de [House et Stevens, 1956] qui infèrent les fonctions d’aire des fosses nasales à partir d’atlas anatomiques, de représentations crâniennes et d’images aux rayons-X latérales. Ils soulignent notamment la difficulté à obtenir les représentations exactes des fosses nasales. Le résultat est affiché à la figure 1.11. La particularité de leur étude est le seuil narinaire fixé à 0.23 cm^2 qui sera largement discuté dans la littérature. La seconde étude quelques années plus tard, détaillée par [Bjuggren et Fant, 1964] mais déjà publiée par [Fant, 1960], est basée sur un moulage des fosses nasales effectué sur un cadavre. Le moulage est ensuite découpé

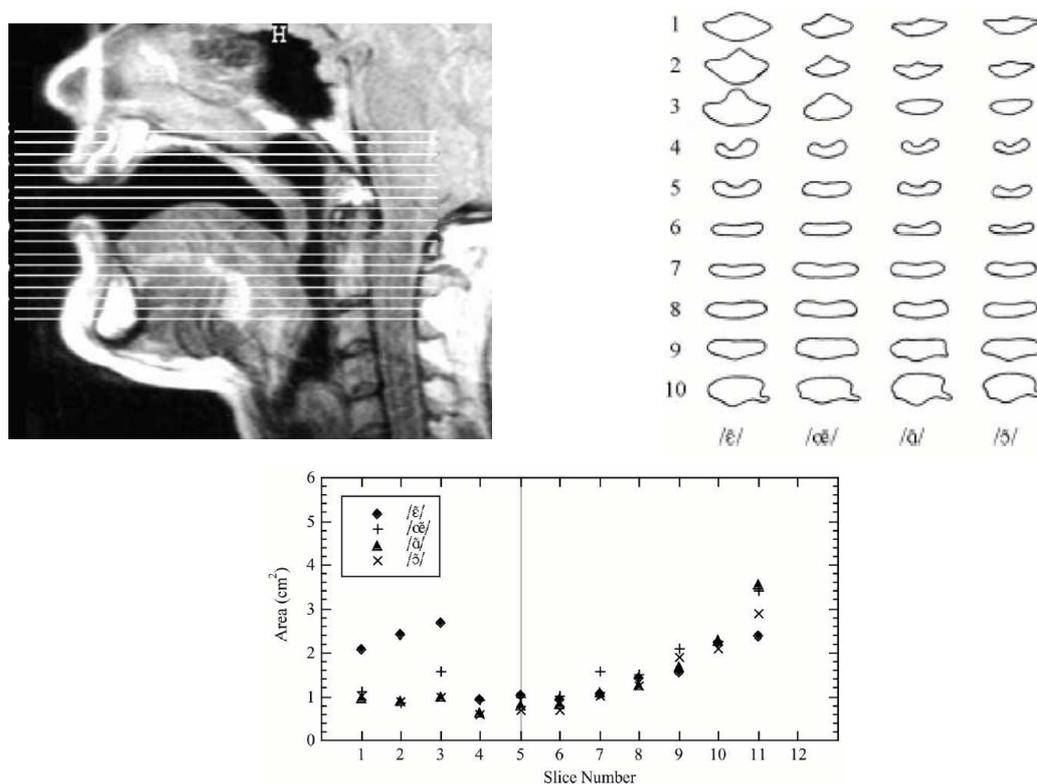


FIG. 1.10 – Image IRM médiosagittale du sujet 1 (haut gauche), sections transverses numérotées 1 à 10 du sujet 2 (haut droite) et aires transverses du sujet 2 correspondantes (bas) pour les quatre voyelles nasales du français dans l'étude de [Demolin *et al.*, 2003]. La section numérotée 1 correspond à la coupe la plus basse, la section numérotée 11 à la plus haute.

à la scie à intervalles réguliers pour obtenir un échantillonnage de l'aire des fosses nasales et la fonction d'aire associée (voir la figure 1.12). La géométrie obtenue est donc très précise et la circonférence complexe de chaque section est calculée et comparée à un cercle : dans la région centrale à mi-distance entre le cavum et les narines la circonférence des sections est 3.5 fois plus importante que celle d'un cercle possédant la même aire ; c'est une mesure de la complexité du conduit. La dissymétrie entre les fosses droite et gauche est également évoquée. La fonction d'aire obtenue est montrée à la figure 1.11. Cette fonction d'aire servira de référence pendant très longtemps pour l'étude des nasales. Les très importantes valeurs obtenues seront discutées par la suite dans la littérature. Il faudra attendre à nouveau l'émergence de l'Imagerie par Résonance Magnétique pour renouer avec les mesures des fosses nasales. [Matsumura, 1992] et [Matsumura *et al.*, 1994] proposent de nouvelles fonctions d'aire obtenues par IRM coronales sur plusieurs sujets vivants (voir la figure 1.11 pour un exemple sur un des sujets). Enfin, [Dang *et al.*, 1994] réalisent une étude très complète à partir d'enregistrement IRM coronaux sur plusieurs sujets. Un exemple est donné à la figure 1.11.

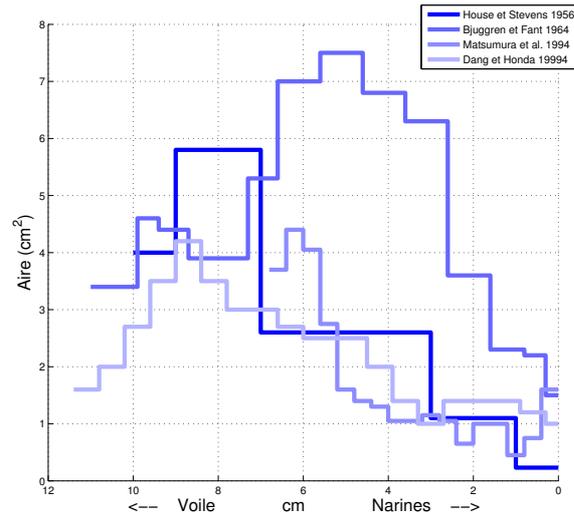


FIG. 1.11 – Fonctions d'aire des fosses nasales trouvées dans la littérature (adapté de [Rossato, 2000]).

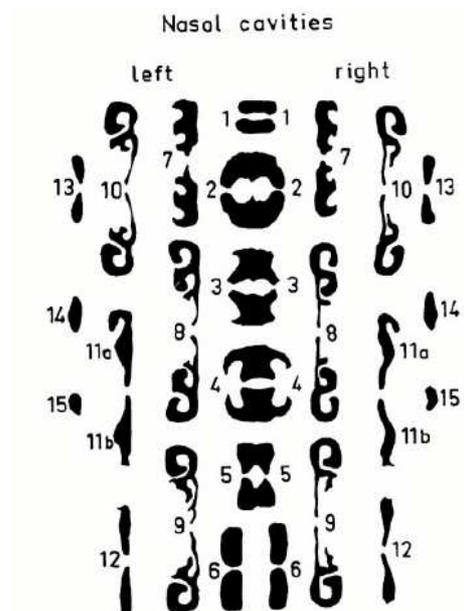


FIG. 1.12 – Sections transversales des fosses nasales obtenues par [Bjuggren et Fant, 1964] en découpant un moulage. Les sections sont numérotées par ordre croissant du cavum vers les narines. Chaque section est représentée par deux contours, ceux de chaque côté de la scie.

1.3.2.2 Les modèles articulatoires du voile du palais et du port vélopharyngé

Il n'existe à proprement parler aucun modèle articulatoire complet du conduit nasal. Notons cependant le modèle de [Mermelstein, 1973] qui décrit le déplacement de la pointe de la luette dans le plan médiosagittal selon un segment de droite déterminé par les deux positions

extrêmes haute et basse du voile du palais définies empiriquement. L'aire de couplage du port vélopharyngé est alors considérée comme proportionnelle au carré de la distance entre la limite supérieure du segment et la position effective de la pointe de la luvette. Notons que cette relation en carré d'une distance médiosagittale n'est pas tout à fait cohérente avec la relation linéaire de [Björk *et al.*, 1961]. Ce modèle a servi de base à quelques études de relations articulatoire-acoustiques des nasales comme celle de [Bell-Berti et Baer, 1983] ou de [Teixeira *et al.*, 2001] qui l'utilisent ensuite pour des tests perceptifs du portugais basés sur la synthèse articulatoire ([Teixeira *et al.*, 2003]).

Le seul véritable modèle de voile du palais est celui de [Rossato, 2000]. Il constitue la première tentative effectuée à l'ICP à l'origine de ce travail. Par une analyse en composantes principales sur un jeu de contours médiosagittaux du voile du palais obtenus sur des images IRM, deux paramètres de contrôle émergent des données : le premier explique 88 % de la variance du voile du palais, le second 8 %. Le voile du palais apparaît donc contrôlé principalement par un unique paramètre. Les caractéristiques de ce modèle sont détaillées plus loin dans ce document (au paragraphe 3.4.1). L'aire de couplage du port vélopharyngé est considérée comme proportionnelle à la valeur d'un paramètre articulatoire représentant la hauteur du voile du palais.

Pour résumer, sans constituer de modèle à proprement du voile du palais, l'aire de couplage du port vélopharyngé peut être estimée par plusieurs méthodes. On a vu ci-dessus le principe qui consiste à trouver une relation plus ou moins réaliste entre l'aire de couplage et une mesure de la position du voile du palais dans le plan médiosagittal : relation linéaire avec le carré de la distance entre la pointe de la luvette et la position de celle-ci en cas de fermeture du conduit nasal ([Mermelstein, 1973]), relation linéaire avec la distance sagittale entre le voile du palais et la paroi du pharynx ([Björk *et al.*, 1961]), combinaison de ces deux relations suivant le degré de fermeture du port vélopharyngé ([Reenen, 1982]) ou relation linéaire avec la valeur d'un paramètre articulatoire ([Rossato, 2000]).

1.3.2.3 Les études articulatoire-acoustiques

Si aucun modèle articulatoire complet du conduit nasal n'existe, un certain nombre d'études par simulation directement basées sur les fonctions d'aire portent sur les relations entre l'articulatoire et l'acoustique. L'objectif de ces simulations est d'expliquer le comportement acoustique observé dans les enregistrements de voyelles et consonnes nasales et de trouver les corrélats acoustiques de la nasalité. La première est celle de [House et Stevens, 1956], basée sur les fonctions d'aire détaillées ci-dessus. Les aires des trois premières sections varient pour simuler la variation de l'aire du port vélopharyngé. [Fant, 1960], sur la base de ses fonctions d'aire décrites ci-dessus, effectue quelques simulations en faisant varier les aires de couplage du port vélopharyngé et du seuil narinaire. Les évolutions des pôles et des zéros induites par ces modifications sont détaillées. [Fant, 1985] donne les résonances obtenues en rajoutant les sinus. [Hecker, 1962] développe un synthétiseur articulatoire pour étudier les consonnes nasales. Les fonctions d'aire adoptées sont proches de celles de [House et Stevens, 1956] avec les deux premières sections

d'aire variable pour simuler la variation de l'aire vélopharyngée. Des tests perceptifs sont ensuite conduits pour étudier l'influence du couplage nasal. De même que [Hecker, 1962], [Mrayati, 1976] s'appuie sur les fonctions d'aire de [House et Stevens, 1956] pour simuler pour la première fois les quatre voyelles nasales du français. Des fonctions de sensibilité succinctes sont fournies par [Guérin et Mrayati, 1977] un an plus tard. Une étude importante pour l'étude des relations articulatoire-acoustiques est celle de [Maeda, 1982] : les résonances des sinus maxillaires comme résonateurs de Helmholtz sont prises en compte dans le calcul des fonctions de transfert et l'influence de l'abaissement du voile du palais sur la fonction d'aire orale est également prise en compte. Les aires des trois premières sections de la fonction d'aire du conduit nasal, obtenu à partir des valeurs de [House et Stevens, 1956] et [Fant, 1960], varient linéairement entre l'aire de couplage du port vélopharyngé, variable, et l'aire du cavum considérée comme fixe ; parallèlement, l'aire de ces trois sections du port vélopharyngé est soustraite à l'aire des trois sections correspondantes du conduit oral lorsque le conduit nasal est fermé. La somme des aires des trois sections orales et nasales dans la région vélaire reste donc fixe et le déplacement du voile du palais assure une diminution de l'aire orale lorsque l'aire vélopharyngée augmente et vice-versa. Des tests perceptifs ont été menés par [Maeda, 1993] en utilisant ce modèle pour synthétiser des sons. Cette approche fut par ailleurs considérée dans une série de travaux réalisés à l'ICP sur les relations articulatoire-acoustiques de l'abaissement du voile du palais : [Lherm, 1984], [Feng, 1986], [Castelli, 1989] et [Rossato, 2000]. Notons également le travail de [Merlier, 1984] sur le couplage externe entre le conduit nasal et le conduit oral. [Feng, 1986] s'est attaché à considérer la nasalité comme le passage du conduit vocal d'une configuration purement orale (c'est-à-dire avec une fermeture du conduit nasal obtenue avec le voile du palais en position haute) à une configuration purement nasopharyngée (c'est-à-dire avec une occlusion du conduit oral obtenue avec le voile du palais en position basse) et a étudié les conséquences formantiques de cette évolution. La majeure partie de ces travaux est résumée dans l'article de [Feng et Castelli, 1996]. Ces études sont reprises par [Castelli, 1989] qui tente de trouver l'origine de la résonance du formant nasal à 250 Hz. Enfin, [Rossato, 2000] tente d'inverser le signal de parole pour retrouver la position du voile du palais. Certains points de tous ces travaux seront repris au cours de ce document, notamment dans le chapitre 5 qui étudie les effets acoustiques de notre modèle de port vélopharyngé.

Toutes les études présentées ci-dessus sont menées sur la base de modèles articulatoires ou de fonctions d'aire plus ou moins réalistes. Le manque de modèle articulatoire réaliste du voile du palais et du conduit nasal en général fait cruellement défaut. Cette thèse vise à compléter ce manque et la section suivante présente notre approche pour atteindre cet objectif.

1.4 Approche modélisatrice

Le développement de modèles de production de la parole – et particulièrement des nasales – plus réalistes suppose la conception de modèles articulatoires tridimensionnels du port vélopharyngé et des cavités nasales plus détaillés. Nous nous plaçons pour cela dans une approche résolument fonctionnelle. Pour concevoir un modèle, le compromis entre le réalisme de

la description du mécanisme et la simplicité du modèle nécessite de faire un certain nombre de choix, que nous détaillons ci-dessous.

Modélisation fonctionnelle Deux objectifs majeurs guident l'élaboration du modèle :

- Comprendre le fonctionnement du port vélopharyngé : sa forme, son mouvement, les co-articulations avec les autres entités du conduit vocal et son rôle acoustique.
- Obtenir une description géométrique précise du conduit nasal en vue de coupler ce modèle à un modèle acoustique pour générer des fonctions de transfert acoustiques du conduit et synthétiser des sons.

Les approches géométriques et biomécaniques supposent l'introduction d'une connaissance *a priori* sur le mouvement des organes. Au contraire, notre approche s'inscrit parfaitement dans une approche fonctionnelle : faire émerger les différents mouvements indépendants des données par une analyse statistique et confronter les données au modèle ; soulignons au passage le choix crucial du corpus d'enregistrement des données (détaillé au chapitre 2).

Modélisation tridimensionnelle Le réalisme du modèle est un critère important. Les techniques qui consistent à calculer la fonction d'aire à partir de contours médiosagittaux du conduit vocal n'apportent pas la précision souhaitée pour notre modèle et ne peuvent pas résoudre les cas limites. C'est le cas par exemple du voile du palais pour lequel la luette, présente dans le plan médiosagittal et son voisinage, ne représente pas la géométrie du voile du palais dès lors qu'on s'écarte du voisinage du plan médiosagittal vers les régions latérales. Pour certaines articulations nasales, comme les voyelles nasales arrières du français, la luette peut être en contact à la fois avec le dos de la langue (voir [Demolin *et al.*, 2003]) et la paroi du pharynx dans le plan médiosagittal (voir la figure 1.13), conduisant ainsi à une occlusion médiosagittale ; les canaux de chaque côté de cette occlusion restent néanmoins ouverts et doivent être considérés par une description géométrique tridimensionnelle. Enfin, les cavités nasales sont séparées par une cloison dans le plan médiosagittal, ce qui implique nécessairement de les considérer en trois dimensions. Ces raisons nous conduisent naturellement à opter pour une modélisation tridimensionnelle plutôt que médiosagittale.

Modélisation de surfaces d'organes La question de l'entité que l'on souhaite modéliser se pose également. Deux options s'offrent à nous : modéliser le *conduit vocal*, c'est-à-dire le conduit résultant de la connexion des divers articulateurs de la parole mais sans référence explicite aux articulateurs eux-mêmes, ou modéliser directement les *articulateurs*. Un de nos objectifs est de comprendre le fonctionnement du conduit vocal et plus précisément du port vélopharyngé. La forme du conduit vocal est obtenue par déformation des divers articulateurs qui constituent les limites. En termes fonctionnels, les articulateurs sont donc les éléments de base constitutifs du conduit vocal. Nous nous intéressons donc à la modélisation des articulateurs. En outre, la déformation du voile du palais modifie conjointement les conduits oral et nasal : c'est bien son mouvement et son influence sur chacun des conduits qu'il convient d'étudier et non les conduits comme deux entités indépendantes. Les articulateurs sont généralement considérés comme des organes (comme la langue, le voile du palais, la mâchoire, etc.) et seront désignés dans la suite du

document indifféremment comme des *articulateurs* ou des *organes*. Le voile du palais et la paroi du pharynx, qui forment les limites du port vélopharyngé, sont considérés à ce titre comme des « organes ». Enfin, une approche basée sur les organes permet un plus grand réalisme dans la détermination de la géométrie du conduit vocal. Pour terminer, dans le but d'obtenir un modèle précis et fonctionnel, nous ne chercherons pas à modéliser le volume de chaque organe mais sa surface, ce qui suffit à décrire le fonctionnement et obtenir des fonctions d'aire réalistes sans alourdir le modèle.

Modélisation monosujet Puisque cette étude constitue une première tentative pour élaborer un modèle articulatoire tridimensionnel, un seul sujet a été considéré : nous avons choisi le locuteur français déjà impliqué dans le développement d'un modèle articulatoire médiosagittal ([Beautemps *et al.*, 2001]) et de modèles tridimensionnels de langue, de lèvres et de visage ([Badin *et al.*, 2002]). De nombreuses données ont par ailleurs été enregistrées sur lui depuis de nombreuses années pour diverses études, notamment sur la nasalité ([Castelli, 1989], [Rossato *et al.*, 2003], [Feng et Kotenkoff, 2004]). Enfin, ce sujet n'est autre que l'encadrant de cette thèse et sa motivation pour servir de cobaye est à toute épreuve !



Notre approche ainsi définie, les chapitres suivants détaillent chaque étape du processus de modélisation articulatoire en direction de la caractérisation articulatoire-acoustique du modèle : le chapitre 2 décrit les données articulatoires et les traitements en vue d'obtenir des formes tridimensionnelles des organes, le chapitre 3 décrit le modèle proprement dit, le chapitre 4 détaille la méthode de calcul des fonctions d'aire à partir du modèle, et enfin le chapitre 5 explore les caractéristiques acoustiques du modèle.

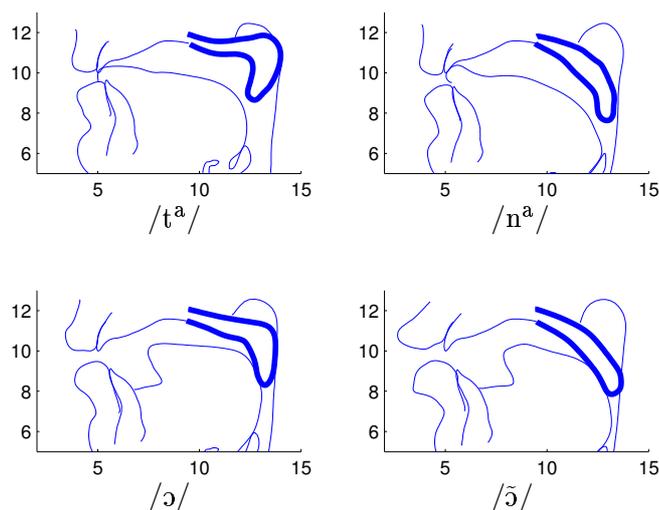


FIG. 1.13 – Contours médiosagittaux du conduit vocal pour les consonnes occlusives orale et nasale $/t^a/$ et $/n^a/$ (haut) et pour les voyelles orale et nasale $/ɔ/$ et $/õ/$ (bas). Le contour épais représente le voile du palais

Chapitre 2

Données articulatoires

L'acquisition et le traitement de données constituent les fondements de la modélisation articulatoire. Les données sont d'autant plus importantes pour notre travail que nous souhaitons faire émerger le modèle directement à partir de celles-ci sans introduire d'autre connaissance *a priori*. Les résultats de ce chapitre constituent donc un maillon essentiel de la chaîne de modélisation. Après un bref aperçu des différentes techniques d'exploration du conduit vocal, nous présentons dans ce chapitre les méthodes d'acquisition et le prétraitement des données utiles par la suite dans le développement du modèle articulatoire. Ces traitements ont pour objectif de construire une base de données de formes tridimensionnelles des articulateurs rigides ou déformables du conduit vocal. Cette base de données sera exploitée pour développer le modèle articulatoire détaillé au chapitre suivant.

2.1 Techniques d'exploration du conduit vocal

Un grand nombre de techniques permettent l'acquisition de données relatives au conduit vocal : de type physiologique (électromyographie, etc.), articulatoire (cinéradiographie, électropalato-graphie, ultrasons, articulographie, vidéo, imagerie médicale, etc.) ou encore acoustiques et aéroacoustiques (mesures de pressions, débits, rhinométrie, etc.). Cette section s'attache à décrire brièvement les caractéristiques des techniques utilisées pour notre étude. Pour une description plus approfondie et un aperçu des autres techniques possibles, il est possible de se référer à [Hardcastle et Laver, 1997] ou à [Baken, 1987] et [Krakow et Huffman, 1993] pour le cas de la nasalité.

L'exploration du conduit vocal et par conséquent l'acquisition de données constitue en modélisation articulatoire un enjeu particulièrement important. Deux caractéristiques principales des données sont à prendre en compte : les résolutions temporelle et spatiale. Les données à faible résolution temporelle (que l'on assimilera à des données *statiques*), et à bonne résolution spatiale dans notre cas, permettent d'obtenir des informations sur le conduit vocal en position fixe, comme lors de la production d'une voyelle ou d'une consonne tenue dans un certain contexte vocalique. Elles servent à mesurer le conduit vocal pour le phonème considéré. À l'inverse, les données à forte résolution temporelle (que l'on assimilera à des données *dynamiques*) permettent un suivi temporel de certains paramètres liés à la production de la parole, comme la position de certains points de surface d'articulateurs, la pression intra-orale, etc. Elles peuvent être utilisées pour étudier la dynamique du conduit vocal : coarticulation, mouvement d'un articulateur particulier, etc. Cette section dresse une présentation des différents types de données utilisées lors de cette étude, ainsi que leurs caractéristiques.

Moulages dentaires (statique) : Il s'agit de moulages couramment utilisés en dentisterie pour modeler les formes des dentitions supérieure et inférieure ainsi que du palais dur. Ils permettent d'obtenir très précisément les formes des surfaces de cette partie du conduit vocal. La figure 2.1 représente un moulage de la partie supérieure.

Images IRM statique : C'est à ce jour la technique la plus couramment utilisée pour obtenir la forme du conduit vocal dans une position fixe. Elle consiste à soumettre le sujet à un



FIG. 2.1 – Données articulatoires statiques : moulage buccal supérieur (gauche), image IRM médiosagittale pour une articulation /a/ (milieu) et image CT axiale pour une position de repos (droite)

champ magnétique élevé qui va orienter le moment magnétique de l'unique proton du noyau des atomes d'hydrogène et à perturber cette orientation forcée par un gradient d'impulsions de champ magnétique dans un plan de coupe donné. Le signal d'énergie généré par le retour des moments magnétiques à leur état d'équilibre forcé est le signal de Résonance Magnétique Nucléaire. Il mesure indirectement spatialement la densité de protons des tissus sur la coupe considérée. L'affichage de niveaux de gris proportionnels à cette densité en fonction de la position dans l'espace forme l'Image par Résonance Magnétique nucléaire (IRM) (voir par exemple [Liang et Lauterbur, 2000]). Un exemple est montré à la figure figure 2.1. Cette technique permet de différencier des tissus n'ayant pas la même densité de protons d'hydrogène. Considérée aujourd'hui comme inoffensive pour le sujet, cette technique est donc largement utilisée pour obtenir la forme du conduit vocal, tant dans le plan médiosagittal que dans d'autres plans sagittaux ou même transverses. Inventée en 1972, son utilisation dans le domaine de la parole et de l'étude géométrique de la forme du conduit vocal s'est développée dans les années quatre-vingt-dix. Cette technique constitue une source d'information essentielle sur la géométrie du conduit vocal ; pour preuve, le nombre d'études en parole centrées sur l'IRM (voir par exemple [Rokkaku *et al.*, 1986], [Baer *et al.*, 1991], [Dang *et al.*, 1994], [Story *et al.*, 1996], [Badin *et al.*, 2002] ou [Demolin *et al.*, 2003]). La possibilité d'obtenir une description spatiale complète de la tête constitue notamment un atout précieux pour l'exploration tridimensionnelle. L'enregistrement d'images IRM présente cependant quelques désavantages qu'il convient d'indiquer. L'un des principaux problèmes des images IRM réside dans l'inexistence de contraste air/os : tous deux faiblement denses en protons d'hydrogène, ils sont indissociables sur les images, ce qui apparaît particulièrement problématique dans la cavité buccale pour distinguer le conduit oral des dents (voir l'image IRM de la figure 2.1). Le second point à noter est le protocole d'enregistrement relativement long pour le sujet, au cours duquel il ne doit pas bouger : plusieurs dizaines de secondes par exemple pour obtenir 25 images sagittales parallèles de la tête espacées de 4 mm chacune. En plus d'une contrainte fastidieuse, la qualité du bloc d'images dépend directement de la capacité du sujet à rester fixe. Il convient donc de faire un compromis entre qualité souhaitée des images (enregistrement d'autant plus long), et confort du sujet. La taille du corpus est bien

évidemment un élément à prendre en compte. Enfin, l'enregistrement s'effectue en position allongée. L'effet de la gravité lorsque le sujet est en position allongée peut modifier légèrement la forme du voile du palais ([Kitamura *et al.*, 2005]) et son activité musculaire ([Moon et Canady, 1995]), bien que les effets sur la géométrie du conduit vocal puissent être considérés comme relativement faibles ([Tiede *et al.*, 2000]). [Kent, 1997] mentionne malgré tout anecdotiquement quelques études indiquant une nasalité légèrement affectée pour les astronautes en apesanteur ! Pour terminer ce paragraphe, je souhaiterais remercier Sylvain Negro dont le rapport de stage m'a fortement aidé à comprendre le mécanisme de l'IRM.

Images tomodensitométriques (CT, statique) : Il s'agit d'un scanner à rayons X sur des plans de coupes de la tête. Les images obtenues permettent notamment contrairement aux images IRM d'identifier les structures osseuses mais pas de discriminer les tissus mous entre eux. L'acquisition s'effectue en position allongée et fixe. De même que pour les images IRM, on obtient une description spatiale tridimensionnelle complète de la tête, dont la qualité est cependant bien supérieure. La figure 2.1 représente une coupe médiosagittale de la tête. Contrairement à l'image IRM les structures osseuses sont particulièrement visibles en blanc tandis qu'il n'est pas possible de distinguer les tissus mous entre eux. Cette technique apparaît donc complémentaire à l'IRM. À base de rayons X, elle est cependant considérée comme nocive pour le sujet et ne peut donc être considérée pour des enregistrements répétitifs.

Suivi de points par Articulographie Électromagnétique (EMA, dynamique) : L'articulographie électromagnétique permet de suivre l'évolution temporelle de la position de plusieurs points de chair du conduit vocal dans le plan médiosagittal ([Perkell *et al.*, 1992], [Hoole et Nguyen, 1997]). Le sujet porte un casque sur lequel sont attachées trois bobines émettrices de champs magnétiques à des fréquences distinctes. Une ou plusieurs bobines réceptrices sont collées sur les surfaces des organes dont on souhaite suivre l'évolution au cours du temps. Ces bobines sont par ailleurs reliées chacune par un fil à une unité réceptrice directement connectée à un ordinateur. La puissance du champ magnétique enregistrée par une bobine réceptrice à une fréquence donnée est proportionnelle à sa distance à la bobine émettrice correspondant à la fréquence considérée. Par triangulation la position de la bobine réceptrice est alors connue par rapport aux positions des trois bobines du casque. La fréquence d'échantillonnage temporel pouvant s'élever à plusieurs kilohertz, ce système permet un suivi précis des points de chair avec une résolution spatiale inférieure au millimètre. La figure 2.2 montre l'auteur de ce document sur lequel est installé l'articulographe électromagnétique et cinq bobines collées (deux sur les lèvres, trois sur la langue et une sur le nez pour avoir une référence fixe). Il est intéressant de noter que d'éventuels mouvements de tête du locuteur n'introduisent pas d'erreur dans le calcul des positions des bobines qui sont des valeurs relatives dépendantes de la position du casque, et donc de la tête.

Chaque technique a donc une ou plusieurs spécificités qui constitue un élément important dans la construction du modèle articulatoire. Les enregistrements des données utiles pour notre étude par l'intermédiaire de ces différentes techniques sont détaillés dans la section suivante.

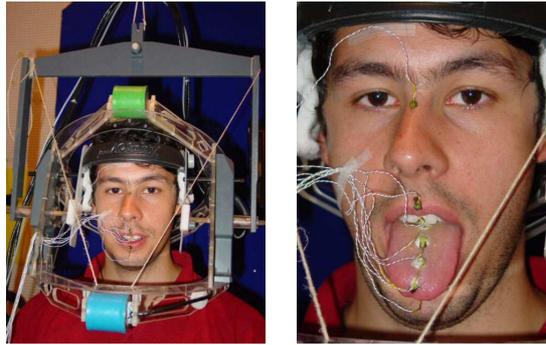


FIG. 2.2 – Sujet sur lequel est installé un casque d’articulographe électromagnétique et collé cinq bobines réceptrices.

2.2 Les données articulatoires

Plusieurs types de données de nature différente sont enregistrées pour des usages spécifiques au cours de cette étude. La représentation des données ne correspond cependant pas toujours exactement aux caractéristiques requises pour l’usage auxquelles elles sont destinées. C’est le cas notamment des images dont l’orientation des plans de coupe dépend principalement de l’outil d’enregistrement et peut ne pas correspondre à l’usage spécifique que nous en faisons dans cette étude. Nous présentons donc également dans cette section, en plus de l’acquisition, un redécoupage des images correspondant à une utilisation plus appropriée à nos besoins.

2.2.1 Le corpus

L’élaboration d’un corpus et l’enregistrement approprié des données constitue naturellement la première étape importante d’une modélisation articulatoire basée sur les données. Puisque le principe sous-jacent de la modélisation est de faire émerger des données les mouvements des articulateurs dont on souhaite faire un modèle, le corpus doit constituer un échantillonnage représentatif de l’espace articulatoire. Une façon d’atteindre cet objectif est d’inclure dans le corpus toutes les articulations que le sujet peut produire dans la langue. Le corpus a donc été constitué du jeu de cibles articulatoires suivant : les dix voyelles orales du français [a ε e i y u ø ɔ œ o], les quatre voyelles nasales du français [ã, õ, ě, õ̃] (que notre locuteur peut produire), et les consonnes artificiellement soutenues [p t k f s ʃ m n ʁ l] produites dans les trois contextes symétriques [a i u], formant ainsi 46 articulations cibles avec les articulations *repos* et *préphonation*. Ce corpus limité est néanmoins suffisant pour développer des modèles articulatoires médiosagittaux avec une précision presque semblable à celle obtenus avec des corpus quarante fois plus grands ([Beautemps *et al.*, 2001]). [Engwall et Beskow, 2003] ont également montré qu’un faible corpus équilibré suffisait pour développer un modèle articulatoire fiable, même s’il conduisait à des représentations hyper-articulées. Ce corpus sera référencé dans la suite comme le *corpus principal*.

2.2.2 Les moulages dentaires

Deux moulages dentaires ont été réalisés en milieu médical, le premier pour obtenir la forme de la dentition supérieure ainsi que le palais dur, et le second la dentition inférieure. La figure 2.3 montre des photos de ces moulages. Ces deux moulages font chacun corps avec un socle qu'il faudra éliminer. Le moulage supérieur représente la géométrie de la surface limitée par la jonction entre la lèvre et la gencive à l'extérieur et le palais mou à l'intérieur ; le moulage inférieur représente la géométrie de la surface limitée par la jonction entre la lèvre et la gencive à l'extérieur et la base de la langue à l'intérieur.



FIG. 2.3 – Moulages dentaires : les deux moulages vus de côté (gauche), le moulage supérieur (milieu) et le moulage inférieur (droite)

2.2.3 Les images tomодensitométriques

Une série d'images tomодensitométriques (CT) axiales ont été acquises au centre de Radiologie du Centre Hospitalier Universitaire de Grenoble le 28 septembre 2002. Le sujet se tient en position de repos. Les caractéristiques des images sont rassemblées dans le tableau 2.1.

Deux images originales sont visibles à la figure 2.4, l'une représentant une coupe de la tête à peu près au niveau de l'ouverture labiale (gauche), l'autre au niveau d'un plan situé entre les narines et les yeux (droite). Ces images permettent de distinguer très clairement les structures osseuses (en blanc sur la figure 2.4) mais pas les tissus mous entre eux. Par ailleurs, l'empilement des 149 images forme un bloc de voxels 3D qui échantillonnent la tête du sujet ; la bonne résolution d'un voxel de $0.05 \text{ cm} \times 0.13 \text{ cm} \times 0.05 \text{ cm}$ permet d'avoir un échantillonnage spatial

Position du sujet	allongée
Nombre d'images	149
Orientation	axiale
Région d'intérêt	du cou jusqu'au milieu du crâne
Taille	25.6 cm \times 25.6 cm ou 512 pixels \times 512 pixels
Epaisseur	0.13 cm
Résolution	0.05 cm / pixels ou 20 pixels / cm

TAB. 2.1 – Caractéristiques des images CT

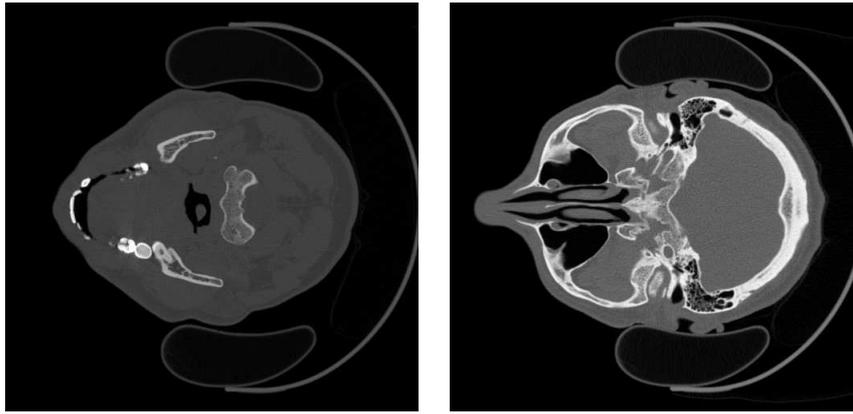


FIG. 2.4 – Exemple d’images CT axiales

Position du sujet	allongée
Nombre d’images	25
Orientation	sagittale
Région d’intérêt	de la base du cou jusqu’en haut de la tête
Taille	25 cm × 25 cm ou 256 pixels × 256 pixels
Epaisseur	0.4 cm
Résolution	0.0977 cm / pixels ou 10.24 pixels / cm
Durée d’enregistrement	environ 35 sec.

TAB. 2.2 – Caractéristiques des images IRM sagittales pour un des 46 blocs

très précis.

2.2.4 Les images IRM

Deux jeux d’images IRM ont été enregistrés pour la construction du modèle articulatoire : un jeu d’images sagittales dont la zone d’intérêt comprend l’ensemble du conduit vocal de la glotte aux lèvres et à l’entrée des fosses nasales et un jeu d’images coronales spécifique pour les fosses nasales dont la géométrie complexe requiert une attention particulière.

2.2.4.1 Les images IRM sagittales

Pour chacune des 46 articulations du corpus, un jeu d’images sagittales a été enregistré à l’aide du scanner IRM Philips GyroScan T10-NT à 1 Tesla du Centre Hospitalier Universitaire de Grenoble le 4 juillet 1998. Les caractéristiques de chacun des 46 jeux d’images sont résumées dans le tableau 2.2. Notons au passage que les résolutions 0.0977 cm / px et 10.24 px / cm seront arrondies à 0.1 cm / px et 10 px / cm dans la suite du document.

Trois images sont visibles sur la figure 2.5. Elles représentent des coupes du plan médiosagittal (milieu) et de plans éloignés de 1.6 cm de part et d’autre de ce plan (gauche et droite) pour

le locuteur prononçant le phonème /a/. Six images IRM espacées de 0.4 cm à partir du plan médiosagittal sont par ailleurs affichées à la figure 2.28 page 74 pour l'articulation /l^a/. On

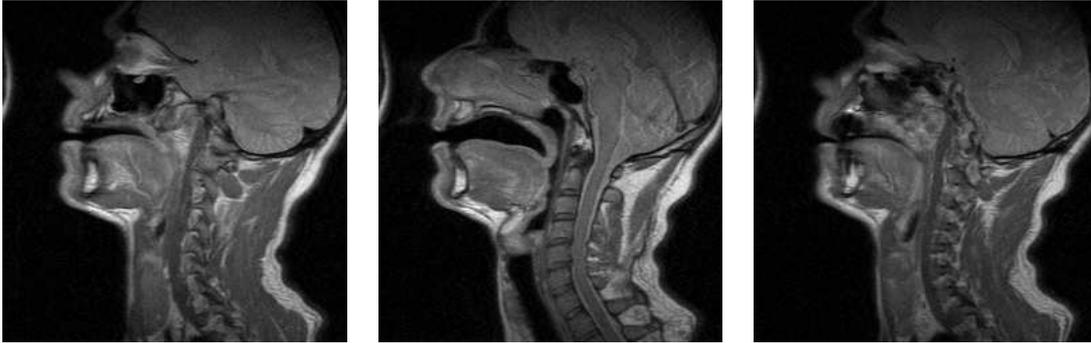


FIG. 2.5 – Exemple d'images IRM sagittales pour l'articulation /a/

observe des niveaux de gris différents proportionnels à la densité de protons d'hydrogène qui permettent de dessiner les tissus mous. Les structures osseuses et l'air se confondent en une même couleur noire (voir par exemple au niveau des dents). Il est à noter que la résolution de ces blocs d'images de 0.1 cm × 0.4 cm × 0.1 cm définit un échantillonnage spatial moins précis que celui du bloc d'images CT.

2.2.4.2 Les images IRM coronales

La complexité des fosses nasales et les très faibles interstices par lesquels l'air s'écoule requiert l'acquisition d'images adaptées à cette géométrie : les images sagittales acquises pour les 46 phonèmes du corpus ne permettent pas de détecter facilement les contours des fosses nasales et leur trop faible résolution latérale (4 mm d'écart entre deux images consécutives) ne permet pas d'envisager leur exploitation pour les fosses nasales. Un autre jeu d'images dédié, d'orientation coronale, donc considéré comme perpendiculaire à l'écoulement de l'air dans les fosses nasales, a donc été enregistré au laboratoire Advanced Telecommunications Research de Kyoto (Japon) le 24 février 1996. Les caractéristiques des images enregistrées sont détaillées dans le tableau 2.3. Notons également que les résolutions 0.0977 cm / px et 10.24 px / cm seront arrondies à 0.1 cm / px et 10 px / cm dans la suite du document.

La figure 2.6 présente deux images coronales extraites de cet enregistrement : à gauche une coupe située au niveau du cavum et à droite une coupe située approximativement à mi-hauteur entre le cavum et les narines. On note sur cette dernière image les fins interstices des fosses nasales qui sont difficilement visibles.

2.2.5 Les recoupes des images

Les orientations des différents jeux d'images enregistrées (un bloc d'images CT axiales, 46 blocs d'images IRM sagittales et un bloc d'images IRM coronales) sont d'une part dépendantes de l'instrument d'acquisition et d'autre part fixées pour l'ensemble de la séquence. Il apparaît

Position du sujet	allongée
Nombre d'images	32
Orientation	coronale
Région d'intérêt	de l'atlas (vertèbre C1) jusqu'à la pointe du nez
Taille	25 cm × 25 cm ou 256 pixels × 256 pixels
Epaisseur	0.4 cm
Résolution	0.0977 cm / pixels ou 10.24 pixels / cm

TAB. 2.3 – Caractéristiques des images IRM coronales acquises pour les fosses nasales.

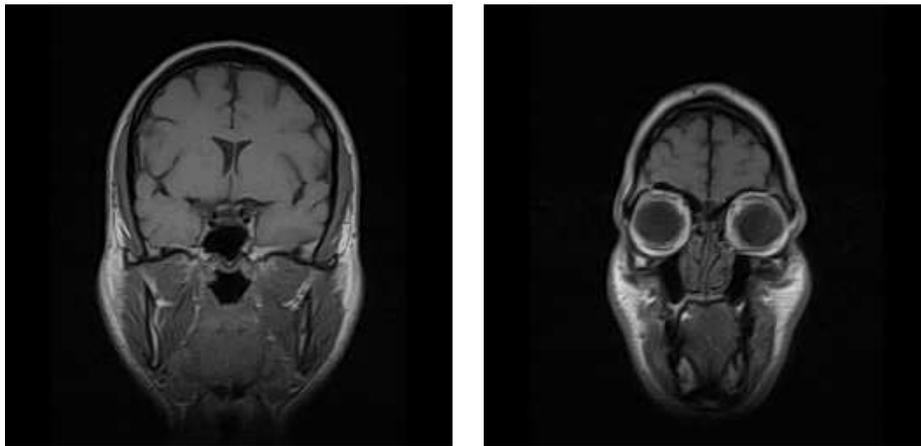


FIG. 2.6 – Exemple d'images IRM coronales acquises pour les fosses nasales.

cependant que l'orientation décidée ne correspond pas toujours à l'utilisation souhaitée. Elles sont utilisées pour extraire les formes 3D des organes (les sinus, les os, le voile du palais, les fosses nasales, etc.) en déterminant les contours de ces derniers sur chaque image, pour obtenir une représentation en contours 3D. Pour maximiser l'exactitude de la détermination des contours, il est de manière générale préférable d'éviter de travailler dans des plans d'image tangents à l'organe lui-même ; il est par contre préférable de travailler dans des plans d'image perpendiculaires à la surface de l'organe à déterminer. Pour donner un exemple, il est nettement plus optimal de détecter les contours d'un cylindre dans des plans perpendiculaires à son axe de révolution plutôt que dans des plans parallèles. Chaque bloc d'images ayant un usage dédié à la détermination d'un type d'organes (structures osseuses, structures déformables ou fosses nasales), les caractéristiques requises sont différentes :

Les images CT sont utilisées pour déterminer la forme tridimensionnelle des structures rigides, comme les différents sinus ou les os. Les formes de ces structures pouvant être très variées, une grande souplesse est nécessaire de manière à toujours posséder un jeu d'images perpendiculaire à la surface considérée. Nous avons donc choisi de rééchantillonner le jeu d'images axiales en un jeu d'images sagittales et un jeu d'images coronales.

Les images IRM sagittales sont utilisées pour déterminer la forme tridimensionnelle des structures déformables dans la région du conduit vocal. Si l'orientation sagittale semble

adéquate dans un voisinage du plan médiosagittal, elle peut poser des problèmes dans les régions latérales tangentes au conduit vocal. Nous avons donc décidé de rééchantillonner les blocs d'images IRM sagittales en images perpendiculaires à l'orientation conduit vocal.

Les images IRM coronales sont utilisées pour déterminer la forme tridimensionnelle des cavités nasales, du cavum aux narines. Au vu de la complexité du *tuyau* nasal, il est donc nécessaire d'avoir un jeu d'images perpendiculaires à la direction du tuyau considéré, y compris dans la région des narines. Les caractéristiques de ce bloc remplissent déjà les conditions requises, sauf dans la région des narines où nous avons décidé de le rééchantillonner perpendiculairement à la direction du tuyau.

Il est à noter que l'opération de rééchantillonnage n'apporte pas d'information nouvelle mais permet simplement d'exploiter les enregistrements effectués de manière plus cohérente avec les objectifs.

Chaque bloc d'images définit dans l'espace un parallélépipède rectangle dont l'unité élémentaire est le voxel. Il suffit donc de recouper ce cube suivant l'orientation souhaitée pour obtenir les nouvelles images. En pratique, la procédure d'intersection d'un bloc d'images par un plan quelconque est la suivante :

1. Le bloc d'images forme un cube composé de voxels. Un échantillonnage 3D de ce cube est obtenu en ne considérant que le centre de chaque voxel. À chaque point est associée la valeur, en niveau de gris de 0 à 255 (proportionnel à la densité moyenne de protons d'hydrogène dans le volume considéré), du voxel qu'il représente.
2. Un échantillonnage 2D de la nouvelle image souhaitée est défini par un quadrillage représentant le centre chaque pixel. Cet échantillonnage dépend directement et uniquement de la taille de la nouvelle image souhaitée et de sa résolution (pour des raisons pratiques on prendra toujours une même résolution suivant les deux axes de l'image).
3. Cet échantillonnage est placé en 3D dans le bloc d'images selon l'orientation souhaitée. Il forme donc un échantillonnage plan 3D intercalé à l'intérieur de l'échantillonnage du cube d'images.
4. Il reste à attribuer à chacun de ces points un niveau de gris ; celui-ci est calculé par interpolation linéaire entre les niveaux de gris correspondant aux centres des voxels du bloc voisins du point considéré.

On obtient une nouvelle image interpolée à partir du bloc d'images et orientée selon le plan considéré.

Comme indiqué précédemment, il est important de noter que ce processus ne permet pas d'affiner la précision des images enregistrées mais de les réinterpréter suivant une orientation plus adéquate. Si la résolution théorique de la nouvelle image peut être définie arbitrairement, sa résolution pratique ne peut donc pas être meilleure que celle du bloc dont elle est issue.

Nombre d'images	512	512
Orientation	sagittale	coronale
Taille	25.6 cm × 19.75 cm ou 512 pixels × 395 pixels	19.75 cm ou 512 pixels × 395 pixels
Epaisseur	0.05 cm	0.05 cm
Résolution	0.05 cm / pixels ou 20 pixels / cm	0.05 cm / pixels ou 20 pixels / cm

TAB. 2.4 – Caractéristiques des images CT réinterpolées



FIG. 2.7 – Images CT : image axiale originale (gauche), image sagittale réinterpolée (milieu) et image coronale réinterpolée (droite). Les deux traits sur l'image axiale originale correspondent aux orientations des deux images réinterpolées présentées ici.

Les trois types d'images sont donc réinterpolées comme suit :

Les images CT selon des orientations sagittales et coronales pour couvrir les trois orientations complémentaires axiale, sagittale et coronale. Pour rester proche des données originales, le bloc d'images axiales a été recoupé selon des plans sagittaux et coronaux exactement perpendiculaires aux plans d'images originaux et espacés d'un pixel. Par ailleurs, pour garder une cohérence entre les trois blocs, la même résolution de 0.05 cm/px a été attribuée aux nouvelles images. Enfin, les tailles des nouvelles images sont fixées à 512×395 px pour couvrir l'étendue du bloc. Le tableau 2.4 résume les caractéristiques des deux nouveaux blocs d'images ainsi obtenus. La figure 2.7 montre deux nouvelles images.

Les images IRM sagittales selon un ensemble de plans perpendiculaires à la direction du conduit vocal. Leurs orientations sont définies dans le plan médiosagittal par une grille semi-polaire fixe pour l'ensemble des 46 blocs d'images héritée de la grille définie par [Beautemps *et al.*, 2001] et visible à la figure 2.8 (gauche). Les 46 blocs d'images sont donc réinterpolés selon des plans perpendiculaires aux plans sagittaux en suivant cette grille semi-polaire. Toujours par souci de cohérence, la résolution du bloc d'images original est conservée et la taille des nouvelles images est fixée de manière à couvrir l'ensemble du bloc d'images original dans l'orientation désirée. Les caractéristiques des nouvelles images sont résumées dans le tableau 2.5. La figure 2.8 présente deux images réinterpolées pour l'articulation / $\tilde{\epsilon}$ /. Il est à noter que l'écart entre deux images sagittales consécutives de 4 mm entraîne par construction une précision relativement faible dans cette dimension

Nombre d'images	27
Orientation	perpendiculaire au conduit vocal selon la grille semi-polaire
Taille	10 cm × 20 cm ou 100 pixels × 200 pixels
Résolution	0.1 cm / pixels ou 10 pixels / cm

TAB. 2.5 – Caractéristiques des images IRM sagittales réinterpolées

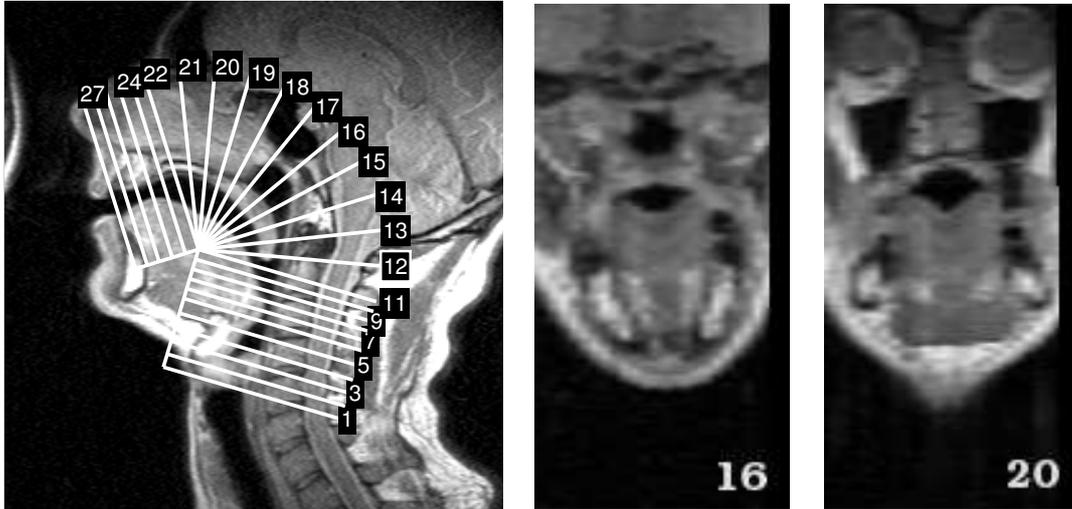


FIG. 2.8 – Images IRM sagittales : image médiosagittale originale de l'articulation / \bar{e} / sur laquelle est superposée la grille de recoupe semi-polaire fixe (gauche), image perpendiculaire réinterpolée correspondant au plan de grille numéro 16 de l'image de gauche (milieu) et image perpendiculaire réinterpolée correspondant au plan de grille numéro 20 de l'image de gauche (droite).

(selon l'axe horizontal des deux images réinterpolées présentées à la figure 2.8).

Les images IRM coronales selon un ensemble de plans perpendiculaires au conduit nasal. Les images originales étant déjà considérées perpendiculaires au conduit nasal dans la région des fosses nasales, il suffit de les recouper selon une grille tournante définie à la figure 2.9 (gauche) pour suivre le mouvement tournant à 90° du conduit dans la région des narines. Néanmoins, étant donnée le faible contraste de ce bloc d'images autour des narines, un des 46 blocs d'images IRM sagittales, arbitrairement celui correspondant à l'articulation /a/, lui est substitué pour cette opération. À nouveau par souci de cohérence, la résolution du bloc d'images original est conservé et la taille des nouvelles images couvre l'ensemble du bloc d'images original dans l'orientation désirée. Les caractéristiques des nouvelles images sont résumées dans le tableau 2.6. La figure 2.8 présente deux images réinterpolées.

Pour résumer, suite aux nombreuses réinterpolations d'images pour obtenir des jeux plus conformes à nos besoins, on dispose de :

Images CT : 149 images axiales, 512 images sagittales et 512 images coronales

Nombre d'images	7
Orientation	perpendiculaire au conduit nasal dans les narines
Taille	25.6 cm × 25.6 cm ou 256 pixels × 256 pixels
Résolution	0.1 cm / pixels ou 10 pixels / cm

TAB. 2.6 – Caractéristiques des images IRM sagittales réinterpolées pour les narines

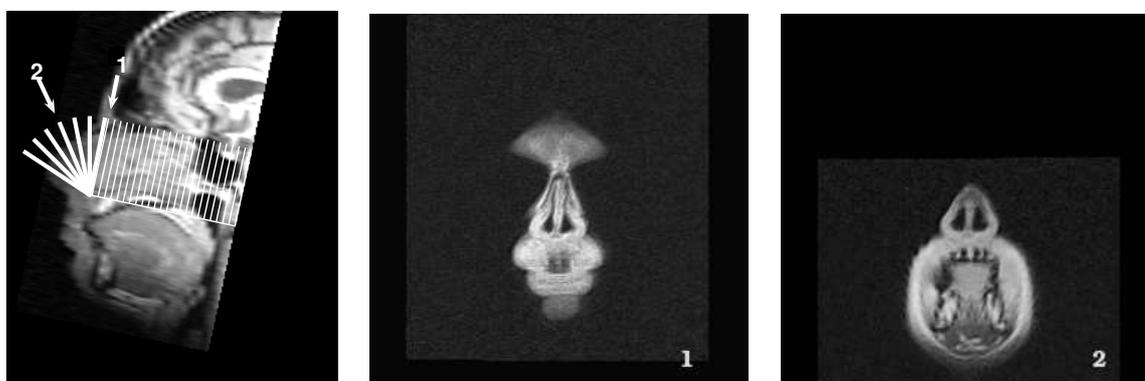


FIG. 2.9 – Images IRM réinterpolées pour les narines : superposition sur l'image médiosagittale de l'orientation des images IRM coronales originales pour la partie des fosses nasales (traits fins) et de l'orientation des images réinterpolées pour les narines (traits épais) à gauche ; deux images réinterpolées dans la région nasinaire (milieu et droite) correspondant aux deux plans de grilles marqués par des flèches sur l'image de gauche.

images IRM sagittales : 46 jeux de 25 images sagittales + 27 images perpendiculaires au conduit vocal pour les 46 phonèmes du corpus

images IRM coronales : 32 images coronales + 7 images perpendiculaires aux conduits nasinaires

2.2.6 Les enregistrements EMA

Des données dynamiques ont été enregistrées sur le sujet par articulographie électromagnétique le 18 janvier 2002 à l'Institut de la Communication Parlée à Grenoble. Cinq bobines réceptrices ont été attachées aux articulateurs du sujet, comme l'illustre la figure 2.10 (gauche) :

- Une bobine sur les incisives supérieures pour servir de référence.
- Une bobine sur les incisives inférieures pour obtenir la position de la mâchoire.
- Deux bobines sur la langue : une à l'avant (à 1 cm de l'apex) et une au milieu (à 4 cm de l'apex) pour obtenir la position de ces deux points de langue.
- Une bobine sur le voile du palais, à mi-distance entre la jonction du voile du palais avec le palais dur et la pointe de la luvette pour obtenir une estimation robuste des mouvements du voile du palais.

L'idéal aurait été de placer à la fois une bobine sur l'arrière de la langue et sur le voile du palais pour suivre plus précisément les mouvements du conduit vocal dans la zone vélaire. Si cette possibilité a été envisagée, elle a néanmoins été écartée pour des raisons techniques, bien qu'elle

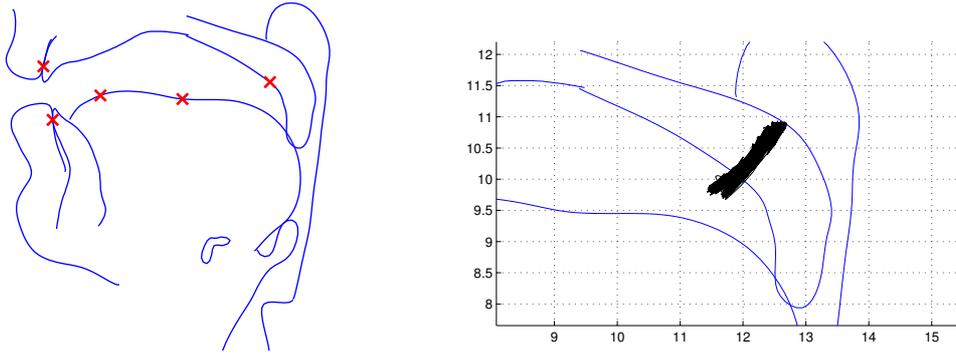


FIG. 2.10 – Superposition sur les contours médiosagittaux du conduit vocal en configuration /a/ de la position des 5 bobines EMA de l’enregistrement (gauche) et de l’ensemble des trajectoires de la bobine du voile du palais (droite)

soit en pratique réalisable ([Kaburagi et Honda, 2001]) : les deux bobines de l’arrière de la langue et du voile du palais seraient très proches, voire en contact, et pourraient ainsi fortement gêner le locuteur, à la fois pour son confort et pour produire l’articulation souhaitée.

Pour couvrir l’ensemble des possibilités articulatoires des séquences VCV de la langue française (V représentant une voyelle et C une consonne), toutes les combinaisons des 14 voyelles de la langue française [a ε e i y u o ø ɔ œ ā ɔ̃ ē œ̃] avec les 16 consonnes [b d g p t k v z ʒ f s ʃ m n ʁ l] ont été enregistrées par l’intermédiaire de logatomes VCV à une fréquence d’échantillonnage temporel de 500 Hz.

Nous avons présenté dans cette section les principales données utiles pour la construction du modèle articulatoire. Il convient cependant de souligner que d’autres données peuvent s’avérer utiles pour certains usages ponctuels au cours de l’étude. Leurs caractéristiques seront alors détaillées plus brièvement si nécessaire.

2.3 Des images aux formes tridimensionnelles

Les données acquises, il est nécessaire d’en extraire les informations utiles. La forme 3D de chaque organe, et pour chaque articulation du corpus si c’est un organe déformable, est entièrement extraite des blocs d’images afin d’être exploitée pour construire le modèle articulatoire. Nous présentons dans cette section la méthode d’extraction des formes tridimensionnelles des organes à partir des blocs d’images. Nous illustrons cette présentation par l’extraction des formes 3D des structures rigides (comme la mâchoire, les sinus, etc.) à partir des images CT.

2.3.1 Jongler avec les coordonnées

La diversité des blocs d’images, de leurs caractéristiques, de leurs résolutions et de leur région d’intérêt nécessite l’usage de jeux de coordonnées adaptés. L’extraction de formes 3D repose sur la détection des contours de l’organe considéré sur les images elles-mêmes. Des contours d’or-

ganes sont donc dessinés directement sur les images. Ces contours en traits continus peuvent être exprimés selon cinq types de coordonnées différentes :

En pixels 2D (*px 2D*) : Ce type de coordonnées dépend directement de l'image considérée.

Le coin en bas à gauche de l'image est arbitrairement défini comme l'origine du repère, la direction horizontale définit l'axe X et la direction verticale l'axe Y. L'unité est naturellement le pixel. Les valeurs du contour peuvent prendre une grandeur quelconque non forcément entière. Il est donc possible de positionner sur l'image un ensemble de points, dont l'enchaînement forme un contour, à des endroits quelconques sans être soumis aux pixels de l'image. La figure 2.11 montre un exemple de contour ainsi positionné sur une image ; ce contour est exprimé en pixels, quelle que soit la résolution de l'image. Ces coordonnées étant les plus proches de celles de l'image, chaque contour est dans un premier temps exprimé dans ce système de coordonnées. Néanmoins, il n'a pas d'existence en dehors de l'image elle-même, le repère lui restant lié.

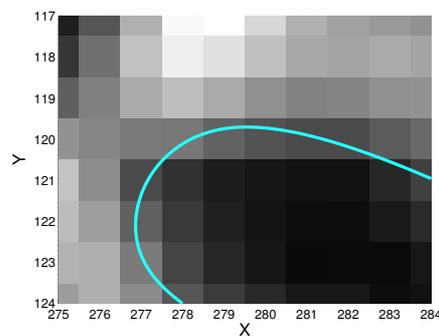


FIG. 2.11 – Superposition d'un contour exprimé en px 2D sur le fragment d'une image

En centimètres 2D (*cm 2D*) : Une première étape pour s'affranchir de la dépendance du contour manipulé à l'image dont il dépend est d'utiliser la résolution de l'image pour l'exprimer en centimètres. Cette unité constitue d'ailleurs la principale unité utilisée dans tout le document. L'origine du repère, le coin en bas à gauche de l'image, reste néanmoins lié à l'image elle-même.

En voxels 3D (*vx 3D*) : De manière similaire, on peut exprimer le contour plan tracé sur une image dans un repère local lié au bloc d'images dans lequel il se situe. Le bloc d'images forme en 3D un bloc de voxels qui représente tout ou partie de la tête du sujet. L'origine du repère 3D est fixée dans le coin en bas, à gauche et en avant par rapport à la tête du sujet. L'unité est maintenant le voxel. Connaissant la position de l'image sur laquelle est dessiné le contour dans ce bloc, il est alors aisé d'exprimer ledit contour plan dans le repère global 3D.

En cm 3D du bloc d'images (*cm 3D bloc*) : En combinant les apports des deux systèmes de coordonnées précédents par rapport au premier, c'est-à-dire le gain d'une unité universelle et d'un repère 3D lié au bloc d'images, il devient aisé d'exprimer le contour plan considéré en centimètres dans le repère 3D du bloc d'images. Pratiquement, connaissant la résolution des voxels, la transition des *vx 3D* (étape précédente) vers les *cm 3D bloc* (étape courante) ne présente pas de difficulté. Ce système de coordonnées reste néanmoins encore lié au bloc d'images considéré.

En cm 3D absolu (*cm 3D*) : La dernière étape consiste à exprimer le contour en centimètres dans un repère absolu. Ce dernier n'est par définition plus lié au bloc d'images dont les bords sont fixés arbitrairement par l'appareil de mesure, mais plutôt à la structure crânienne du sujet qui constitue un ensemble de repères se retrouvant sur chaque bloc d'images (voir paragraphe 2.4.1). Le contour plan considéré, jusqu'à présent exprimé en centimètres dans un repère lié au bloc d'images, s'exprime maintenant en centimètres dans le repère absolu.

Les différents systèmes de coordonnées décrits ci-dessus permettent alors de passer, en plusieurs étapes, d'un contour exprimé en pixels dans un plan d'image à un contour exprimé en centimètres 3D dans un repère absolu lié à la structure crânienne du sujet. La suite du document se réfère indifféremment à l'un ou l'autre de ces systèmes de coordonnées. Les relations entre les différents systèmes sont schématisés sur la figure 2.12.

2.3.2 Tracer un contour

Les différents systèmes de coordonnées étant rigoureusement définis, il est désormais possible d'extraire les formes 3D des différents organes. La première étape consiste à tracer manuelle-

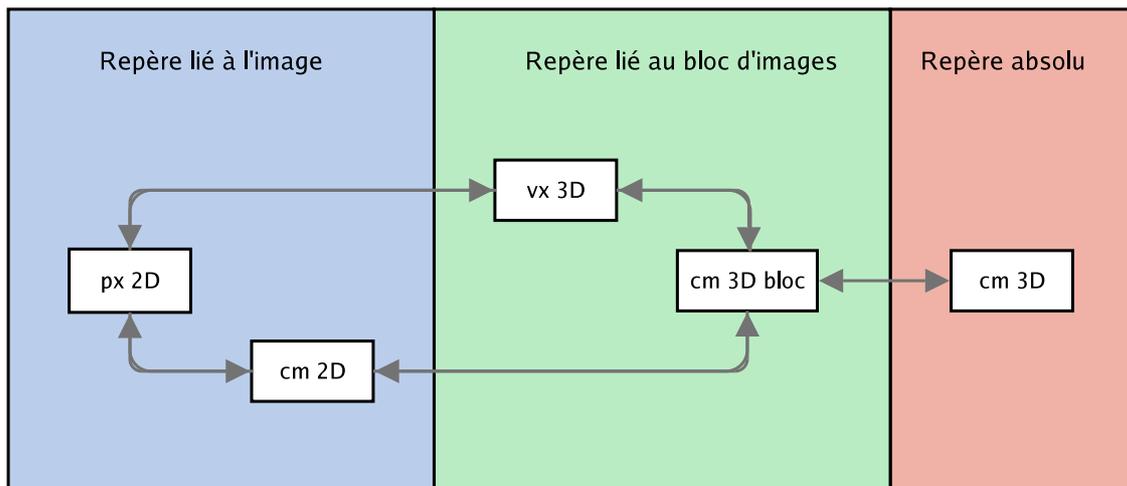


FIG. 2.12 – Relations entre les divers systèmes de coordonnées pour les contours tracés sur les images (px = pixel ; vx = voxel ; cm = centimètre).

ment les contours des organes sur les images. Le choix du bloc d'images considéré dépend de l'organe à tracer : l'objectif est de sélectionner les images dans une orientation la plus perpendiculaire possible à la surface de l'organe. Ainsi par exemple la mâchoire, reconstruite à partir des images CT puisqu'il s'agit d'une structure osseuse, est considérée sur les images axiales pour rester perpendiculaire à l'orientation des régions latérales ; une orientation coronale perpendiculaire à la surface dans la région du corps de la mâchoire aurait aussi pu être envisagée. Enfin, dans les cas complexes comme les fosses nasales, plusieurs orientations sont mêlées pour optimiser la détermination des structures fines. Mathématiquement, le contour, tracé directement sur l'image, donc exprimé en px 2D continus (voir par exemple la figure 2.11), est le résultat d'une interpolation en splines cubiques à partir d'un certain nombre de points de contrôle placés manuellement sur l'image. Plus le nombre de points de contrôle est élevé, plus la courbe résultante épousera finement les entrelacs du bord de l'organe et inversement, moins le nombre de points de contrôle est grand, plus la courbe résultante sera lisse et arrondie. En pratique, le contour spline interpolé est actualisé en continu pour aider l'opérateur à tracer. Le tracé est donc réalisé itérativement en déplaçant, en enlevant, ou en rajoutant des points de contrôle jusqu'à obtenir un contour fiable de l'organe aux yeux du traceur. La figure 2.13 présente comme exemple les tracés de la mâchoire sur les deux images axiales dont la position est visible sur l'image médiosagittale de gauche. Les points représentent les points de contrôle tandis que les traits continus représentent l'interpolation en splines cubiques. Les points interpolés sont à peu près équidistants, d'un nombre 5 à 20 fois supérieur au nombre de points de contrôle. Sur la figure 2.13, l'image du milieu présente un contour de 312 points pour 30 points de contrôle tandis que l'image de droite présente deux contours, l'un de 89 et l'autre de 68 points pour 12 et 13 points de contrôle. Notons au passage qu'un organe peut être représenté sur une image par plusieurs contours, appelés sous-contours dans la suite du document, comme c'est le cas de la mâchoire sur l'image de droite de la figure 2.13.

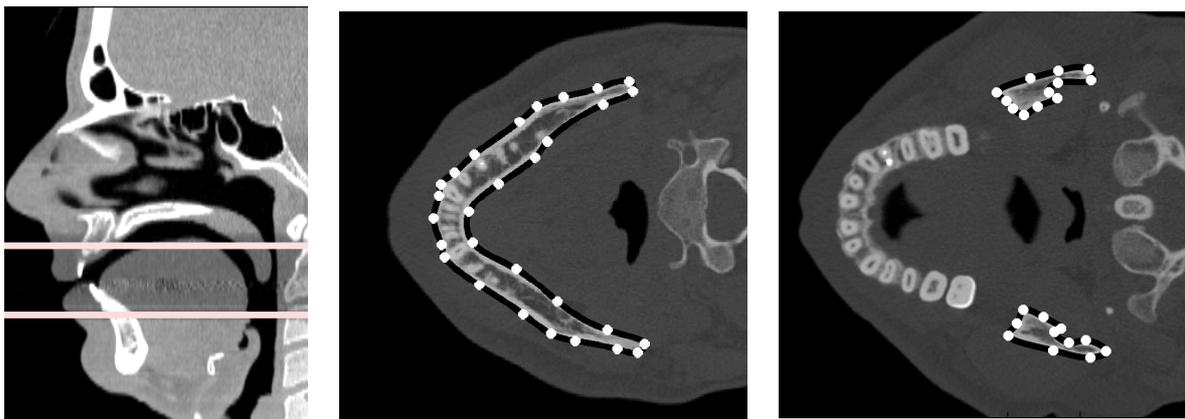


FIG. 2.13 – Illustration des contours de la mâchoire dans deux plans d'images CT axiales dont les positions sont visibles sur l'image médiosagittale de gauche. Les traits continus épais représentent les contours réinterpolés en splines cubiques à partir des points de contrôle représentés par des points clairs.

Suivant la méthodologie décrite ci-dessus, les contours de la mâchoire ont été tracés sur les images axiales du bloc d'images CT, le sinus sphénoïdal sur les images sagittales, les deux sinus maxillaires sur les images coronales et enfin les fosses nasales à la fois sur les images axiales et coronales.

Il est important de rappeler ici que tous les contours sont tracés manuellement et dépendent donc de l'habileté de l'expert chargé de tracer. Néanmoins, [Soquet *et al.*, 2002] ont montré que les tracés manuels des bords du conduit vocal sur des images IRM sont fiables et reproductibles avec de faibles erreurs. La qualité des images CT étant supérieure aux images IRM, les tracés manuels présents peuvent donc être considérés comme représentatifs des contours réels et seront utilisés comme références dans la suite de l'étude.

Enfin, rappelons également le travail important que représentent ces tracés : que Géraldine Rougier et Pierre Philippe, ayant effectué ce travail avant mon arrivée à l'ICP, en soient infiniment remerciés !

2.3.3 Obtenir une surface 3D

Les contours tracés sur les images sont par définition exprimés en px 2D. Conformément à la procédure résumée sur le schéma de la figure 2.12, ils peuvent être exprimés en cm 3D bloc correspondant au bloc d'images CT. En rassemblant tous les contours tracés pour chaque organe, on obtient donc un échantillonnage des organes en contours plans (voir les figures 2.14 (gauche) et 2.16 (gauche) pour la mâchoire et les fosses nasales).

L'étape suivante consiste à reconstruire la surface de l'organe considéré à partir des contours plans qui en constituent un échantillonnage. Cette opération est effectuée grâce à un logiciel de reconstruction de surface à partir de points (voir [Fabri, 2002]) développé par l'équipe *Geometrica* à l'INRIA (France) (il existe une version en ligne à l'adresse suivante : <http://cgal.inria.fr/Reconstruction/index.html>). Les points des contours plans constituent les données d'entrée du logiciel : 6706 pour la mâchoire, 3047 pour le sinus sphénoïdal, 13171 pour les sinus maxillaires et 45250 pour l'ensemble des fosses nasales dont les points sont issus à la fois des tracés axiaux et coronaux. La surface obtenue est un maillage triangulaire¹ : voir la figure 2.14, milieu, pour la mâchoire et la figure 2.15 pour un détail de cette surface. Comme le montre cette dernière figure, chaque sommet est le sommet de plusieurs triangles et chaque arête est la bordure de deux triangles, exception faite des bords s'il y en a, le maillage pouvant être ouvert ou fermé pour définir une surface ouverte ou fermée. On obtient par ailleurs une surface *texturée* en considérant les triangles pleins (voir les figures 2.14, droite, et 2.16, milieu). L'ordre de grandeur du nombre de triangles de chaque surface est d'environ quatre fois le nombre de points originaux. La densité du maillage dépend directement de celle des points de départ. Ainsi, la distance moyenne entre deux sommets est par exemple inférieure au millimètre pour les sinus et de l'ordre de 1.5 mm pour la mâchoire, dont les tracés sont plus distants entre eux car obtenus des images CT axiales

¹en format *Virtual Reality Modeling Language 2* (VRML 2)

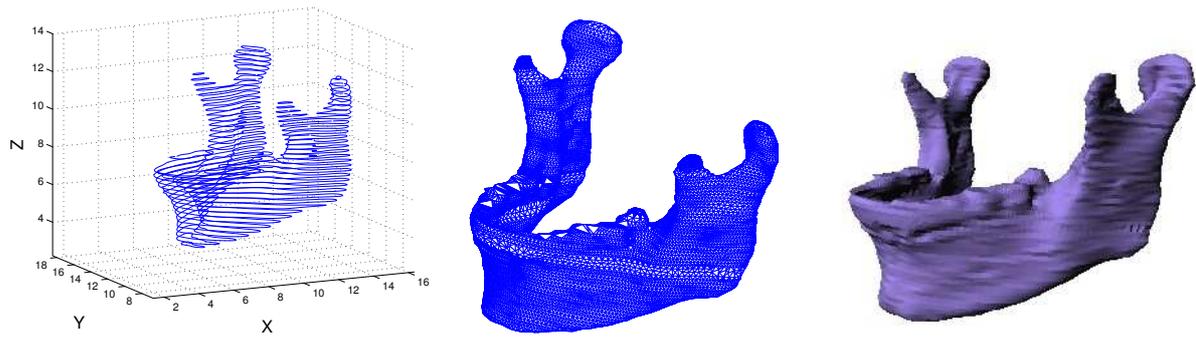


FIG. 2.14 – Contours plans axiaux de la mâchoire (gauche) : seul un contour sur deux est affiché pour plus de clarté. Maillage de la mâchoire obtenu à partir des contours plans (milieu et droite). Une texture est ajoutée au maillage de la figure droite.

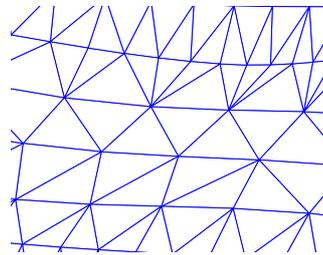


FIG. 2.15 – Zoom sur un fragment du maillage de la mâchoire

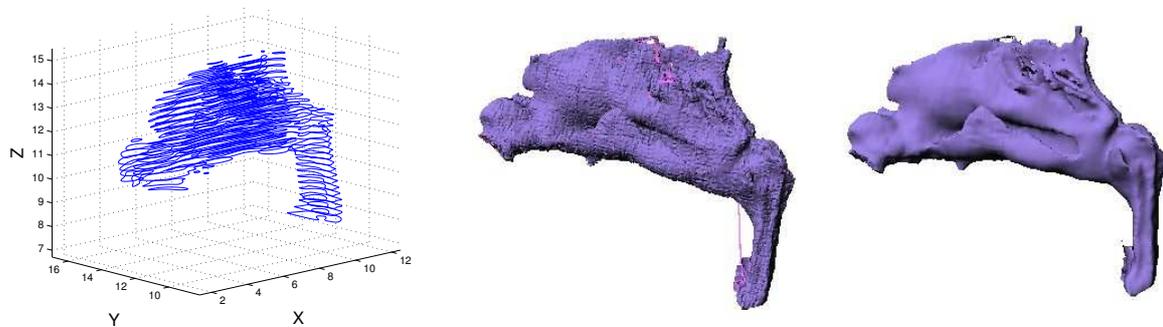


FIG. 2.16 – Contours plans axiaux des fosses nasales (gauche, seul un contour sur deux est affiché pour plus de clarté), maillage surfacique texturé avant (gauche) et après lissage (droite).

espacées de 0.13 cm.

La géométrie particulièrement complexe des fosses nasales a nécessité l'introduction des points à la fois issus des tracés axiaux et coronaux. Les tracés pouvant être légèrement décalés d'une orientation à l'autre, ce quadrillage a conduit à un maillage assez anguleux. Pour résoudre ce problème, la surface a été lissée par le logiciel *SmoothMesh* ([Huber, 1999]) développé au laboratoire VMR de l'université Carnegie Mellon (États-Unis). Le résultat de cette opération

est visible à la figure 2.16 (droite).

Pour évaluer l'erreur introduite par les étapes de maillage puis de lissage, la distance de l'ensemble des points originaux d'un organe (c'est-à-dire les points des tracés) à la surface de son maillage a été calculée : chaque triangle étant considéré comme plein, la distance calculée est celle du point jusqu'à son projeté sur la surface du maillage. Comme expliqué dans la section 2.3.2, les tracés sont considérés comme la référence de la forme 3D d'un organe ; la distance calculée entre ces points de référence et le maillage donne ainsi une évaluation de l'erreur de reconstruction des formes 3D des organes par les maillages VRML. L'erreur ainsi calculée étant inférieure au dixième de millimètre pour tous les organes, les maillages obtenus par le procédé décrit dans cette section sont considérés comme une approximation très fiable de la forme réelle des structures rigides du sujet, à l'exception des fosses nasales dont l'erreur avoisine le demi-centimètre. Cette dernière s'explique par la forme complexe et sinueuse de l'organe que le logiciel de maillage ne peut pas reproduire fidèlement, ce qui conduit à forte erreur ; néanmoins les résultats visuels de superposition des recoupes des fosses nasales sur les images (dont le principe est détaillé au paragraphe 2.3.4) montrent des résultats très satisfaisants. Enfin, le maillage des fosses nasales présenté ici n'a qu'un caractère informel car l'étude montrera l'intérêt de travailler sur des images IRM plutôt que sur des images CT pour cette structure en raison de la présence de muqueuses non visibles sur les images CT (voir le paragraphe 2.5.7).

Les maillages du palais dur et les dentitions inférieure et supérieure sont issus des moulages dentaires dont la précision est supérieure à celle des images CT. Ils ont donc été scannés pour obtenir deux maillages surfaciques exploitables numériquement et dont les mailles sont espacées en moyenne d'un tiers de millimètre. La figure 2.17 illustre les scans obtenus des moulages. La position des moulages dans l'espace est arbitraire car non lié à un bloc d'images. Pour rendre l'ensemble des maillages des structures rigides cohérent, ils ont été alignés manuellement à la place qui aurait été la leur s'ils avaient été reconstruits à partir des images CT. Ce processus de calage des données est décrit dans la section 2.4.3.

À ce point du processus on dispose donc des formes 3D de la mâchoire, des sinus sphénoïdal et maxillaires, du palais dur, des fosses nasales et des dentitions supérieure et inférieure calées entre elles et exprimées en grandeurs réelles (cm) dans un repère lié au bloc d'images CT.

2.3.4 Manipuler les maillages

Les maillages obtenus représentent des surfaces dont l'intersection avec des plans quelconques forme par définition des contours : l'intersection d'un plan avec un triangle forme un segment, avec plusieurs triangles plusieurs segments et avec plusieurs triangles accolés (c'est-à-dire un maillage) plusieurs segments accolés (c'est-à-dire un contour). Il est donc maintenant aisé d'obtenir les contours des structures rigides dans n'importe quel plan et n'importe quelle orientation. Cette propriété sera largement utilisée dans la suite de l'étude pour superposer sur les images IRM les contours des structures rigides qui n'y sont pas visibles. Un exemple de recoupe est

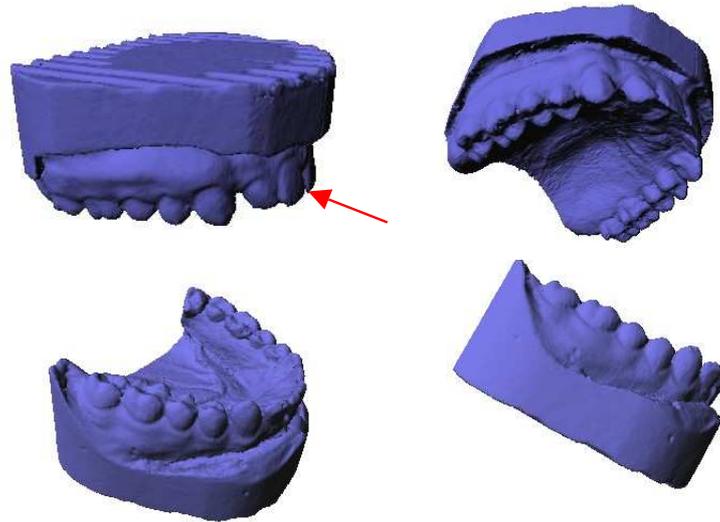


FIG. 2.17 – Scans des moulages dentaires : dentition supérieure et palais dur (haut) et dentition inférieure (bas). La flèche indique la pointe centrale inférieure des incisives supérieures

visible à la figure 2.18 avec la superposition dans le plan médiosagittal des contours tracés manuellement sur l'image CT de la figure 2.7 (milieu) page 49 et des intersections des structures rigides avec le plan d'image correspondant. On note au passage la bonne correspondance entre les intersections des structures rigides et les tracés manuels pour le palais dur et la mâchoire.

Une autre propriété intéressante de ces maillages est qu'ils peuvent être *découpés* en plusieurs sous-surfaces en ne conservant par exemple qu'une des deux moitiés du maillage obtenu

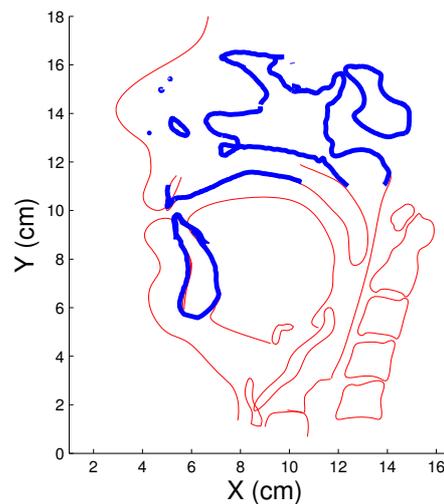


FIG. 2.18 – Superposition des contours tracés manuellement sur l'image CT médiosagittale (traits fins) et des contours obtenus par intersection du plan d'image avec les maillages des structures rigides (traits épais).

par l'intersection avec un plan. Les maillages des dentitions inférieure et supérieure et du palais dur sont ainsi séparés du maillages des socles pour ne conserver que la partie correspondant à l'anatomie du sujet. Le résultat est visible à la figure 2.19.

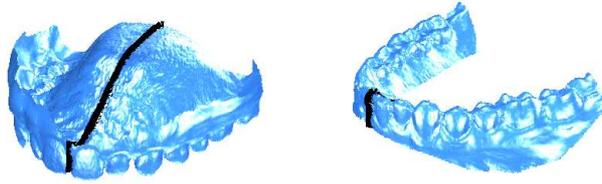


FIG. 2.19 – Maillages issus des moulages dentaires pour lesquels les parties correspondant aux socles ont été enlevées ; le trait continu représente le contour médiosagittal recoupé dans ces maillages

Le processus décrit dans cette section permet d'obtenir les formes tridimensionnelles d'organes à partir d'images anatomiques de la tête du sujet. Quatre étapes sont identifiées : (1) tracé des contours de l'organe dans les plans d'images, (2) passage des contours 2D planaires en contours 3D planaires formant un échantillonnage 3D de l'organe, (3) maillage des contours planaires pour former une surface et (4) lissage de cette surface si cela s'avère nécessaire. Le processus donne de très bons résultats pour les structures rigides issus des images CT. Il peut être étendue aux structures déformables à extraire sur les images IRM. Néanmoins, avant cette étape, les maillages des structures rigides ainsi obtenus sont utilisés pour placer l'ensemble des blocs d'images dans le même repère commun absolu attaché à ces structures rigides.

2.4 Alignement dans un repère commun

Nous avons souligné dans le paragraphe 2.3.1 l'importance d'un repère absolu commun qui ne soit pas lié au bloc d'images considéré. Nous présentons dans cette section la méthode d'alignement de l'ensemble des données articulaires dans un repère absolu. L'alignement dans ce repère permettra alors de comparer et d'utiliser les différents types de données de manière conjointe.

2.4.1 Définition d'un repère absolu

Un repère commun à l'ensemble des données articulaires doit logiquement être lié à l'anatomie du sujet. Il paraît donc naturel de l'attacher à la structure crânienne du sujet. Pour des raisons historiques, le repère 3D défini est une extension du repère médiosagittal défini par [Beautemps *et al.*, 2001] ; il permet notamment de rester cohérent avec l'ensemble des données articulaires acquises sur le même sujet lors de précédentes études ([Badin *et al.*, 1998], [Beautemps *et al.*, 2001], [Badin *et al.*, 2002], [Rossato *et al.*, 2003]).

Le repère orthonormé tridimensionnel absolu est défini comme suit :

1. L'axe X est orienté de l'avant vers l'arrière de la tête du sujet dans le plan médiosagittal et approximativement dans le plan occlusal, l'axe Y de la gauche vers la droite et l'axe Z des pieds vers la tête.
2. La pointe inférieure centrale des incisives supérieures indiquée par une flèche sur la figure 2.17, référencée dans la suite du document comme le point *Upper Teeth* (UT), est fixée comme le point (5,0,10). Il s'ensuit que l'ensemble du conduit vocal a des coordonnées positives dans le plan médiosagittal (X, Z).

L'ensemble des données peuvent maintenant être placées dans ce repère. Cet alignement s'effectue en plusieurs étapes :

- Le maillage du palais dur et de la dentition supérieure est positionné manuellement dans le repère absolu.
- Ce maillage est ensuite positionné semi-automatiquement dans le repère cm 3D lié au bloc d'images CT. La transformation de passage du repère cm 3D absolu vers le repère cm 3D lié au bloc CT est alors connue, donc également la position de toutes les structures rigides dans le repère absolu.
- Les structures rigides sont positionnées sur chacun des 46 blocs d'images IRM sagittales et du bloc d'images IRM coronales par un procédé semi-automatique. Les transformations de passage du repère cm 3D absolu vers les repères cm 3D liés aux blocs IRM sont alors connues.
- Enfin la mâchoire et la dentition inférieure dont la position par rapport aux autres structures osseuses change en fonction de l'articulation sont placées semi-automatiquement sur chacun des 46 blocs d'images IRM sagittales.

Ces étapes sont décrites en détail dans la suite de la section.

2.4.2 Alignement du palais dur dans le repère absolu

Cette tâche est effectuée manuellement par itérations. Le maillage est recoupé itérativement dans un certain nombre de plans présentant des propriétés géométriques intéressantes : plan médiosagittal, plans sagittaux, axiaux et coronaux. Plan par plan le contour est approché sur des critères géométriques de sa position idéale dans le repère absolu. À chaque étape la transformation déterminée en 2D est appliquée à l'objet 3D en entier. Le maillage est donc placé par approximations successives à sa place dans le repère absolu 3D.

Cette étape permet de disposer du maillage du palais positionné dans le repère absolu défini ci-dessus. L'étape suivante consiste à le placer dans le repère correspondant au le bloc d'images CT de manière à connaître la transformation mathématique permettant de passer d'un repère à l'autre. Ce processus s'inscrit dans le cadre plus général de l'alignement semi-automatique d'un maillage de structure rigide sur un bloc d'images.

2.4.3 Alignement semi-automatique d'un maillage de structure rigide sur un bloc d'images

Ce processus s'appuie sur l'alignement du maillage du palais dur et de la dentition supérieure sur le bloc d'images CT. Le principe de calage des structures rigides sur un bloc d'images servira de base au recalage de l'ensemble des blocs IRM. Il est donc préférable d'en faire un processus automatique, ou au moins semi-automatique comme nous le proposons.

L'objectif de ce paragraphe est de trouver la transformation géométrique de passage de la position d'un objet rigide dans un repère (le maillage du palais dur dans le repère absolu) vers sa position dans un autre repère (lié au bloc d'images). La transformation géométrique globale d'un objet rigide en 3D pour passer d'une position à une autre consiste en une rotation 3D et une translation 3D ; elle sera donc nommée dans la suite du document une *rototranslation* (3D). Cette *rototranslation* possède par définition six degrés de liberté : trois angles signés de rotation et trois distances signées de translation. La rototranslation de passage du maillage de palais dur dans le repère absolu à sa position dans un repère lié à un bloc d'images définit la rototranslation de passage entre les deux repères. Il suffit donc de caler le maillage d'une ou plusieurs structures rigides, en l'occurrence le palais dur dans notre exemple, dans le repère lié à un bloc d'images pour connaître la rototranslation de passage d'un repère à l'autre. Notons que cette rototranslation est réversible et que tout objet peut être exprimé indifféremment dans l'un ou l'autre des repères. Cette transformation définit le passage entre les deux systèmes de coordonnées les plus à droite de la figure 2.12 page 54.

Les différentes étapes du processus de calage d'un maillage de structure rigide sur un bloc d'images sont alors les suivantes :

1. La position du maillage de la structure rigide considérée est initialisée approximativement dans le repère 3D du bloc d'images en cm ; ceci revient donc à initialiser grossièrement les six paramètres de la rototranslation recherchée.
2. Un certain nombre de points (typiquement entre 2-3 et 20 par image) sont déterminés manuellement sur certaines images du bloc comme des points de la surface considérée. Ces points sont naturellement exprimés en px 2D de l'image correspondante. Pour plus de flexibilité, il est possible selon les besoins de reconstruire de nouvelles images à partir du bloc permettant un choix plus judicieux des points d'ancrage. Une attention particulière est portée sur la symétrie des points choisis : pour ne pas introduire de déséquilibre, l'opérateur s'efforce au maximum de répartir équitablement les points sur l'image. La figure 2.20 montre des points d'ancrage dans un plan d'image sagittal (haut) et coronal (bas).
3. Ces points sont recalculés en cm 3D du bloc d'images conformément au schéma de la figure 2.12.
4. Les six paramètres de la rototranslation recherchée sont déterminés par optimisation, c'est-à-dire en minimisant la distance 3d entre les points cibles et leur projeté sur le maillage.

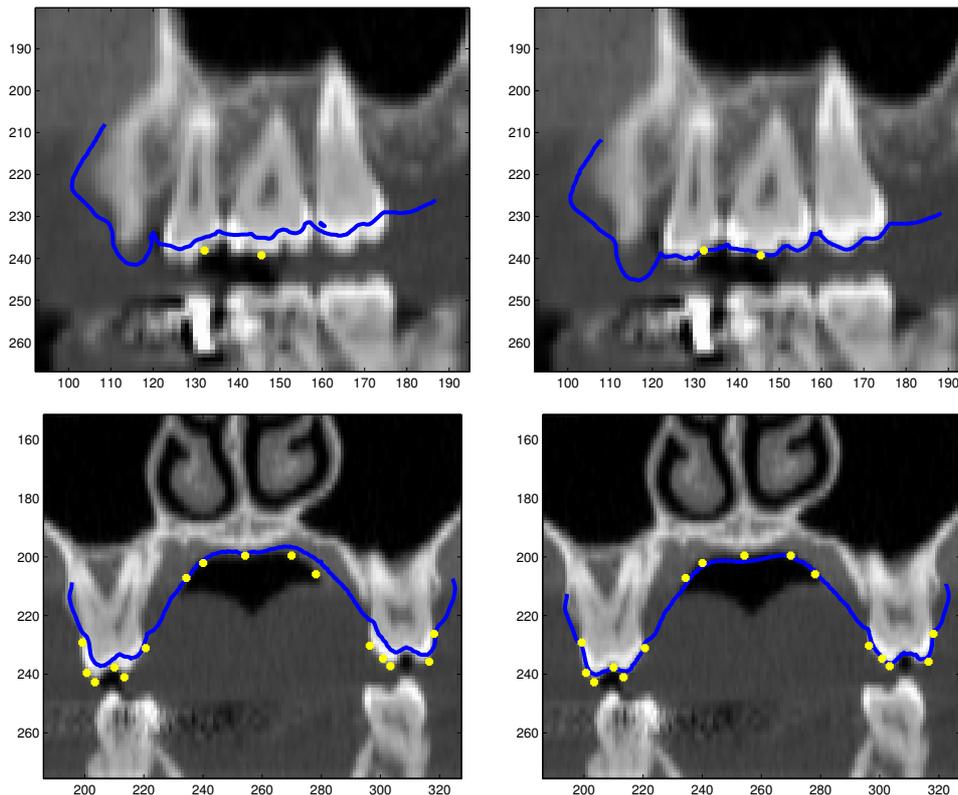


FIG. 2.20 – Superposition sur une image CT sagittale à 2.5 cm du plan médiosagittal environ (haut) de la recoupe du palais placée initialement à la main (haut gauche) et de sa recoupe après recalage semi-automatique (haut droite); les points représentent les points d’ancrage dans ce plan d’image. *Idem* pour les deux images du bas dans un plan coronal.

Cette étape de minimisation est réalisée par la fonction MATLAB *fminunc* de minimisation non linéaire sans contrainte.

Un exemple de résultat de recalage est affiché à la figure 2.20 où les deux images de droite montrent les contours du palais dur et de la dentition supérieure recalés par rapport aux positions initiales correspondantes sur les images de gauche; les points d’ancrage utilisés pour le recalage dans ces images sont également affichés.

Comme souligné précédemment, la rototranslation peut être utilisée indifféremment dans les deux sens : du repère 3D absolu vers le repère 3D du bloc d’images ou inversement. Par ailleurs, ce procédé est considéré comme semi-automatique dans le sens où les points d’ancrage sont marqués à la main sur les images par un opérateur mais l’étape de détermination de la rototranslation elle-même par optimisation est automatique.

L’amélioration apportée par ce processus semi-automatique sur le calage du maillage du palais dur sur les images CT est mesurée par la réduction sensible de la distance des points d’ancrage (qui constituent la référence anatomique) au maillage : l’*Erreur Quadratique Moyenne*

(*Root Mean Square Error* en anglais, erreur RMS dans la suite du document) passe de 0.13 cm à 0.06 cm. On notera par ailleurs l'attention particulière qui a été portée à l'étape d'initialisation dans le cas du palais, comme l'attestent les deux images de gauche de la figure 2.20.

L'opération de recalage semi-automatique est finalement validée par deux critères :

- Le critère d'erreur RMS, défini ci-dessus.
- Le critère visuel, en observant les recoupes de la surface rigide recalée dans différents plans d'images. Les deux images de droite de la figure 2.20 montrent ainsi clairement un recalage correct du palais par rapport aux deux images de gauche, même si celui-ci pourrait encore être amélioré.

En se basant sur le même principe, la dentition inférieure est également recalée sur le bloc d'images CT. L'erreur RMS tombe à 0.03 cm. Elle est ensuite *soudée* à la mâchoire issue des images CT, c'est-à-dire qu'un déplacement de l'une entraîne immédiatement le déplacement de l'autre. On peut donc considérer le maillage de la dentition inférieure simplement comme une partie plus précise du maillage de la mâchoire au niveau des dents. Il ne sera par conséquent plus fait de différence entre les deux maillages dans la suite de l'étude.

[Takemoto *et al.*, 2004] ont également proposé une méthode pour replacer sur des images (IRM cette fois) un volume 3D correspondant à la mâchoire et au palais dur avec leur dentition. La méthode proposée est très différente de la nôtre sur plusieurs points. Les deux principaux sont que les structures rigides ne sont plus considérées en tant que surfaces mais en tant que volumes de voxels et que la détection du volume est effectuée automatiquement sur le bloc d'images par seuillage. Pour trouver les six paramètres de la rototranslation, ils s'appuient sur la minimisation d'une erreur de volume calculée comme le volume de superposition, en voxels, de la structure rigide à positionner sur les tissus mous. L'erreur résiduelle est de 1.18 cm^3 avec un palais dur et une dentition supérieure (au sens anatomiquement large) de 12.44 cm^3 . Pour tenter de comparer cette erreur avec les mesures présentées ici, si tant est que cela soit possible, on peut considérer le palais dur comme une boule de 12.44 cm^3 de volume ; une variation de 1.18 cm^3 de ce volume correspond environ à une variation de son rayon de 0.05 cm, à comparer à l'erreur de 0.06 cm dans l'étude présente. Cependant, il reste clair que ce ne sont que des ordres de grandeur et qu'il paraît difficile de comparer les deux approches.

Le palais dur est maintenant calé sur le bloc d'images CT. La rototranslation de passage du repère lié au bloc d'images CT vers le repère de référence est donc connue. En l'appliquant à l'ensemble des autres structures rigides, issues et donc calées par construction dans le repère du bloc d'images CT, on obtient l'ensemble des structures rigides calées dans le repère absolu. L'étape suivante consiste donc à placer ces structures rigides dans chacun des 46 blocs d'images IRM sagittales et du bloc d'images IRM coronale de manière à obtenir la rototranslation de passage des 47 repères liés aux blocs d'images IRM vers le repère absolu.

2.4.4 Alignement des images IRM

L'alignement des images IRM dans le repère 3D absolu s'effectue par un processus semi-automatique similaire à celui présenté ci-dessus. L'ensemble des structures rigides, à l'exception de la mâchoire, sont à ce point du processus placées et gardent une position fixe dans le repère absolu. Les maillages sont donc considérés comme solidaires. Ces structures rigides sont ensuite calées par le procédé semi-automatique décrit ci-dessus dans chacun des 47 repères 3D en cm liés aux 47 blocs d'images IRM. Les points d'ancrage doivent être choisis judicieusement. Dans le plan médiosagittal, ils sont choisis le long du palais dur et du sinus sphénoïdal, pour bénéficier d'une étendue maximale et caler ainsi au mieux les structures selon l'orientation sagittale. Dans les autres plans (sagittaux sur le côté, coronaux, axiaux et transverses divers) ils sont choisis le long des sinus maxillaires qui présentent l'avantage, en plus d'une grande étendue, d'avoir une anatomie relativement *ronde* qui évite à l'algorithme d'optimisation de trouver des minima locaux qui correspondraient aux calages des mauvaises faces des structures rigides sur les points d'ancrage. Ce pourrait par exemple être le cas des fosses nasales dont les différentes faces sont très imbriquées et très rapprochées les unes aux autres; elles sont donc exclues du processus de calage.

Quelques exemples de superpositions sur des images IRM de contours obtenus par l'intersection des maillages des structures rigides ainsi calées avec les plans d'images correspondant à l'articulation /1^a/ sont montrés à la figure 2.22. On observe que la superposition des structures rigides sur les images IRM clarifient grandement la lecture de celles-ci.

Les résultats de ces calages sont visibles à la figure 2.21. On observe que l'erreur RMS entre les points d'ancrage, considérés comme la référence, et leur projeté sur les maillages calés, varie entre 0.08 et 0.15 cm. L'erreur RMS moyenne sur l'ensemble des 47 blocs est de 0.11 cm.

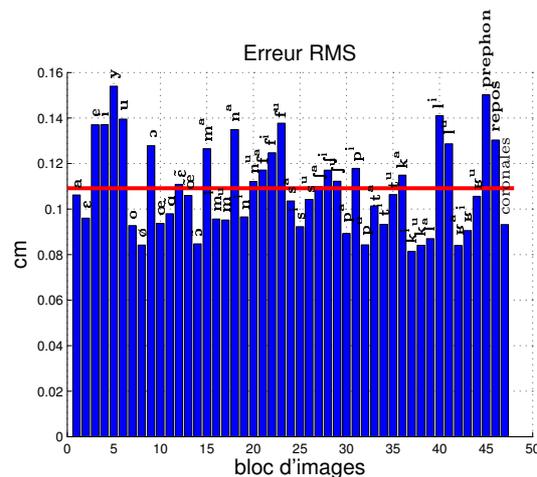


FIG. 2.21 – Erreur RMS entre les points d'ancrage et les maillages calés des structures rigides sur les 47 repères liés aux blocs d'images IRM sagittales et coronale. Le trait continu désigne la moyenne.

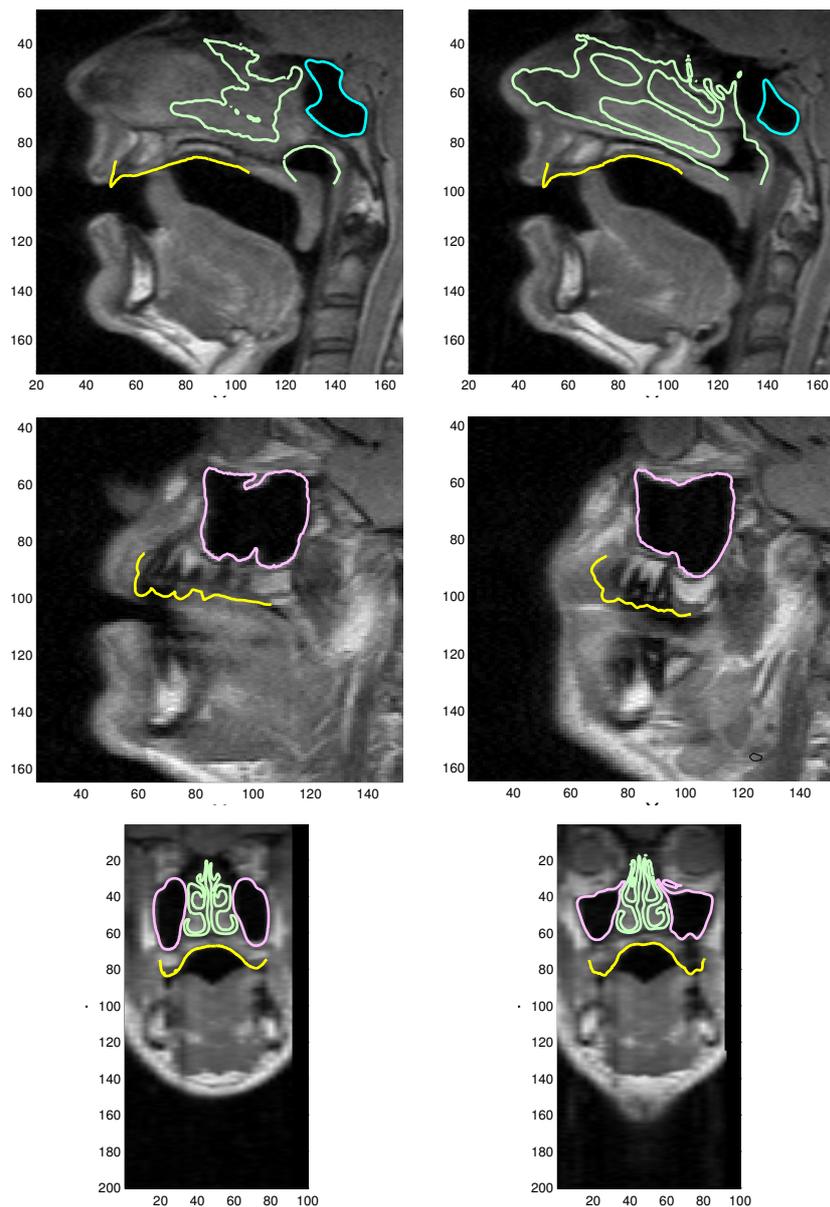


FIG. 2.22 – Superposition des recoupes des structures rigides fixes dans le repère absolu sur des images IRM du bloc correspondant au phonème /l^a/ après le recalage semi-automatique de ces structures : deux plans sagittaux proche du plan médiosagittal (haut), deux autres plans sagittaux éloignés de 2 à 2.5 cm du plan médiosagittal (milieu) et deux plans coronaux correspondant aux deux plans de grille numérotés 19 et 20 de la figure 2.8 (bas).

Il est donc maintenant possible d'exprimer l'ensemble des structures rigides fixes par rapport au crâne du sujet dans le repère absolu ou dans les différents repères liés aux blocs d'images. En d'autres termes, la rototranslation 3D de passage des repères liés aux blocs d'images IRM au repère absolu est connue. Il reste encore à déterminer la position de la mâchoire qui varie en fonction de l'articulation.

mâchoire sur chacun des 46 blocs d'images IRM sagittales. L'évaluation du calage obtenu est visible à la figure 2.23. La moyenne des erreurs RMS sur l'ensemble des 46 blocs d'images est légèrement inférieure à 0.05 cm tandis que la plus mauvaise erreur RMS se situe autour 0.07 cm. Ces chiffres montrent la fiabilité du processus pour le calage de la mâchoire. Quelques exemples de contours de la mâchoire calée sur le bloc d'images IRM correspondant à l'articulation /1^a/ sont visibles à la figure 2.24.

La connaissance de toutes les rototranslations de passage des repères liés aux blocs d'images vers le repère absolu permet de continuer le processus de détermination des formes 3D à partir des images IRM : d'une part les contours des structures rigides sont superposables sur les images IRM pour aider à leur interprétation et à la détection des structures déformables et d'autre part tous les contours extraits seront comparables entre eux dans un même repère. Avant de passer à l'étape de détermination des structures déformables, il est nécessaire de terminer le calage des données avec les enregistrements dynamiques EMA.

2.4.6 Alignement des enregistrements EMA

Le placement des enregistrements des bobines EMA dans le repère absolu ne pose pas de problème. La bidimensionnalité des données limite la rototranslation de calage à une rototranslation dans le plan médiosagittal, c'est-à-dire une translation et une rotation 2D. Pour déterminer la position des structures rigides, une bobine a été placée au milieu des incisives supérieures en guise de référence (voir figure 2.10). Sa position par rapport au point UT a été mesurée manuellement sur le sujet le jour de l'enregistrement. La translation est donc fixée par le vecteur calculé entre le point UT de référence et le point UT déterminé à partir de la bobine des incisives. Rappelons au passage que cette bobine est supposée fixe par rapport aux trois bobines émettrices, sa position au cours de l'enregistrement se limite donc à un point. Pour fixer la rotation, une des cinq bobines a été déplacée manuellement à la fin de l'enregistrement au contact du palais dur dans le plan médiosagittal, des incisives jusqu'au palais mou. L'enregistrement obtenu a donc permis de *dessiner* le palais dur dans le plan médiosagittal. Cette forme a ensuite été placée manuellement sur la recoupe du maillage du palais dur dans le plan médiosagittal, ce qui permet finalement d'obtenir la valeur de la rotation autour du point UT.

2.5 Les formes 3D des structures déformables

Contrairement aux structures rigides extraites sur les images CT qui présentaient des contours clairs, les structures déformables sont extraites à partir des images IRM pour lesquelles les contours entre tissus sont plus diffus. La méthode d'extraction nécessite donc l'usage d'informations supplémentaires, comme la superposition des contours des structures rigides déjà connus sur les images IRM pour aider la détermination des contours. Le principe d'extraction est détaillé au début de la section et chaque étape de la méthode est détaillée par la suite. Une approche similaire simplifiée est finalement proposée pour les contours médiosagittaux 2D.

2.5.1 Principe de détermination des structures déformables

Les formes 3D des structures déformables sont obtenues des images IRM de manière similaire à celles des structures rigides sur les images CT, c'est-à-dire par détermination manuelle des contours des organes sur les images elles-mêmes. Le passage de ces contours en px 2D vers les cm 3D donne un échantillonnage des surfaces 3D des organes en contours plans. La répétition de ce processus sur les 46 phonèmes du corpus procure une base de données de formes 3D de structures déformables échantillonnées en contours plans. Ces surfaces doivent posséder un nombre de points identique et cohérent sur l'ensemble de la base de données pour un même organe afin de pouvoir construire le modèle proprement dit. Ceci signifie en d'autres termes qu'un point sur la surface d'un organe dans une position précise doit avoir un correspondant sur le même organe dans une autre position. Il serait donc possible de *suivre* chaque point sur l'ensemble des formes extraites. Pour obtenir une telle représentation, un *maillage générique* unique de la surface de chacun des organe est déterminé. Il est ensuite *déformé* pour épouser chacune des 46 formes 3D de l'organe définie par ses contours plans. Cette méthode assure donc la représentation des 46 formes 3D des organes par l'intermédiaire d'un unique maillage pour chaque organe et permet ainsi un suivi des points de sa surface. Les différentes étapes du procédé sont alors les suivantes :

- Détermination des formes 3D des organes à partir des IRM suivant le processus décrit à la section 2.3. Les 46 formes, non maillées, sont appelées les *cibles*.
- Construction d'un maillage générique unique pour chacune des structures déformables.
- Déformation des maillages génériques vers les 46 cibles. Le résultat forme un jeu de maillages génériques déformés représentant chacune des 46 articulations du corpus.

Rappelons ici que les tracés sont effectués à la main directement sur les images IRM et dépendent donc de l'habileté de l'expert chargé de tracer. Comme expliqué au paragraphe 2.3.2, les résultats de [Soquet *et al.*, 2002] nous incitent cependant à considérer cette tâche assez fiable et les tracés obtenus seront considérés comme une référence absolue pour notre travail.

2.5.2 Le tracé : choix et évidences

La première étape consiste donc à extraire les formes 3D des structures déformables à partir des blocs d'images IRM sagittales. L'objectif consiste à obtenir les formes du voile du palais et de la paroi du pharynx dans la région du port vélopharyngé. Si les tracés des structures rigides sur les images CT avaient un caractère évident, ce n'est plus le cas pour les images IRM pour lesquelles d'une part la qualité est inférieure et d'autre part il s'agit de discriminer plusieurs tissus mous entre eux. Le travail de marquage est un travail difficile très important pour la suite de l'étude : il fournira un jeu de données important pour la modélisation articulaire. Un soin particulier lui est donc apporté.

L'enjeu de ces tracés est de déterminer la forme complète des organes déformables pour toutes les articulations du corpus de manière à construire un modèle articulaire et obtenir une géométrie précise du conduit vocal. On peut donc supposer que seules les régions des organes

constituant la limite du conduit vocal représentent un enjeu important et doivent être modélisées correctement. Néanmoins l'approche de modélisation par organe adoptée dans ce travail vise à représenter de manière réaliste les mouvements des organes dans le conduit vocal. Ceci implique donc une vue globale de l'organe constituant un tout cohérent et réaliste. Se pose alors la question cruciale des limites surfaciques des organes : si les régions des organes dont la surface constitue la limite du conduit vocal sont évidentes (l'air est en noir, les tissus en gris), les autres régions dans lesquelles de nombreux tissus s'entrecroisent (muscles, nerfs, veines, etc.) peuvent particulièrement poser problème. En particulier, la question de savoir où est la limite exacte de l'organe ne peut pas être résolue, cette limite n'existant pas réellement, puisque les différents tissus concernés sont imbriqués avec d'autres. Il est donc nécessaire de faire des choix. Ces choix sont par définition arbitraires mais guidés par un certain nombre de règles :

1. **Cohérence** : les tracés doivent être cohérents entre eux pour une même articulation du conduit vocal et cohérents avec les tracés des autres articulations. En d'autres termes, les mêmes structures doivent être identifiées sur toutes les articulations et les mêmes règles de traçage concernant ces structures doivent être appliquées.
2. **Lissage** : les tracés doivent décrire des formes réalistes sans aspérités ni variations qui ne seraient pas possibles pour un organe du conduit vocal (il est par exemple peu probable de trouver une pliure anguleuse pour un organe du conduit vocal). La multiplication de ce genre de cas rendrait par ailleurs hasardeuse l'étape de déformation du maillage générique vers une telle cible.
3. **Vraisemblance** : les choix effectués pour les tracés doivent être anatomiquement vraisemblables. En d'autres termes, ils doivent, ou doivent sembler lorsque l'interprétation est difficile, correspondre à des critères anatomiques définis et précis. Le contact entre les tracés de deux organes pour obtenir la fermeture effective du conduit vocal dans les régions latérales, qui est une réalité physique, en est une illustration.

Ces règles générales orientent les choix à effectuer pour déterminer les tracés. Elles sont néanmoins laissées à la libre interprétation du traceur à qui revient *in fine* de décider les tracés.

En pratique chaque bloc d'images est composé de 52 images, 25 images sagittales originales et 27 images transverses recoupées (voir la figure 2.8 pour l'orientation des 27 images recoupées). Pour augmenter les chances de détection correcte des contours de l'organe et augmenter leur densité spatiale, ceux-ci sont tracés dans les deux jeux d'images à la fois. Ce choix permet d'exploiter la complémentarité de ces deux jeux :

- Le jeu d'images sagittales permet une bonne détection du voile du palais et de la paroi du pharynx dans un voisinage du plan médiosagittal.
- Le jeu d'images transverses permet une bonne détection dans la région vélaire transverse à l'exclusion des extrêmes pour cette orientation que sont la luette et le contact avec palais dur. Il permet notamment de clarifier la détermination des contours dans les régions latérales où les surfaces du voile du palais et de la paroi du pharynx deviennent tangentes aux plans sagittaux.

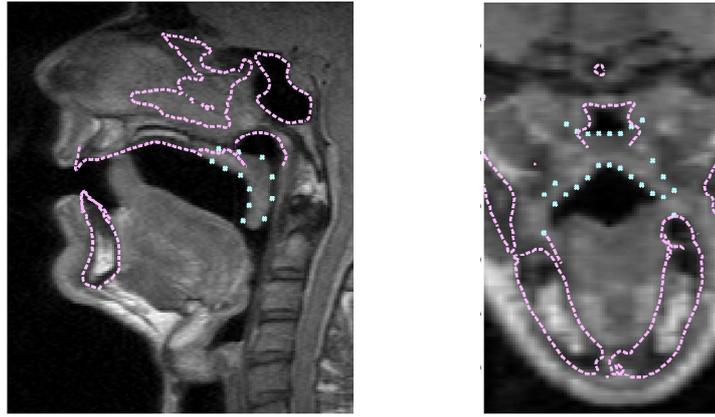


FIG. 2.25 – Superposition sur l'image médiosagittale (gauche) et sur l'image transverse numéro 16 de la figure 2.8 (droite), pour le phonème /l^a/, des recoupes des structures rigides (pointillés) et des contours du voile du palais (points) tracés dans les plans d'images du jeu complémentaire à l'image concernée.

Pour exploiter efficacement l'information apportée par l'un des jeux dans l'autre jeu, tout contour tracé sur une image est immédiatement recoupé dans les plans d'images constituant le jeu complémentaire. Ces points sont alors superposés sur l'image du jeu complémentaire. L'ensemble des recoupes des contours d'un jeu fournit donc dans l'autre jeu un ensemble de points qui constituent une description de l'organe fort utile pour aider à tracer ; ils *dessinent* dans le second jeu l'organe tel qu'il ressort sur le premier jeu. Il reste à tracer un contour qui s'appuie sur les points. La figure 2.25 montre l'information apportée à l'image par ces recoupes.

Par ailleurs, pour aider à l'interprétation des images, les contours suivants, en plus des points correspondants aux recoupes des tracés du jeu complémentaire, sont systématiquement superposés aux images :

- Les recoupes des structures rigides (voir par exemple les contours en pointillés de la figure 2.25)
- Les tracés déjà disponibles sur la même image des autres structures déformables (voir par exemple les contours en traits-pointillés de la figure 2.28)

On notera au passage que le *processus ptérygoïde* a été tracé sur les images CT et maillé, conformément à la méthode décrite dans la section 2.3 ; il a ensuite été ajouté à la liste des structures rigides dont les intersections sont superposées sur les images. Cette structure osseuse entourant le cavum et descendant sur le côté jusqu'au voile du palais permet de circonscrire les limites de ce dernier dans les régions latérales.

Pour faire converger les tracés issus des deux jeux de manière cohérente en respectant les règles définies ci-dessus, les tracés sont effectués méthodiquement et itérativement du plus évident vers le moins évident : d'abord dans le plan médiosagittal, puis dans les autres plans sagittaux en s'écartant petit à petit du plan médiosagittal, puis dans les plans transverses qui fournissent d'importantes informations pour les plans sagittaux, puis à nouveau dans les

plans sagittaux encore plus éloignés du plan médiosagittal, puis à nouveau dans les plans transverses, etc... puis itérativement dans les plans sagittaux et transverses pour faire converger petit à petit de manière cohérente les deux jeux de tracés.

Deux organes sont à tracer : le voile du palais et le pharynx dans la région du port vélopharyngé. Les caractéristiques de leurs tracés sont détaillés ci-dessous.

Le voile du palais

La surface du voile du palais est considérée comme une surface fermée. Les tracés sagittaux du voile du palais autour du plan médiosagittal ne posent pas de problème. L'édition du voile du palais dans deux plans proches du plan médiosagittal est visible sur les deux premières images de la figure 2.28 pour une articulation /l^a/. En s'éloignant encore du plan médiosagittal, le port vélopharyngé se referme : il convient donc de définir sa frontière avec la paroi du pharynx. Les tracés sont alors guidés par l'interprétation des images IRM transverses proposées à la figure 2.26. Nous cherchons à inclure les régions correspondant aux muscles définis comme appartenant au voile du palais (ici le levator veli palatini et le palatoglosse) et à contourner les autres structures. Une interprétation anatomique des images IRM est proposée aux figures 2.26 et 2.27. La résolution relativement faible des images associée à la précision requise rendent l'interprétation difficile. Néanmoins, les interprétations paraissent cohérentes avec celles de [Ettema *et al.*, 2002]. Suivant l'interprétation proposée sur l'image transverse présentée à la figure 2.26, le contour du voile du palais dans sa partie supérieure doit alors remonter pour entourer le levateur en passant sous l'ouverture cartilagineuse de la trompe d'Eustache ; il redescend ensuite en contournant le muscle masséter pour aller entourer le palatoglosse et rejoindre le bord du conduit oral où le tracé est alors évident. Les tracés sont par ailleurs étendus dans les régions latérales au-delà de la limite indiquée du palatoglosse, jusqu'à la mâchoire, pour assurer la fermeture du conduit oral et accrocher le contour sur une référence fixe et sûre (voir figure 2.29). Ces tracés dans les plans transverses se traduisent dans les plans sagittaux par une remontée du contour du voile du palais depuis la pointe du bas vers le haut pour aller entourer l'îlot constitué par le levator veli palatini visible à la figure 2.27. La figure 2.28 présente quelques exemples de tracés dans ces cas.

Ces interprétations définissent les grandes lignes des tracés pour le voile du palais. Ceux-ci restent néanmoins dépendants des décisions de l'opérateur lors de la tâche en elle-même. La cohérence entre les différentes articulations du corpus est par ailleurs assurée par la reconnaissance des mêmes formes grisées sur les images qui permettent de déduire le tracé. Un exemple de tracés pour l'articulation /l^a/ est montré sur les figures 2.28 et 2.29. Deux remarques peuvent être formulées en regardant ces tracés :

- Le contour du voile du palais peut être constitué de plusieurs sous contours, en particulier dans les plans transverses au niveau de la luette, comme le montre la première image de la figure 2.29.
- La surface du voile du palais étant considérée comme une surface fermée, l'intersection par

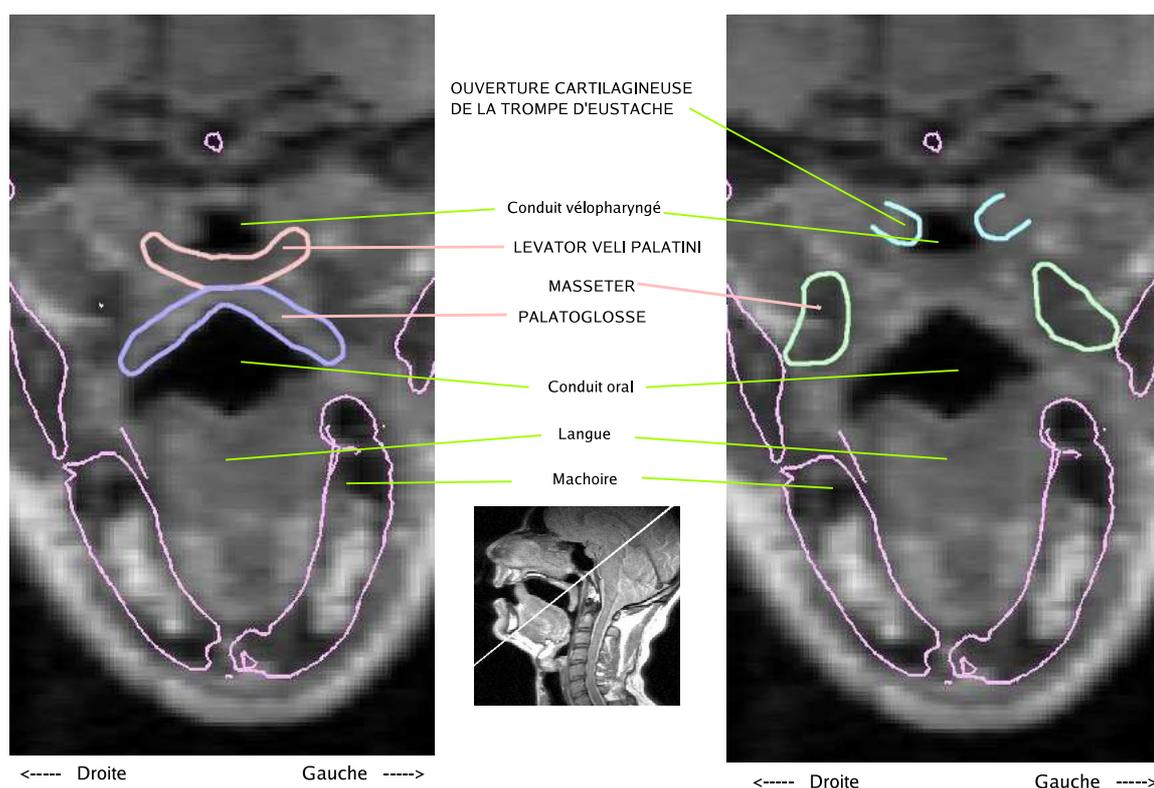


FIG. 2.26 – Proposition d'interprétation anatomique de l'image IRM transverse numérotée 16 de la grille de la figure 2.8 pour le phonème /l^a/. Les contours rigides sont obtenus par recoupe des maillages et superposés à l'image tandis que les contours des structures déformables sont dessinés manuellement pour les besoins de l'explication.

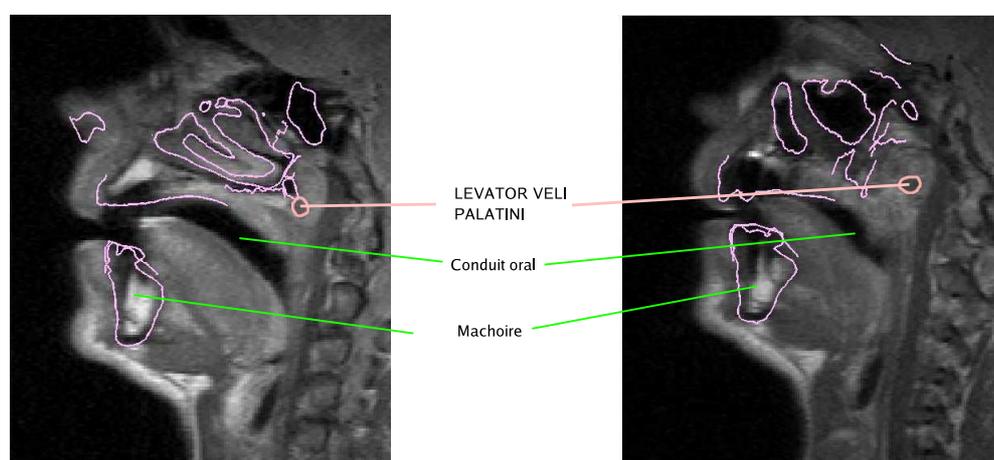


FIG. 2.27 – Recoupes des structures rigides et position du muscle levateur du voile du palais sur les images IRM sagittales à 1.2 cm (gauche) et 1.6 cm (droite) du plan médiosagittal pour l'articulation /l^a/. Le contour de ce dernier est dessiné manuellement pour les besoins de l'explication, à l'aide des tracés obtenus sur la figure 2.26.

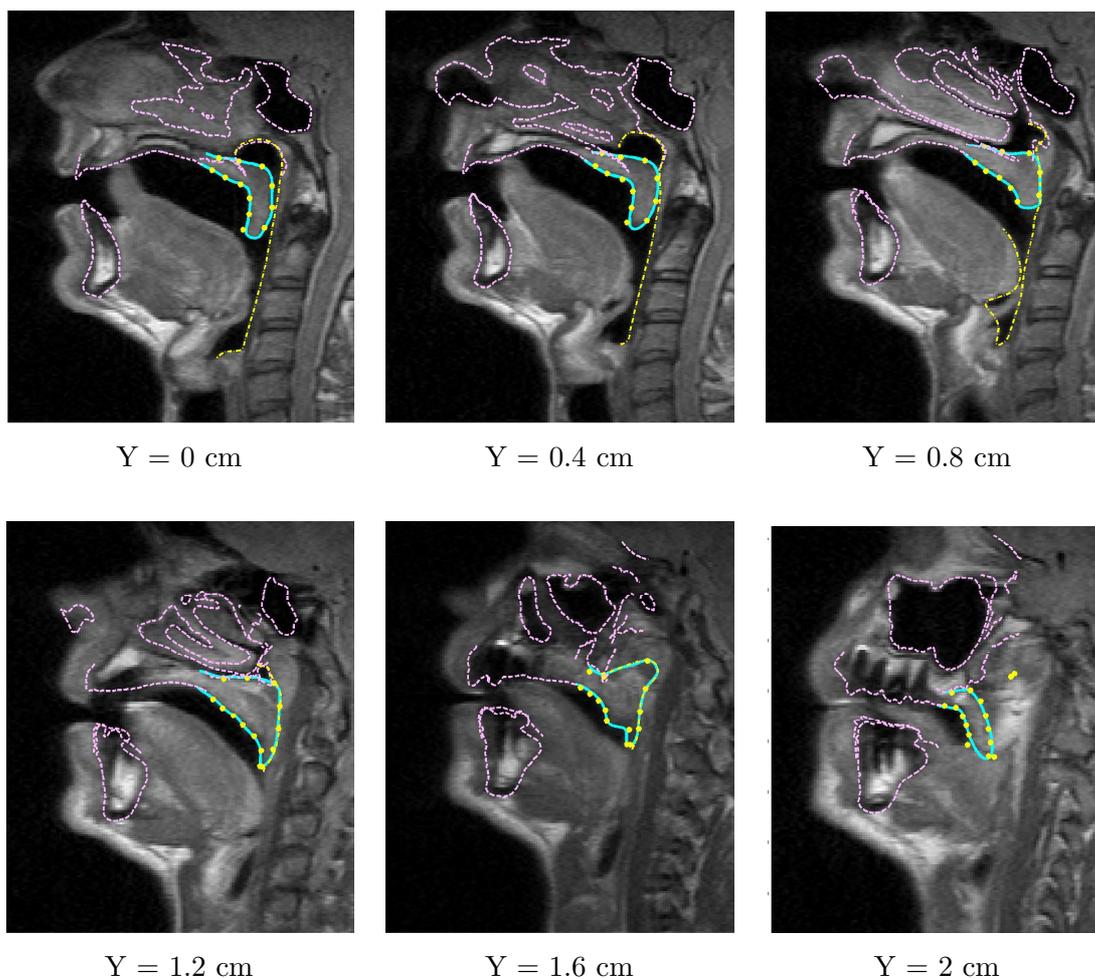


FIG. 2.28 – Contours du voile du palais pour le phonème /l^a/ dans les plans d’images sagittales espacées de 0.4 cm entre elles. Le trait plein représente le contour du voile du palais ; les contours en pointillés représentent les recoupes des structures rigides superposées sur l’image ; les points représentent les recoupes des tracés du voile du palais dans les plans transverses et les contours en points-pointillés représentent les contours des autres structures déformables (c’est-à-dire la paroi du pharynx) tracées sur cette image.

un plan donne un contour fermé. Tous les contours tracés sur les plans d’image sont donc fermés, à l’exception de la connexion avec le palais dur.

La paroi du pharynx

Les parois arrière et latérales sont considérées comme une unique surface qui est, à l’inverse du voile du palais, ouverte. L’étude portant sur les nasales, nous ne considérerons que la région du pharynx autour du port vélopharyngé. Elle s’étend donc verticalement du bas de la luette jusqu’aux structures rigides de l’os sphénoïdal et autant que nécessaire pour assurer la fermeture du conduit vélopharyngé horizontalement. Les choix de tracés s’avèrent beaucoup plus simples pour les contours de la paroi du pharynx qui vont compléter ceux du voile du palais de manière

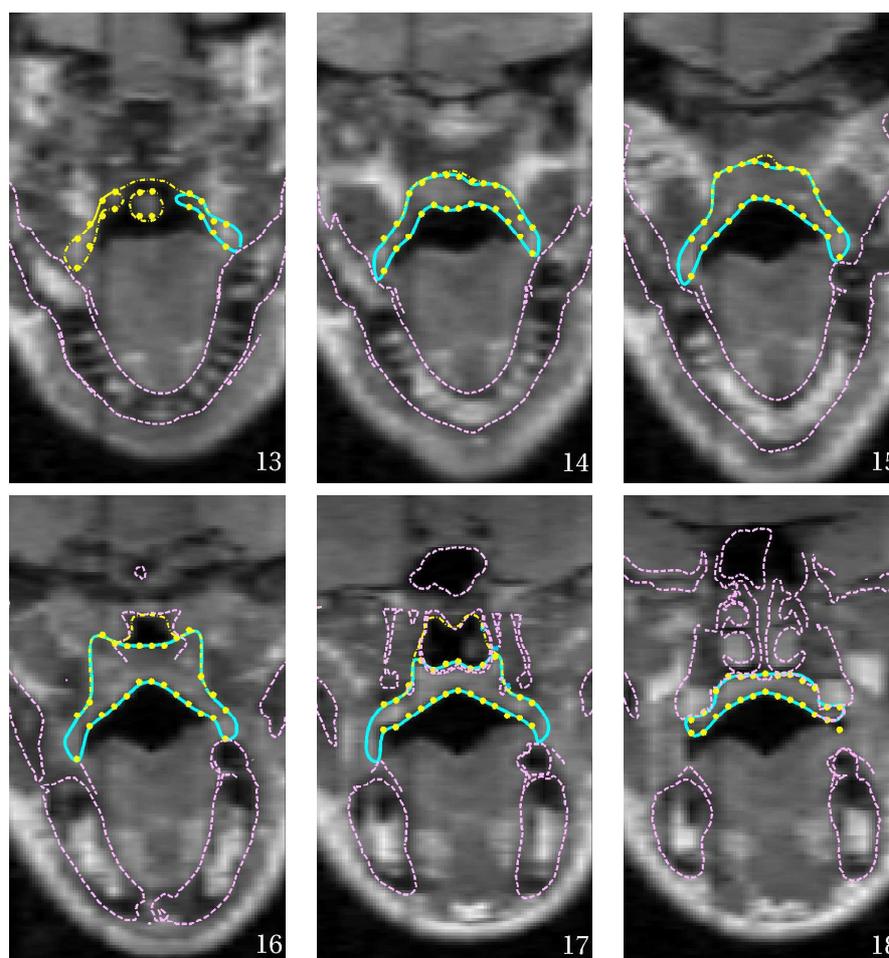


FIG. 2.29 – Contours du voile du palais pour le phonème /l^a/ dans les plans d'images transverses correspondant aux plans 13 à 18 de la région vélaire visible figure 2.8. Le trait plein représente le contour du voile du palais (sur la première image, il représente seulement une portion du voile du palais constitué de trois sous-contours) ; les contours en pointillés représentent les recoupes des structures rigides superposées sur l'image ; les points représentent les recoupes des tracés du voile du palais dans les plans transverses et les contours en points-pointillés représentent les contours des autres structures déformables (c'est-à-dire la paroi du pharynx) tracées sur cette image.

à former un conduit vélopharyngé fermé. Ainsi, en dehors des régions où le contour de la paroi du pharynx délimite clairement la limite du conduit vocal (frontière tissu / air clairement visible sur les IRM), celui-ci est au contact de celui du voile du palais. Dans le cas des phonèmes oraux, la fermeture du port vélopharyngé se produit effectivement grâce au contact entre la paroi du pharynx et le voile du palais, contrairement aux voyelles et consonnes nasales. Dans les régions latérales où il n'y a plus de conduit, la fermeture est encore une fois assurée par le contact entre les contours du voile du palais et de la paroi du pharynx. La notion de contact entre les deux organes est artificielle ici car en réalité les fibres des deux organes se mélangent pour former un tissu unique. La figure 2.30 illustre les divers cas de figures mentionnés ci-dessus.

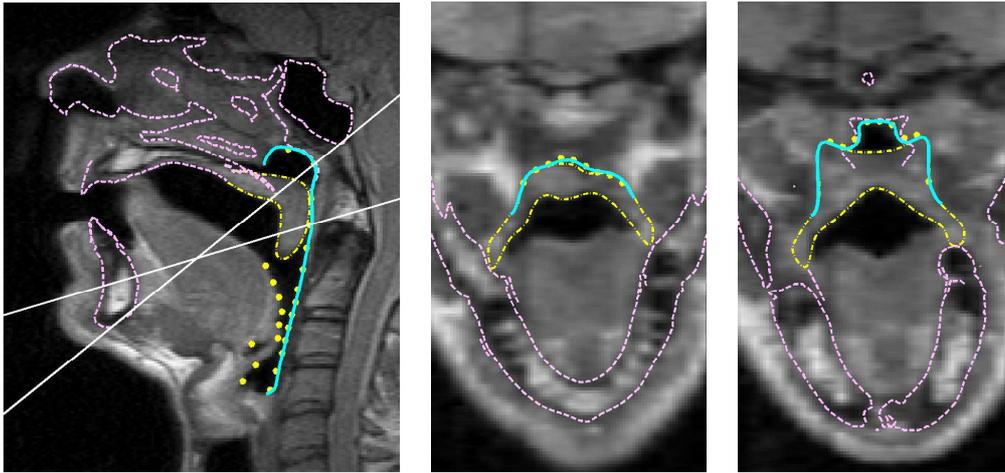


FIG. 2.30 – Contours de la paroi du pharynx pour le phonème /l^a/ dans un plan sagittal proche du médiosagittal (gauche) et dans les deux plans transverses marqués par des lignes droites sur l'image de gauche (milieu et droite). Le trait plein représente le contour de la paroi du pharynx (sur la première image, elle s'étend de la glotte au cavum mais seule la région du port vélopharyngé fait l'objet de l'étude présente) ; les contours en pointillés représentent les recoupes des structures rigides superposées sur l'image ; les points représentent les recoupes des tracés de la paroi du pharynx dans les plans transverses et les contours en points-pointillés représentent les contours des autres structures déformables (c'est-à-dire du voile du palais) tracées sur cette image.

La description précédente des choix effectués pour déterminer les tracés des organes déformables a permis de définir une forme générale considérée comme la surface de l'organe. Toutes les régions de cette surface globale ne sont cependant pas utiles pour la modélisation propre. Comme nous l'avons expliqué précédemment, chaque organe doit être vu comme un tout logique et cohérent. Cette vision est très importante dans la phase de déformation du maillage générique vers les cibles. Néanmoins, une fois le maillage générique complet de l'organe déformé vers l'ensemble des positions du corpus, il peut être étudié par morceaux, sachant que le mouvement de chaque morceau, en tant qu'entité appartenant au maillage complet, est tout à fait cohérent. Cette propriété est importante pour exclure de la construction du modèle les régions qui risquent de rajouter beaucoup de bruit sans apporter d'information pour la détection des fonctions d'aire. C'est le cas en particulier des régions latérales supérieures du voile du palais qui sont toujours en contact avec des tissus mous autres que la langue et la paroi du pharynx. Pour exclure de la construction du modèle (mais non pas de l'étape de déformation du maillage générique vers les cibles) ces régions schématisées par des portions en pointillés sur la figure 2.31 et qui représentent le passage entre les muscles du voile du palais et d'autres tissus comme le muscle masséter (voir figure 2.26), la partie considérée comme utile pour l'analyse sera limitée à une région comprise entre - 1.5 cm et + 1.5 cm du plan médiosagittal. Ceci correspond à la région comprise entre les deux lignes verticales en pointillées sur la figure 2.31. Cette sélection exclut cependant certaines

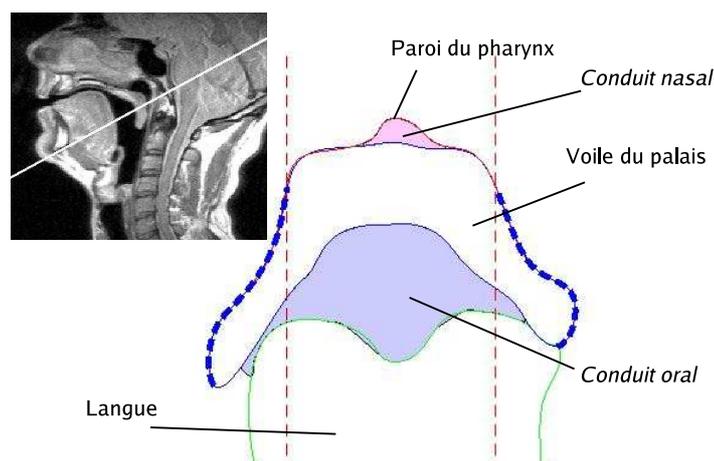


FIG. 2.31 – Contours du voile du palais, de la paroi pharyngée et de la langue dans le plan transverse indiqué sur l'image IRM du coin supérieur gauche pour une articulation /i/. Les conduits oral et nasal sont légèrement grisés. Les lignes verticales horizontales à 1.5 cm des deux côtés du plan médiosagittal délimitent la région utilisée pour la construction du modèle articulaire de voile du palais ; les contours du voile du palais en pointillés au-delà de ces lignes verticales sont considérés comme non pertinents pour le modèle.

régions pourtant pertinentes situées au-delà de la limite, comme la partie inférieure du voile du palais proche de la langue. On fera l'hypothèse que la surface du voile du palais au contact du conduit oral possède un mouvement continu de la région médiosagittale jusqu'aux extrémités latérales : ceci permet de supposer que les mouvements extérieurs, exclus de la construction du modèle, peuvent être efficacement extrapolés à partir des mouvements centraux. La paroi du pharynx quant à elle, dont les tracés ne dépassent pas les limites fixées à - 1.5 cm et + 1.5 cm du plan médiosagittal est conservée dans sa totalité.

Le travail de tracé constitue une étape importante du processus de détermination des formes 3D à partir des IRM. Il nous a permis de constituer une base de donnée de 46 cibles du voile du palais et de la paroi de pharynx dans chacune des positions du corpus. Il a également représenté une part importante en termes de temps, mais également en termes de dilemme, ce qui s'est traduit par plusieurs versions de tracés. On peut considérer que le travail complet de calage des données entre elles et d'obtention de tracés fiables pour la construction d'un modèle s'est ainsi étalé sur un an à un an et demi. Ces tracés constituent maintenant une donnée précieuse pour la modélisation articulaire. L'étape suivante consiste à créer les deux maillages génériques qui seront ensuite déformés vers chacune de ces 46 cibles.

2.5.3 Les maillages génériques

Les maillages génériques du voile du palais et de la paroi du pharynx ont été obtenus de la même manière que ceux des structures rigides : par maillage des points formant les contours

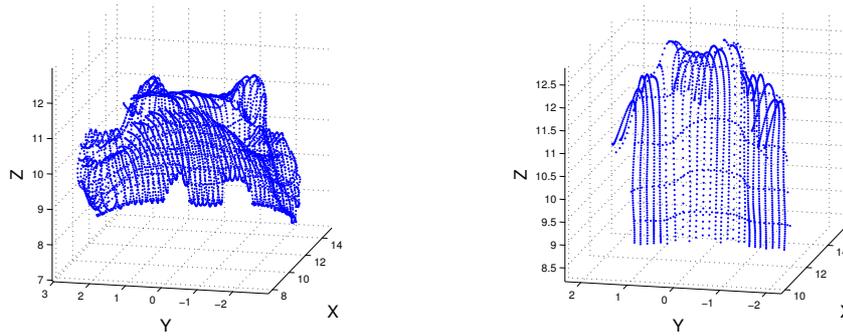


FIG. 2.32 – Vue 3D antérieure dans le repère absolu de l'ensemble des points formant les contours plans du voile du palais (gauche) et de la paroi du pharynx (droite) tracés pour obtenir le maillage générique.

plans de l'organe et définissant ainsi un échantillonnage 3D. Nous avons choisi les points des tracés de la voyelle nasale $/\bar{a}/$ pour laquelle la position basse du voile du palais assure un contact minimal avec les structures environnantes et permet ainsi des tracés plus étendus et plus fiables. Cependant, en comparaison des images CT, les images IRM ont une faible résolution. Pour augmenter à la fois la fiabilité de l'étape de maillage et la densité du maillage générique résultant, le bloc d'images IRM correspondant à l'articulation $/\bar{a}/$ a été réinterpolé de 25 à 44 images sagittales, introduisant deux nouvelles images sagittales entre deux originales, en plus des 27 images déjà réinterpolées au paragraphe 2.2.5. Les tracés présentés au paragraphe précédent ont alors été étendus aux nouvelles images. Tous ces tracés forment alors un échantillonnage des organes exprimé dans le repère de référence (figure 2.32). Ils sont ensuite maillés puis légèrement lissés comme les structures rigides l'ont été au paragraphe 2.3.3. Le maillage du voile du palais est composé de 10464 triangles et 5239 points espacés en moyenne de 0.13 cm tandis que celui de la paroi du pharynx de 4037 triangles et 2110 points espacés en moyenne de 0.14 cm. L'erreur RMS de reconstruction du maillage est de 0.03 cm pour le voile du palais et de 0.04 cm pour la paroi du pharynx, ce qui permet de considérer comme presque négligeable l'erreur introduite par cette étape. Les maillages génériques sont visibles à la figure 2.33.

2.5.4 Déformations des maillages génériques

Cette étape est une étape clé dans l'obtention pour chaque organe des 46 formes 3D cohérentes entre elles. Elle est réalisée par le logiciel de déformation élastique *TestRigid* développé au laboratoire TIMC à Grenoble. Le principe de déformation est détaillé par [Szeliski et Lavallée, 1996]. On supposera que les déformations appliquées par le logiciel aux maillages génériques surfaciques sont conformes aux déformations réelles de l'organe en parole. Le logiciel déforme élastiquement le maillage de référence pour l'ajuster au mieux sur un jeu de points constituant la cible. Naturellement le maillage n'est pas déformable à l'excès et une certaine rigidité est donnée à l'organe. Le principe est de découper le maillage en cubes élémentaires ; les cubes sont alors déformés en fonction d'un algorithme de minimisation de distance tout en restant soudés les uns aux autres ; les nouveaux points du maillage sont alors calculés en fonction de leur position par

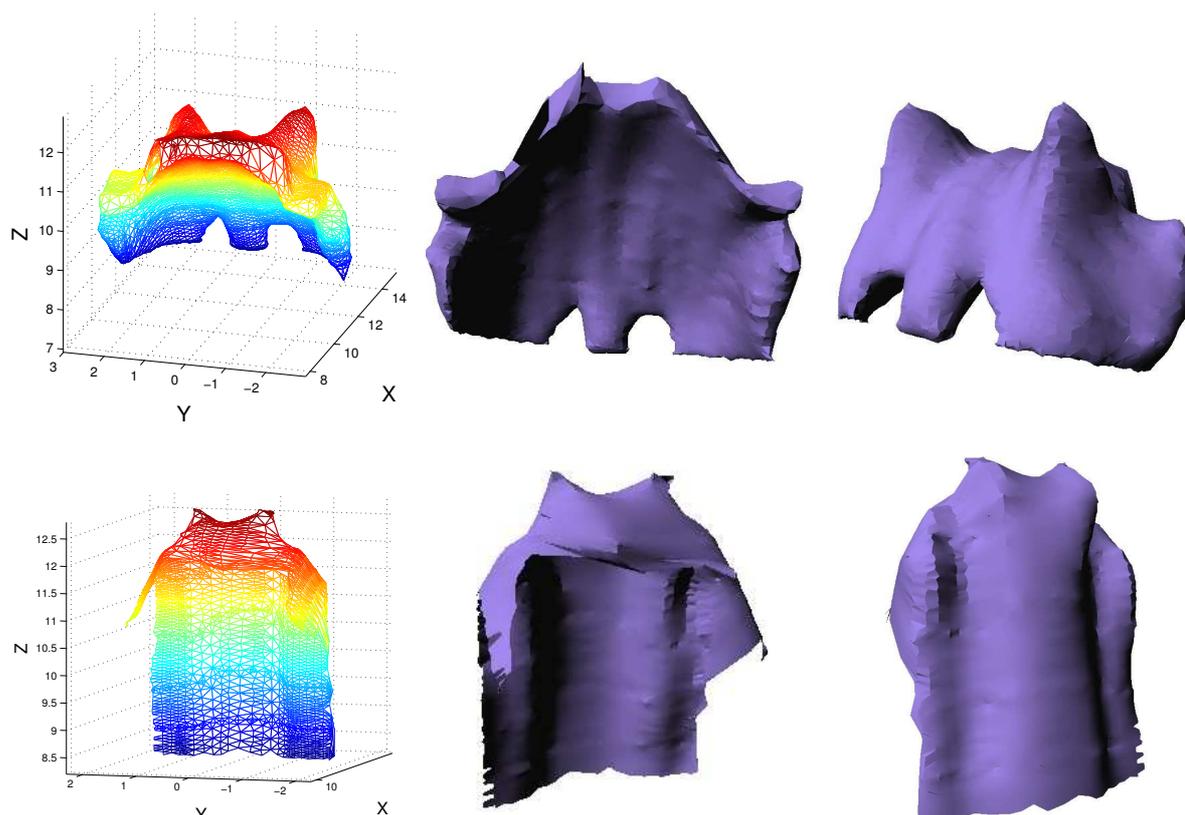


FIG. 2.33 – Maillages génériques du voile du palais (haut) et de la paroi du pharynx (bas) : vue antérieure non texturée (gauche), vue antérieure texturée (milieu) et vue postérieure texturée (droite).

rapport aux sommets du cube élémentaire.

La tâche demandée à l'algorithme est relativement compliquée, en particulier pour le voile du palais : le passage par exemple du maillage en position $/\bar{a}/$ vers une position $/p^a/$ déplace dans le plan médiosagittal le contour du voile du palais d'une distance supérieure à son épaisseur ; il y a donc un risque que les parois du maillage se collent à la mauvaise cible, ce qui correspondrait à un minimum local. Pour éviter ces cas et pour améliorer le processus de déformation, ce dernier est décomposé pour chaque articulation du corpus en plusieurs ajustements sur des cibles intermédiaires interpolées entre la référence la cible réelle. Ces ajustements successifs permettent de *guider* le maillage vers la cible finale. En pratique, les cibles intermédiaires sont calculées par interpolation linéaire des contours sources vers les contours cibles dans les plans contenant les contours cibles. Rappelons en effet que la cible est composée de points qui forment les contours tracés dans les plans d'images tandis que la source est un maillage dont l'intersection avec n'importe quel plan peut être calculée, et en particulier avec les plans contenant les contours de la cible. Les contours sont donc interpolés dans les plans d'images sagittales et transverses. La réunion des $i^{\text{èmes}}$ contours intermédiaires de chaque plan constitue la $i^{\text{ème}}$ cible intermédiaire 3D pour le maillage générique en direction de la cible finale. La figure 2.34 montre par exemple

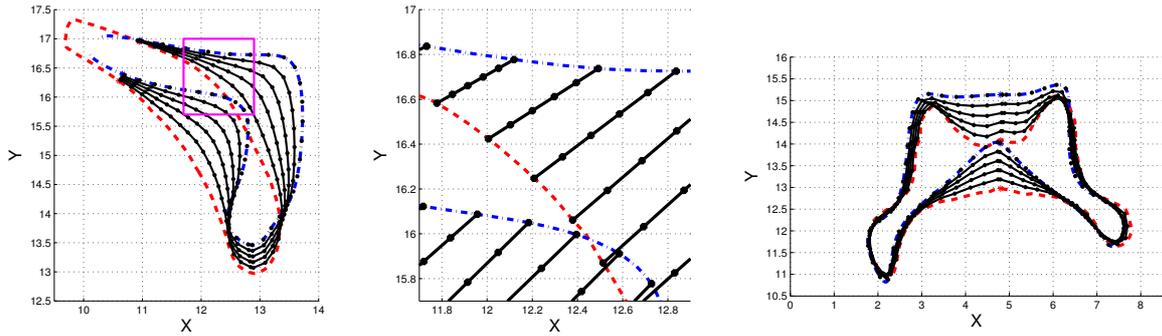


FIG. 2.34 – Contours du maillage générique du voile du palais (pointillés en position *basse*) et de la cible correspondant à l’articulation $/p^a/$ (trait-pointillés en position *haute*) dans le plan médiosagittal (gauche) et un plan transverse (droite). Les contours intermédiaires sont obtenus en reliant entre eux les points obtenus par une interpolation linéaire entre les points du contour du maillage générique et de la cible, comme montré sur la figure centrale (zoom de la région encadrée sur la figure gauche). Le plan de représentation de la figure droite correspond à la grille numérotée 16 de la figure 2.8.

quatre contours interpolés entre les contours génériques et les contours tracés de la cible $/p^a/$. On note au passage que tous les contours se rejoignent bien comme il se doit au contact du palais dur (image gauche). L’assemblage des contours intermédiaires sur l’ensemble des plans d’interpolation forme les cibles intermédiaires. Elles ne constituent que des indications de guidage et seule la cible finale, composée des vrais tracés effectués sur les images comme expliqué au paragraphe 2.5.2, est considérée comme la réalité. Par ailleurs, pour forcer le logiciel à bien converger pour certains points de référence du maillage générique, il est possible d’associer manuellement des points sources et des points cibles. Cette propriété du logiciel a notamment été utilisée pour la pointe de la luvette et les deux pointes latérales (visibles par exemple sur la figure 2.33 en haut à droite).

Le même procédé a été appliqué à la paroi du pharynx qui converge plus facilement en raison d’un écart relativement faible entre le maillage générique et les cibles.

Un exemple de maillages génériques du voile du palais et de la paroi du pharynx déformés par ce procédé vers la cible $/p^a/$ est montré à la figure 2.35.

L’erreur RMS entre les points cibles de chacune des 46 articulations et les maillages génériques du voile du palais et de la paroi du pharynx déformés par le logiciel vers ces cibles est visible à la figure 2.36. L’erreur RMS moyenne pour le voile du palais est de 0.06 cm tandis qu’elle est de 0.04 cm pour la paroi du pharynx. L’erreur RMS du voile du palais calculée dans la région centrale limitée à 1.5 cm de chaque côté du plan médiosagittal, région utilisée pour la future construction du modèle, est légèrement inférieure ; la moyenne de la distance maximale calculée sur chaque articulation entre les points cibles et le maillage déformé passe notamment de 0.31 cm à 0.21 cm. Ceci suggère que les points les moins bien atteints par le maillage déformé ont

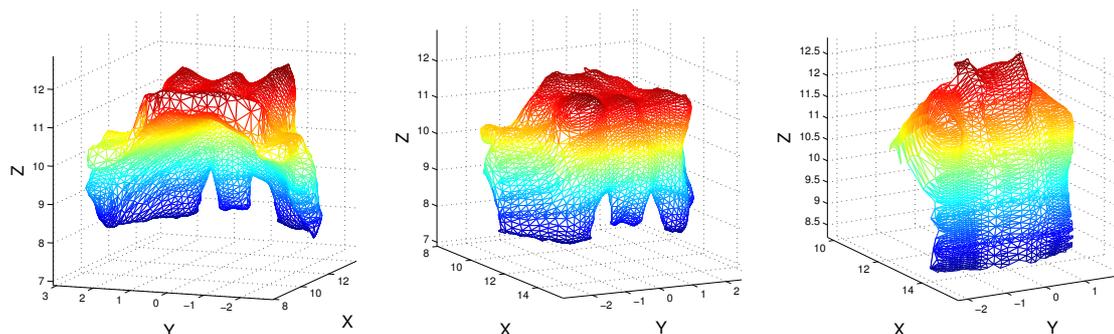


FIG. 2.35 – Vue antérieure (gauche) et postérieure (milieu et droite) des maillages génériques du voile du palais (gauche et milieu) et de la paroi du pharynx (droite) déformés vers la cible $/p^a/$.

tendance à se situer en dehors de la région située entre ± 1.5 cm du plan médiosagittal.

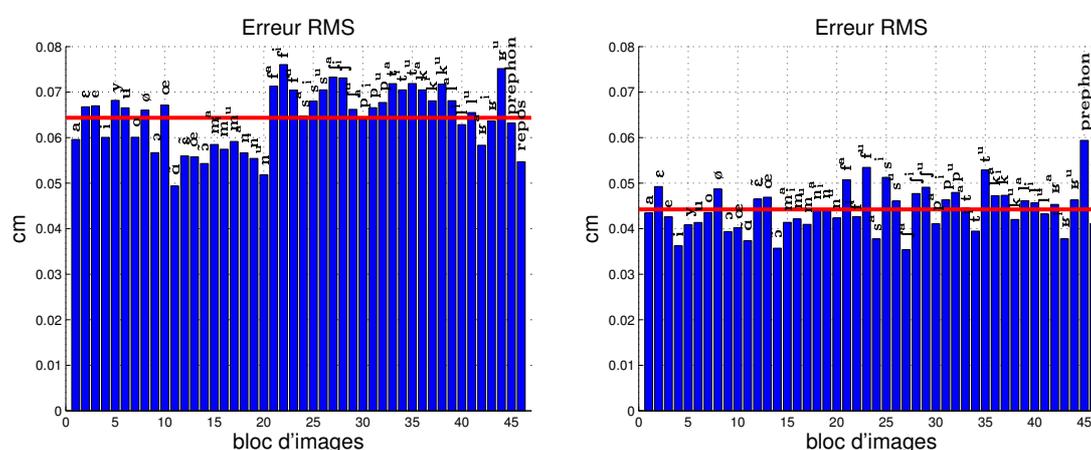


FIG. 2.36 – Erreurs RMS entre les points cibles des 46 articulations considérés comme la référence et les maillages génériques du voile du palais (gauche) et de la paroi du pharynx (droite) déformés dans les 46 positions correspondantes. Les traits continus désignent les moyennes.

2.5.5 Une approche similaire en deux dimensions

L'approche proposée dans le paragraphe précédent permet d'obtenir un jeu cohérent de formes 3D du voile du palais et de la paroi du pharynx. Une approche similaire est appliquée aux contours médiosagittaux pour construire par la suite un modèle purement médiosagittal qu'il sera intéressant de comparer au modèle tridimensionnel. L'ensemble des contours médiosagittaux du voile du palais et de la paroi du pharynx sont obtenus par la recoupe dans le plan médiosagittal des maillages génériques déformés dans chacune des 46 positions du corpus (figure 2.37 gauche). Chaque contour est alors décomposé en deux sous-contours : un contour inférieur, du palais dur jusqu'à la pointe de la luette, dessinant dans cette région le bord supérieur du conduit oral, et un contour supérieur, de la pointe de la luette jusqu'au palais dur, dessinant dans cette région le

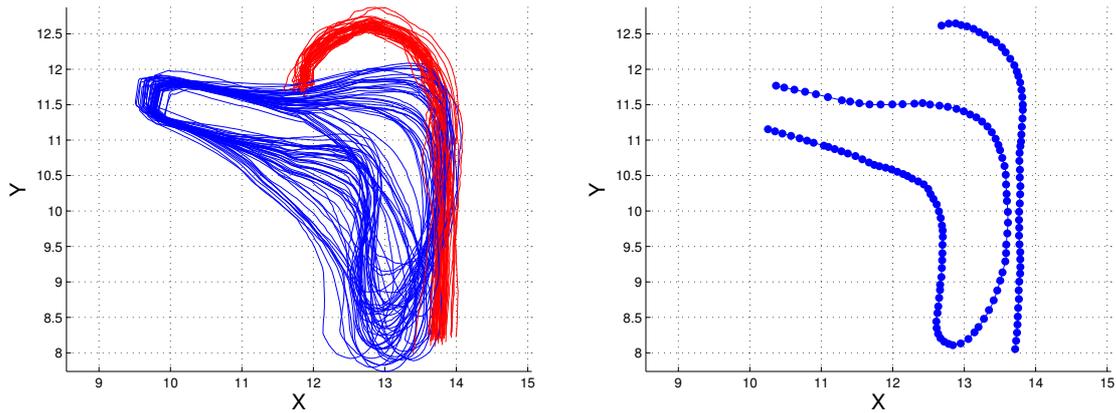


FIG. 2.37 – Superposition des 46 recoupes médiosagittales des maillages génériques du voile du palais et de la paroi du pharynx déformés vers les 46 cibles du corpus (gauche) et contours rééchantillonnés en 2D pour l’articulation $/p^a/$ dont un point sur deux a été affiché pour plus de lisibilité (droite).

bord antérieur du port vélopharyngé, décrivant ainsi le contour du voile du palais dans le sens trigonométrique. Pour obtenir un nombre similaire de points pour les 46 articulations, chacune des deux courbes est réinterpolée par une spline contenant exactement 101 points équirépartis. Il est ainsi possible d’associer entre eux les points issus des diverses articulations. Si on assimile ces points à des points de chair, on suppose d’une part que le point de chair correspondant à la pointe de la lchette est toujours le point le plus bas et d’autre part que la déformation élastique du voile du palais dans le plan médiosagittal est équirépartie entre la pointe de la lchette et le palais dur. En d’autres termes, on suppose une élasticité uniforme des contours médiosagittaux : il n’y a pas de régions qui se *tendent* plus ou moins que d’autres lors des mouvements du voile du palais. Un exemple de rééchantillonnage correspondant à l’articulation $/p^a/$ est visible à la figure 2.37 (droite).

La même approche est appliquée à la paroi arrière du pharynx. Chacun des 46 contours de la paroi du pharynx est réinterpolé par une spline de 101 points équirépartis entre le point du pharynx situé au niveau de la pointe de la lchette et le haut du pharynx au niveau du cavum. En assimilant les points obtenus à des points de chair, on suppose également que les points extrêmes correspondent toujours aux mêmes points de chair réels et que la déformation de la paroi est élastiquement uniforme.

Ce procédé fournit des contours exactement superposables aux contours obtenus des recoupes des maillages génériques : la moyenne des erreurs RMS sur l’ensemble des 46 articulations entre les points réinterpolés et les contours est inférieure au dixième de millimètre pour les deux organes.

2.5.6 Comparaison avec les données articulatoires dynamiques

Une première validation de l'approche peut être effectuée en comparant les résultats obtenus ci-dessus en 2D et en 3D avec les données dynamiques EMA. Ces dernières offrent l'avantage de représenter les trajectoires réelles de certains points de surface des organes, contrairement aux images IRM qui permettent d'obtenir des formes géométriques sans notion de point de chair. Elles sont donc complémentaires. La déformation du maillage générique du voile du palais vers les 46 cibles articulatoires 3D et le rééchantillonnage des contours 2D présentés ci-dessus peuvent être considérés comme une *imitation* de la déformation réelle de la surface de l'organe. La comparaison des résultats obtenus avec les données EMA, qui sont des enregistrements de trajectoires de points de chair réels, permet d'évaluer cette approche. On s'intéresse pour cela à l'enregistrement EMA de la trajectoire d'un point de la surface du voile du palais situé dans le plan médiosagittal environ à équidistance entre la pointe de la luvette et le palais dur (voir figure 2.10 page 52). La démarche présentée ici consiste à trouver le point du maillage générique du voile du palais qui correspond au point fictif sur lequel est collé la bobine EMA puis de comparer son domaine de déplacement sur les 46 articulations du corpus, censées couvrir l'ensemble des capacités articulatoires du sujet pour une tâche de parole, avec les enregistrements EMA.

Pour chacune des 44 articulations statiques du corpus IRM correspondant à de la parole, c'est-à-dire les 46 articulations sauf celles du *repos* et de la *préphonation*, une position de la bobine EMA du voile du palais a été calculée comme la moyenne de 10 réalisations en position cible pour les 14 voyelles et d'une réalisation en position cible pour les 30 consonnes. Chaque articulation statique du corpus IRM a alors été associée avec la position EMA correspondante. Pour chaque articulation, les points du maillage générique du voile du palais sont ordonnés en fonction de leur distance croissante au point EMA correspondant. Un rang général qui correspond à son rang le plus élevé sur l'ensemble des articulations dans le classement précédent a été attribué à chaque point du maillage. Le *point équivalent EMA* du maillage a finalement été choisi comme celui ayant un rang général minimal. Ceci revient à calculer le point dont la distance maximale au point EMA sur l'ensemble des 44 articulations est la plus faible. Une erreur RMS des distances entre le *point équivalent EMA* dans les 44 articulations possibles et la position de la bobine EMA est calculée. Elle s'élève à 0.33 cm. Cette forte erreur doit cependant être relativisée en considérant que les deux corpus, celui de l'IRM et des enregistrements EMA, sont difficilement comparables directement. Le premier représente un jeu de 46 phonèmes statiques hyper-articulés enregistrés en position allongée tandis que le second représente des enregistrements de pseudo mots de parole dynamique en position verticale du sujet. Les cibles atteintes dans chacun des deux enregistrements sont donc difficilement comparables. La méthode proposée ici pour calculer le *point équivalent EMA* du maillage a simplement pour ambition de calculer un point réaliste sans signifier pour autant que les réalisations du même phonème cible dans les deux contextes coïncident exactement.

La même approche est adoptée en deux dimensions sur les contours rééchantillonnés au paragraphe précédent. Ces deux *points équivalent EMA* en 2D et en 3D sont visibles à la figure 2.38.

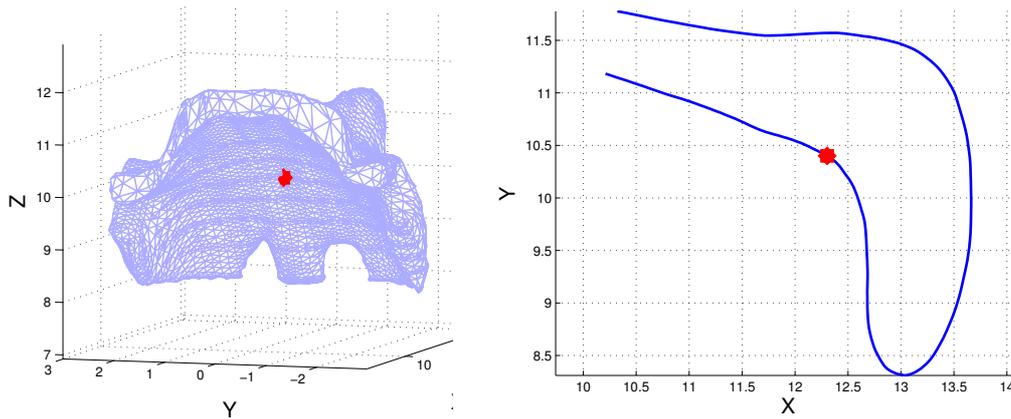


FIG. 2.38 – Maillage générique 3D (gauche) et moyenne des contours médiosagittaux rééchantillonnés (droite) du voile du palais sur lesquels sont positionnés les *points équivalent EMA* calculés en 3D et en 2D.

Ils se situent logiquement environ à mi-distance dans le plan médiosagittal entre la pointe de la luvette et le palais dur, sur la face inférieure du voile du palais.

Les 46 positions prises par ce *point équivalent EMA* en 2D et en 3D sont alors superposées à l'ensemble des positions prises par la bobine EMA collée sur le voile du palais. Les résultats sont visibles à la figure 2.39. On observe un bon recouvrement pour les données 3D mais un recouvrement plus médiocre pour les données 2D. De manière intéressante, on note que l'orientation des axes principaux de l'ensemble des positions EMA et du *point équivalent EMA* en 3D, malgré un léger décalage, coïncident exactement ; ce qui est un peu moins vrai pour le 2D. Ceci signifie que le *point équivalent EMA* obtenu en 3D se déplace de la même manière que le vrai point de chair sur lequel est positionné la bobine EMA du voile du palais. Ce résultat encourageant valide donc, au moins partiellement pour la région centrale du voile du palais qui bouge le plus, l'approche de déformation tridimensionnelle du maillage générique vers les 46 cibles du corpus et nous incite en particulier à assimiler les points du maillage à des points de chair. Le rééchantillonnage effectué sur les contours médiosagittaux semble quant à lui un peu moins convainquant et suggère une déformation du voile du palais dans le plan médiosagittal qui ne serait pas exactement élastiquement uniforme comme nous l'avons proposé. Si les contours ainsi rééchantillonnés restent néanmoins exploitables pour une modélisation mathématiquement tout à fait valide, les mouvements obtenus seront à considérer avec précaution dans leur interprétation et leur comparaison avec la réalité. Il sera en particulier préférable d'exploiter les données tridimensionnelles sur ce point.

2.5.7 Le cas des fosses nasales

Les fosses nasales présentent une particularité qui les rend difficilement classifiables : elles sont bordées de tissus mous... non déformables en parole. Entre structure rigide déformable et structures molles non déformable donc. Elles ont été considérées jusqu'ici comme des structures

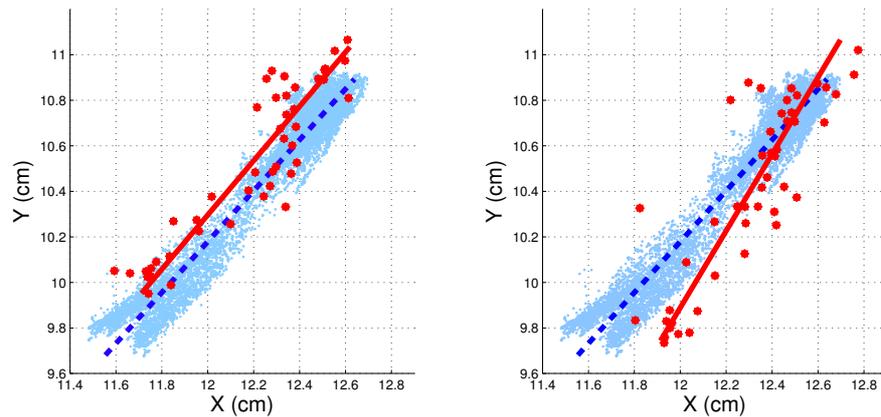


FIG. 2.39 – Superposition sur l'ensemble des positions enregistrées par la bobine EMA du voile du palais des 46 positions du *point équivalent EMA* en 3D (gauche) et en 2D (droite) représentés par les points foncés. Les traits en pointillés représentent l'axe principal des positions EMA obtenu par ACP sur tous les points et les traits continus l'axe principal des positions du *point équivalent EMA* en 3D (gauche) et en 2D (droite) obtenu par ACP sur les 46 points.

rigides non déformables au même titre que les structures osseuses ou les différents sinus. Leur maillage 3D a donc été extrait à ce titre à partir des images CT. Il se peut néanmoins, et nous le vérifierons plus loin dans notre étude, que les muqueuses, de faible densité, ne soient pas visibles sur les images CT. [Dang *et al.*, 1994] ont souligné le volume important occupé par ces muqueuses dans les cavités nasales. La détermination des formes 3D des fosses nasales à partir des images CT conduit donc probablement à l'obtention de cavités de volumes surévalués. Il apparaît donc nécessaire d'extraire la géométrie des fosses nasales à partir des images IRM, sur lesquelles apparaissent les muqueuses. La qualité des recoupes transverses des images IRM sagittales ne répondant pas aux critères nécessaires à la visibilité et à l'extraction de contours fins comme ceux des fosses nasales, le jeu d'images IRM coronales dédié a été utilisé. Néanmoins, étant donnée la géométrie complexe des fosses nasales, la relative *difficulté* à mailler leurs points issus des images CT (voir le paragraphe 2.3.3) et la perte de résolution entre le bloc d'images CT ($0.05 \times 0.05 \times 0.13$ cm / voxels) et le bloc d'images IRM coronales ($0.4 \times 0.1 \times 0.1$ cm / voxels), il semble difficile avec notre outil de maillage d'obtenir une surface maillée complète des fosses nasales issue des images IRM. La démarche adoptée ici consiste donc à déterminer les contours des fosses nasales sur les images IRM coronales ainsi que leurs recoupes dans la région narinaire, de manière à obtenir un échantillonnage précis proche de la réalité mais non interpolable. Les deux types de données, le maillage issu des images CT d'une part et les contours plans issus des images IRM d'autre part, sont finalement complémentaires :

- Le maillage offre une géométrie complète des fosses nasales, qui permet notamment des recoupes par des plans d'orientations quelconques, mais n'est exploitable qu'à caractère informel.
- Les contours offrent une géométrie partielle précise proche de la réalité mais ne sont pas exploitables en tant que surface complète. En d'autres termes, ils n'ont pas la *maniabilité*

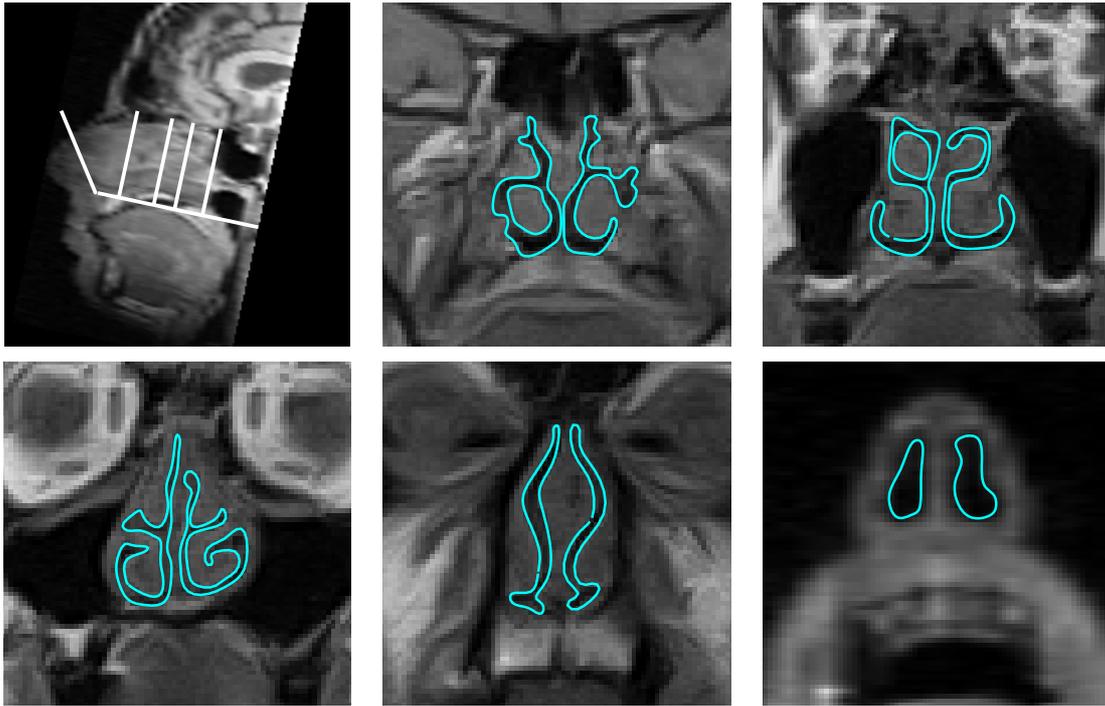


FIG. 2.40 – Tracés manuels des contours des fosses nasales sur 5 images dont les intersections avec l’image médiosagittale sont visibles en haut à gauche ; les images sont ordonnées de gauche à droite et de haut en bas et se déroulent de l’entrée des fosses nasales côté cavum vers la sortie narinaire. Les quatre premières images sont des images IRM coronales originales et la dernière (en bas à droite) est une recoupe d’un bloc d’images IRM sagittales dans la région narinaire détaillée au paragraphe 2.2.5.

qu’offrent les maillages.

En pratique, les contours du conduit nasal sont extraits manuellement sur les plans d’images coronales et leurs recoupes dans la région narinaire (visibles sur la figure 2.9 page 51), selon la même méthode que pour les autres organes. Un exemple de tracés pour cinq images est donné à la figure 2.40.

2.6 Conclusion

Nous avons présenté dans ce chapitre l’ensemble des étapes de l’acquisition des données jusqu’à la constitution d’une base de travail utilisable pour la construction de modèles, en passant par la mise en cohérence des nombreuses données et par l’extraction d’informations sur celles-ci. Un certain nombre de données de natures différentes ont été acquises sur un même sujet : données statiques, dynamiques, images, moulages, etc. Ces données articulatoires ont été positionnées dans un même repère tridimensionnel de manière à être exploitées conjointement. Ainsi les formes tridimensionnelles des structures rigides extraites sur plusieurs jeux de données ont servi de base à l’extraction des structures déformables sur un autre jeu. Une approche originale

de déformation des maillages génériques des structures déformables a permis de construire une collection cohérente de formes 3D du voile du palais et de la paroi du pharynx représentatives des capacités articulatoires du sujet. Leurs points de surface peuvent par ailleurs être assimilés à des points de chair : cette approche a été validée par la comparaison avec un dernier jeu de données dynamiques.

Suivant une approche parfois fastidieuse mais toujours nécessaire pour extraire des informations précises et fiables, les apports de ce travail sont nombreux. Dans cette optique, la mise en place d'une méthodologie permettant la mise en commun et l'exploitation croisée de plusieurs types de données articulatoires de natures très diverses est à souligner. Elle a notamment conduit à la construction d'une base de donnée articulatoire inédite en parole et précieuse pour la construction de modèles basés sur les données. Le chapitre suivant exploite les résultats acquis jusqu'ici pour développer un modèle articulatoire des nasales.

Les nombreuses données utilisées lors de ce travail n'auraient pu être acquises sans l'aide ni l'investissement de nombreuses personnes. Nous souhaiterions remercier Christoph Segebarth, Jean-François Lebas et Monica Baciou pour l'acquisition des images CT et des images IRM sagittales du corpus principal au Département de Radiologie du CHU de Grenoble, et Kiyoshi Honda, Shinobu Masaki, Sayoko Takano, Yasuhiro Shimada et Ichiro Fujimoto pour l'acquisition à ATR (Japon) des images IRM coronales et d'inclinaison de la tête (voir section 3.3). Les données EMA ont été enregistrées à l'ICP à Grenoble grâce à l'aide d'Alain Arnal, Christophe Savariaux et Solange Rossato. Nous remercions également l'équipe GMCAO au TIMC (Grenoble) pour l'utilisation du logiciel de déformation élastique *TestRigid* (Yohan Payan, Maxime Béjar), Daniel Huber pour le logiciel de lissage *SmoothMesh* (laboratoire VMR Lab à Carnegie Mellon University, États-Unis) et Andreas Fabri (équipe de recherche Geometrica à l'INRIA, France) pour le logiciel de maillage 3D.

Chapitre 3

Modélisation articulatoire

Le travail présenté dans le chapitre précédent a permis de construire une base de données de formes tridimensionnelles des divers organes du conduit nasal. La modélisation articulatoire concerne la représentation géométrique des formes variées des surfaces des articulateurs impliqués dans les tâches de parole. Nous nous intéressons dans ce chapitre aux diverses structures constituant le conduit nasal, du bas du port vélopharyngé jusqu'à l'ouverture narinaire. L'approximation du *conduit* nasal en *tuyaux* pour la modélisation acoustique fera l'objet du chapitre suivant.

Deux parties sont distinguées dans le conduit nasal : une partie déformable, le port vélopharyngé, délimité en bas par la connexion avec le conduit oral et en haut par le cavum, impliquant le voile du palais et la paroi du pharynx, et une partie rigide, les cavités nasales, délimitées en amont par le cavum et en aval par l'ouverture narinaire, impliquant les fosses nasales et les narines. Pour des questions de temps, nous n'avons pas considéré les divers sinus dans cette étude.

Dans ce chapitre nous présentons la modélisation de la partie rigide du conduit nasal, la modélisation linéaire tridimensionnelle et bidimensionnelle des parties déformables, la relation entre les formes dans le plan médiosagittal et la géométrie tridimensionnelle et les relations avec les autres organes du conduit vocal.

3.1 Les cavités nasales

Cette section concerne la description des données issues du traitement décrit dans le chapitre précédent concernant la partie rigide du conduit nasal, à savoir les fosses nasales et les narines. Deux types de données sont disponibles :

- Un maillage complet de la surface issu des images CT dont le volume est surévalué par rapport à la réalité, mais qui donne malgré tout un aperçu intéressant.

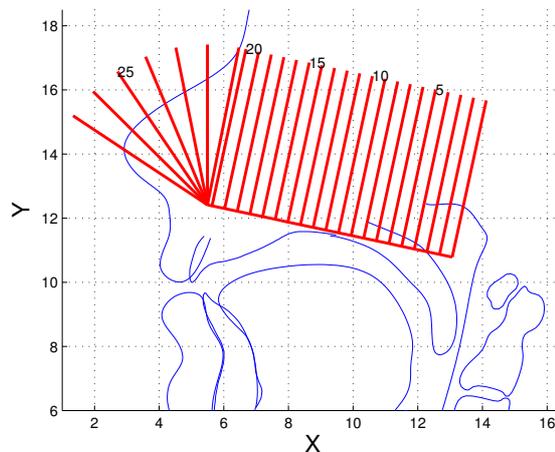


FIG. 3.1 – Grille d'échantillonnage de la partie rigide du conduit nasal.

- Des contours plus réalistes tracés sur les images IRM coronales qui donnent un échantillonnage des fosses.

La différence entre les deux types de données est lié au volume et à l'importance des muqueuses dans les fosses nasales.

Les contours des fosses nasales tracés sur les plans de grilles de la figure 3.1 sont visibles plan par plan sur la figure 3.2 et en 3D sur la figure 3.3. Ces représentations permettent d'apprécier la complexité des deux cavités. La région du cavum, correspondant aux plans de grille situés en amont de l'origine sur la figure 3.2, est bordée dans sa région inférieure par le voile du palais. Sa section dépend donc de la position de ce dernier et sera étudié avec le port vélopharyngé en tant que région déformable. On note néanmoins la géométrie plutôt circulaire des sections. La complexité apparaît dans les choanes, correspondant aux plans de grilles référencés de 0.4 cm à 6.4 cm sur la figure 3.2, et dans une moindre mesure dans les narines, au-delà de 6.4 cm. On peut clairement distinguer les entrelacs des méats entourant les cornets associés (sur les plans de

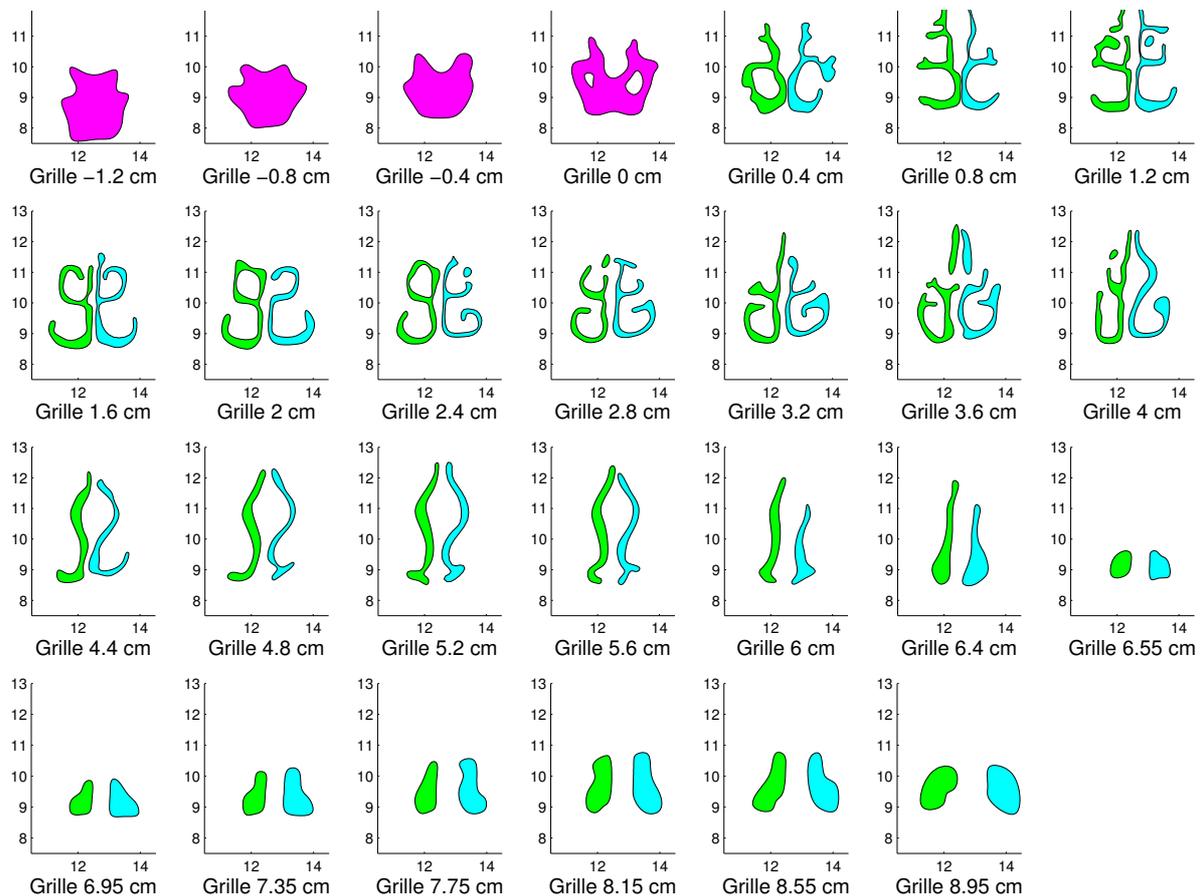


FIG. 3.2 – Contours transverses de la partie rigide du conduit nasal le long des plans de la grille visible à la figure 3.1 tracés à partir des images IRM : cavum (de -1.2 cm à 0 cm), choanes gauche et droite (de 0.4 cm à 6.4 cm) et narines gauche et droite (de 6.55 cm à 8.95 cm). L'origine est arbitrairement fixée au point de séparation du conduit nasal en deux choanes.

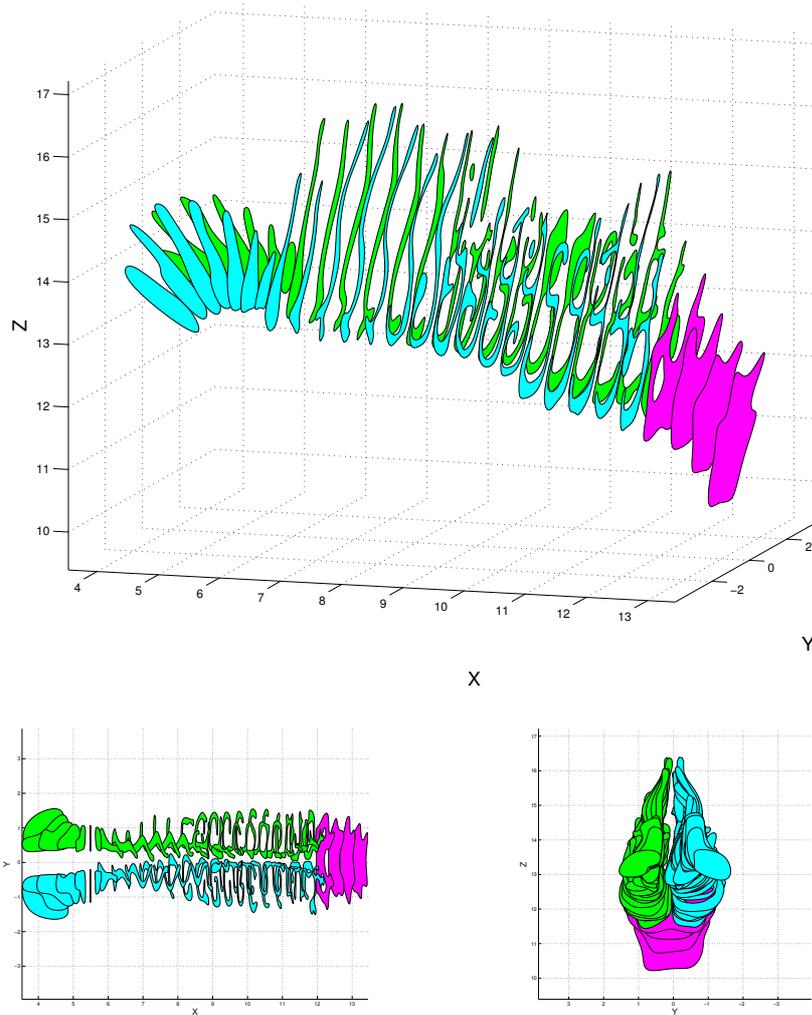


FIG. 3.3 – Contours transverses tracés à partir des images IRM de la partie rigide du conduit nasal le long des plans de la grille visible à la figure 3.1 et affichés en cm 3D dans le repère absolu : vues 3/4 postérieure (haut), supérieure (bas gauche) et antérieure (bas droite).

grilles référencés de 0.4 cm à 3.6 cm), similaires aux contours de [Bjuggren et Fant, 1964] (voir la figure 1.12 page 33), [Matsumura, 1992] ou [Dang *et al.*, 1994]. Pour comparer ces résultats avec ceux issus des images CT et savoir si les muqueuses sont visibles sur ces images, nous avons recoupé le maillage des fosses nasales dans les plans de grille de la figure 3.1 correspondant à l'échantillonnage issu des images IRM. Les résultats comparés en termes d'aires transverses des cavités gauche et droite sont visibles à la figure 3.4. Les résultats en termes de périmètre des contours transverses donnent un graphique similaire.

La première observation est que les aires (comme les périmètres) sont environ une fois et demie plus importantes sur les contours issus des images CT que sur ceux issus des images IRM. [Dang *et al.*, 1994] observent à partir d'images IRM une différence d'aires deux à trois

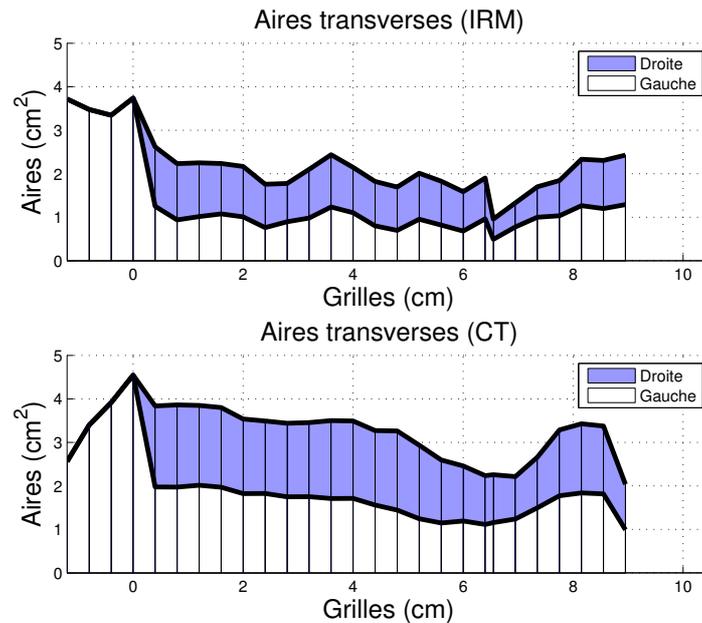


FIG. 3.4 – Aires transverses le long des plans de grille de la partie rigide du conduit nasal extrait des images IRM (haut) et CT (bas). L'axe horizontal représente la distance en cm du plan de grille au point de séparation du conduit nasal en deux choanes, choisi comme origine. L'écart entre les grilles est fixé à 0.4 cm, ce qui constitue l'écart exact entre deux images consécutives dans les régions du cavum et des choanes et une extension de cet écart dans la région narinaire tournante.

fois supérieure pour des enregistrements effectués avec rétraction des muqueuses induite par un astringent par rapport aux conditions normales pour plusieurs sujets. Cette comparaison avec nos résultats suggère qu'une partie des muqueuses serait malgré tout visible sur les images CT. Les ordres de grandeur à la fois des aires et des périmètres issus des images IRM sont quant à eux tout à fait comparables à ceux de [Dang *et al.*, 1994]. Ils sont par contre nettement inférieurs aux aires trouvées par [Bjuggren et Fant, 1964]. Mais, comme l'avancent [Dang *et al.*, 1994], cette différence s'explique par une probable rétraction des muqueuses asséchées du cadavre ayant servi de référence aux mesures.

L'allure générale de nos données est cohérente avec celles proposées à la fois par [Bjuggren et Fant, 1964] et [Dang *et al.*, 1994] : si la région du cavum dépend de la position du voile du palais lors de l'acquisition des données, les aires cumulées des deux fosses varient plus ou moins régulièrement entre 1.5 cm² et 2.5 cm² avec une légère constriction de 1 cm² environ, située à peu près à 7 cm de l'entrée des deux fosses côté cavum et à 2.4 cm de la sortie narinaire. La taille de cette constriction a suscité l'intérêt de plusieurs études : elle varie de 0.23 cm² pour [House et Stevens, 1956] à 2 cm² pour [Bjuggren et Fant, 1964] ; [Feng et Castelli, 1996] l'estiment entre 0.5 cm² et 1 cm² et la fixent à 0.6 cm² pour leurs simulations acoustiques ; enfin, sur la base de données IRM, [Dang *et al.*, 1994] considèrent une constriction d'environ 1 cm² et [Matsumura *et al.*, 1994] entre 0.5 cm² et 1 cm². Par ailleurs, la relative irrégularité des

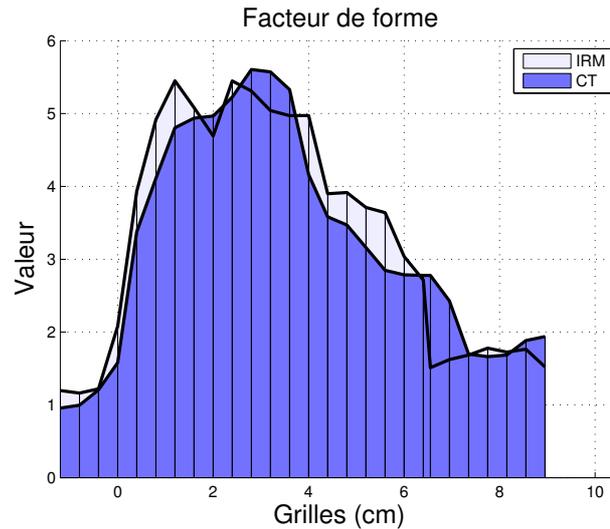


FIG. 3.5 – Facteurs de forme le long des plans de grille de la partie rigide du conduit nasal extrait des images IRM et CT. L’axe horizontal est similaire à celui de la figure 3.4.

aires et des périmètres mesurés le long des plans de grille constatée pour les données IRM par rapport aux données CT pourrait s’expliquer d’une part par la difficulté à détecter le contour exact du conduit sur les images IRM (voir paragraphe 2.5.7) et d’autre part par des variations d’aires proportionnellement plus importantes sur les données issues des images IRM, puisque les interstices sont plus fins.

Le *facteur de forme* FF défini par [Fant, 1960] comme une proportion entre le périmètre P d’un contour et la racine carrée de son aire A est une mesure possible de la complexité d’un contour ; il vaut 1 pour un disque parfait :

$$FF = \frac{P}{\sqrt{4\pi A}}$$

Les facteurs de forme par plan de grille pour les deux types de données sont superposés sur la figure 3.5. Conformément aux attentes, le facteur est très élevé, aux alentours de 5, dans les régions centrales entre 0.4 cm et 4 cm de l’origine (voir la figure 3.2), très légèrement supérieur aux valeurs de 3.5 pour [Bjuggren et Fant, 1964] et autour 4-5 pour [Dang *et al.*, 1994]. Il est proche de 1 dans la région du cavum et de 2 dans celle des narines. Aucune différence marquée n’est à noter entre les deux types de données.

Enfin [Bjuggren et Fant, 1964], [Matsumura, 1992] et [Dang *et al.*, 1994] soulignent une dissymétrie entre les cavités droite et gauche. En calculant le *facteur d’asymétrie* défini par [Dang *et al.*, 1994] sur les valeurs des aires, des périmètres et des facteurs de forme, on mesure une dissymétrie inférieure à 10% sur les périmètres et les formes et inférieure à 20% sur les aires. Ces valeurs sont légèrement inférieures aux observations de [Dang *et al.*, 1994] sur les formes.

Nous avons présenté dans cette section la modélisation articuloire tridimensionnelle, c'est-à-dire l'obtention des formes géométriques 3D, de la partie rigide du conduit nasal, du cavum aux sorties narinaires. L'étape suivante consiste à construire un modèle articuloire tridimensionnel de la région déformable du conduit nasal, en d'autres termes du port vélopharyngé. Cette modélisation, qui fait l'objet de la suite du chapitre, s'appuie sur une décomposition linéaire des mouvements présentée dans le section suivante.

3.2 Principe de modélisation des structures déformables

L'objectif de la modélisation articuloire proposée dans cette étude est de déterminer les degrés de liberté articuloires des organes déformables. Suivant une approche initiée par [Maeda, 1979] et déjà utilisée au laboratoire, notamment par [Beautemps *et al.*, 2001] ou [Badin *et al.*, 2002], les gestes des organes sont décomposés en plusieurs mouvements indépendants associés à ces degrés de liberté et décorrés les uns des autres sur l'ensemble des tâches considérées. La contribution de chaque degré de liberté au mouvement de l'articulateur correspond à une composante articuloire. Un paramètre de contrôle permet de fixer la valeur de la contribution d'un mouvement spécifique lié à un degré de liberté articuloire précis. Ces degrés de liberté sont déterminés principalement par l'observation des corrélations entre les différentes variables qui décrivent les formes et les positions des articulateurs pour ne conserver que les paramètres indépendants. Les corrélations reposent sur une série de contraintes implicites et explicites : continuité physique des articulateurs, propriétés biomécaniques des tissus, nature de la tâche en relation avec son contrôle, etc. Dans notre étude, nous souhaitons que la décomposition des mouvements en composantes indépendantes soit guidée par une nécessité de *vraisemblance* biomécanique : les composantes doivent correspondre à des mouvements plausibles en termes biomécaniques et ne doivent pas représenter la stratégie de contrôle du sujet. Cette nécessité peut ainsi nous amener à introduire quelques connaissances *a priori* bien ciblées pour guider la décomposition en composantes indépendantes basée sur l'analyse des corrélations. En outre, un des enjeux de la modélisation est de réduire le nombre de paramètres de contrôle en gardant un niveau de réalisme le plus élevé possible. La modélisation présentée ici vise donc à obtenir la meilleure décomposition possible biomécaniquement vraisemblable. Cette approche peut conduire à une décomposition mathématiquement sous optimale mais plus réaliste.

La linéarité du modèle découle des techniques linéaires mises en œuvre pour la décomposition des données en composantes indépendantes. Malgré leur relative simplicité, les modèles articuloires linéaires permettent une description fidèle des données ([Maeda, 1979], [Maeda, 1991], [Beautemps *et al.*, 2001], [Badin *et al.*, 2002], [Engwall, 2000]). La forme et la position d'un articulateur pour une configuration précise s'exprime alors comme une combinaison linéaire des composantes du modèle, pondérée par le jeu de valeurs des paramètres associés correspondant à cette configuration. En d'autres termes, en supposant la forme d'un articulateur définie par le jeu des coordonnées de l'ensemble de ses points, la variation de la valeur d'un seul paramètre à la fois conduit à un déplacement de chaque point de l'articulateur sur une ligne droite.

L'Analyse en Composantes Principales (ACP) sur des données articulatoires proposée à l'origine par [Maeda, 1978] et largement utilisée par la suite en modélisation articulatoire (par exemple [Maeda, 1991], [Badin *et al.*, 2002], [Hoole et Nguyen, 1999] ou [Engwall, 2000]) permet d'exploiter la corrélation des données pour obtenir une décomposition en composantes linéairement indépendantes. Ces dernières forment une nouvelle base de représentation des articulateurs. La modélisation articulatoire par ACP consiste à ne conserver pour représenter les articulateurs qu'un petit nombre de vecteurs judicieusement choisis dans cette nouvelle base. La partie suivante du paragraphe présente le principe en s'appuyant sur les explication de [Lebart *et al.*, 1995] et [Ramsay, 2001], que je souhaiterais remercier pour ses explications détaillées. Le détail des calculs est fourni dans l'annexe A.

Considérons n observations d'un ensemble de m variables. Prenons le voile du palais comme exemple. Son maillage étant composé de 5239 points, il est défini par $m = 5239 \times 3 = 15717$ variables. On possède $n = 46$ observations de ce même maillage sur l'ensemble du corpus. Les observations sont donc définies par n vecteurs $\{Y_j, j = 1..n\}$ de taille $m \times 1$. D'après [Lebart *et al.*, 1995], les résultats mathématiques obtenus en considérant m observations de n données sont identiques. On considère donc pour la suite $n \leq m$. La première opération consiste à centrer les données ; on considère pour la suite de l'explication que les données sont donc centrées.

Le principe de la modélisation linéaire est de représenter une réalisation quelconque Y_j des données de manière linéaire comme une somme pondérée de p facteurs :

$$\forall j \in \{1..n\}, Y_j = CX_j + Z_j \quad (3.1)$$

où C est une matrice déterministe rectangulaire de taille $m \times p$ dont les p colonnes sont les vecteurs définissant chaque composante du modèle ; c'est donc elle qui définit le modèle. X_j , de taille $p \times 1$ est un vecteur de pondération des p composantes du modèle ; il change pour chaque réalisation et constitue le vecteur de contrôle (ou d'entrée) du modèle. Enfin Z_j , de taille $m \times 1$ constitue l'erreur de modélisation pour la réalisation j . Un des critères de modélisation étant de réduire au maximum le nombre de facteurs tout en minimisant l'erreur de modélisation, on choisira $p < m$.

L'objectif de la modélisation est de trouver C en minimisant un critère d'erreur, forcément relié à Z , matrice de taille $m \times n$ contenant les erreurs de l'ensemble des observations. On choisit un critère des moindres carrés qui revient à calculer la somme des variances des m composantes individuelles de l'erreur Z . Mathématiquement, elle s'exprime comme suit :

$$\varepsilon = \sum_{i=1}^m \frac{1}{n} \sum_{j=1}^n z_{ij}^2 \quad (3.2)$$

où z_{ij} est l'élément (i, j) de la matrice Z de taille $m \times n$. Cette erreur s'exprime également comme la somme des valeurs diagonales, généralement appelée la *trace*, de la matrice de covariance de Z :

$$\varepsilon = tr(E\{ZZ^t\}) \quad (3.3)$$

où t définit l'opération de transposition. Cette erreur définit donc une erreur globale de l'approximation de l'ensemble des n vecteurs Y par CX .

L'analyse en composantes principales revient à diagonaliser la matrice $E\{YY^t\}$ (voir l'annexe A). Soit $\{\lambda_i, i = 1..m\}$ ses m valeurs propres et $V_i, i = 1..m$ les vecteurs propres associés. Ils sont mutuellement orthogonaux entre eux et la matrice V , de taille $m \times m$, est une matrice de changement de base. Une modélisation linéaire à $p < m$ facteurs consiste à sélectionner p vecteurs propres pour former la matrice C . On montre alors (voir l'annexe A) que cette solution minimise l'erreur globale de modélisation ε qui vaut alors :

$$\varepsilon = \sum_{i=p+1}^m \lambda_i \quad (3.4)$$

L'erreur globale de modélisation correspond donc à la somme des $(m - p)$ valeurs propres de $E\{YY^t\}$ non utilisées par le modèle. Les vecteurs propres sont alors classés dans la matrice V par ordre d'importance de leur valeur propre correspondante et le modèle de la matrice C est toujours construit grâce aux p premiers vecteurs de V . Ainsi, les valeurs propres restantes dont la somme forme l'erreur globale sont les plus petites. Ceci revient à faire ressortir les composantes les plus importantes en termes articulatoires. Ce terme d'erreur peut être exprimé en relatif en *pourcentage* de l'erreur totale possible, soit la somme de toutes les valeurs propres. Or on montre également (voir l'annexe A) que la variance globale des données Y est égale à la somme des valeurs propres $\{\lambda_i, i = 1..m\}$, soit l'erreur totale possible précédemment définie. On peut donc parler par cette erreur relative de *pourcentage de variance des données non expliqué*. Au cours de ce travail, son complémentaire, c'est-à-dire le *pourcentage de variance des données expliqué*, servira à évaluer le modèle.

Au cours de l'étude, le modèle sera évalué quantitativement par deux valeurs :

1. Le pourcentage de variance expliquée défini ci-dessus.
2. L'erreur RMS de reconstruction en cm sur toutes les observations en approchant Y par CX , correspondant à la distance globale en cm entre les points de données et leurs estimés.

En outre, on peut introduire en marge des données brutes utilisées pour la construction du modèle, une connaissance *a priori* du mouvement qui sert à *guider* la construction du modèle, par exemple pour faire émerger des composantes induisant des mouvements biomécaniquement vraisemblables. C'est le cas dans les modèles classiques de langue (voir par exemple [Maeda, 1979], [Badin *et al.*, 2002] ou [Badin et Serrurier, 2006]) pour lesquels la première composante doit représenter la contribution de la mâchoire. La mesure de la hauteur de la mâchoire est alors introduite dans le processus de modélisation et l'objectif est de déduire une composante du modèle relative à ce paramètre.

Par ailleurs, il peut s'avérer judicieux de n'appliquer l'ACP que sur une partie physique restreinte de l'organe pour faire émerger une composante particulière. Ceci revient à appliquer l'ACP à des données Y de taille $q \times n$ avec $q < m$. C'est également le cas pour les modèles de langue (voir

par exemple [Badin *et al.*, 2002]) pour lesquels les diverses composantes sont extraites par ACP *locales* sur l'ensemble des réalisations.

Ces deux approches permettent d'obtenir un paramètre X_0 , de taille $1 \times n$, qui peut être vu comme une connaissance supplémentaire *a priori* des données et qui est alors introduit dans le modèle. On note qu'il couvre bien l'ensemble des observations. On cherche par *régression linéaire* la composante correspondante du modèle sur l'ensemble des données, c'est-à-dire la matrice C de taille $m \times 1$. Elle se réduit à un vecteur que nous noterons U . Le critère d'ajustement pour trouver U est toujours celui de minimisation de l'erreur de modélisation exprimée à l'équation (3.3). On montre qu'il est égal à un facteur près au vecteur $Y X_0^t$ (voir l'annexe A). Le vecteur U représente finalement sur l'ensemble de l'organe la composante liée au paramètre X_0 introduit dans le modèle.

Enfin, de la même manière, il est parfois nécessaire de construire un modèle sur un corpus réduit d'observations $r < n$ présentant des caractéristiques intéressantes. Il devient alors parfois nécessaire de calculer sur le reste du corpus la valeur des paramètres associés aux composantes déjà calculées. C de taille $m \times p$ est connue (il s'agit du modèle), et on cherche par régression linéaire le vecteur de paramètre X de taille $p \times 1$ correspondant à l'observation Y de taille $m \times 1$ non utilisée pour la modélisation. Toujours par régression linéaire, en utilisant le critère de minimisation de l'erreur de modélisation définie à l'équation (3.3), on montre que X est égal à $(Y^t C (C^t C)^{-1})^t$ (voir l'annexe A).

Dans la suite de l'étude, le modèle est constitué de la matrice C . En entrée du modèle, un jeu particulier de valeurs des *paramètres de contrôle* permet d'obtenir en sortie une forme particulière de l'articulateur en fonction du modèle. La variation de la valeur d'un seul paramètre de contrôle permet d'obtenir en sortie un ensemble de formes de l'articulateur qui décrit le mouvement associé au paramètre considéré. Ces formes sont alors appelées des *nomogrammes*.

Nous avons basé la modélisation articulatoire linéaire tridimensionnelle et bidimensionnelle sur l'approche décrite ci-dessus. Comme nous l'avons détaillé dans le chapitre précédent, nous disposons pour cela des données suivantes :

1. 46 observations d'un maillage 3D de voile du palais composé de 5239 points, dont 2812 situés dans une région maximalelement éloignée de 1.5 cm du plan médiosagittal sont utilisés pour la construction du modèle (voir paragraphe 2.5.2).
2. 46 observations d'un contour médiosagittal de voile du palais composé de 202 points.
3. 46 observations d'un maillage 3D de paroi du pharynx composé de 2110 points.
4. 46 observations d'un contour médiosagittal de paroi du pharynx composé de 101 points.

Certaines autres données seront nécessaires au cours de l'analyse. Leurs caractéristiques seront alors détaillées au cours de l'explication suivant les besoins.

Avant de présenter les modèles du port vélopharyngé, la contribution de l'inclinaison de la tête est analysée sur les données décrites ci-dessus.

3.3 Influence de l'inclinaison de la tête

L'alignement des différents blocs d'images, détaillé dans la section 2.4, dans un repère absolu attaché à la structure crânienne du sujet, permet de compenser un éventuel mouvement de tête du sujet. L'ensemble du bloc d'images étant recalé suivant la position de ses structures rigides, nous avons supposé que le mouvement était rigide sur l'ensemble de la tête, c'est-à-dire que le recalage mesuré sur les structures rigides s'appliquait également à l'ensemble de la tête, du cou jusqu'au sommet du crâne, et en particulier au conduit vocal. Cependant, un changement d'inclinaison propre à la tête peut introduire une variation de la forme du conduit vocal, suggérée par [Kitamura *et al.*, 2005], pour une inclinaison d'orientation avant-arrière, dite sagittale. La figure 3.6 montre de la déformation du conduit vocal dans le plan médiosagittal en fonction de l'inclinaison sagittale. On note en particulier la distance entre la langue et la paroi du pharynx juste au-dessus de l'épiglotte qui varie considérablement entre les deux positions antagonistes.



FIG. 3.6 – Images IRM médiosagittales du /i/ pour trois positions d'inclinaison sagittale différentes : en avant (gauche), normal (milieu) et en arrière (droite).

Ce mouvement de tête peut augmenter la variabilité des données sans correspondre à une tâche de parole ; il est donc considéré comme un degré de liberté articulaire global de la tête non spécifiquement lié à la production de la parole. Avant toute modélisation des articulateurs du conduit vocal, il paraît donc nécessaire de retirer des données la variabilité introduite par ce degré de liberté. Notre méthode consiste à calculer la composante correspondant à ce mouvement sur l'ensemble du conduit vocal et à enlever sa contribution aux données de chaque articulateur individuellement.

Anatomiquement, un changement d'inclinaison d'orientation sagittale correspond à une rotation de la structure crânienne, selon des plans sagittaux, autour de la pointe supérieure des vertèbres cervicales, auxquelles elle est reliée par divers ligaments. La position de la structure crânienne étant fixe dans le repère absolu, ce changement d'orientation se traduit dans ce repère par une rotation de la colonne vertébrale autour de la pointe supérieure des vertèbres cervicales. Comme le montre la figure 3.6, ce mouvement introduit une déformation du conduit vocal dans la région pharyngée. La figure 3.7 montre la superposition de contours dans des plans assimilés au plan médiosagittal pour les configurations /i/ et /t^a/. On observe une rotation des vertèbres

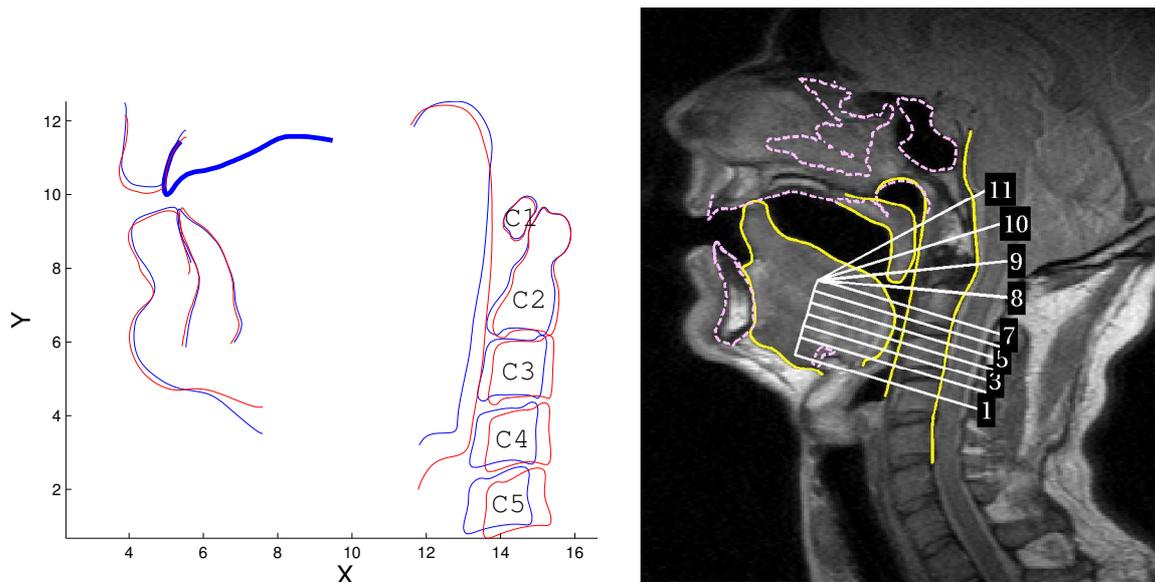


FIG. 3.7 – À gauche, contours superposés pour les configurations $/i/$ et $/t^a/$ des vertèbres cervicales, de la paroi du pharynx, du palais dur, des lèvres et de la mâchoire tracés dans les plans d’image sagittaux les plus proches du plan médiosagittal. À droite, superposition sur l’image sagittale de la configuration $/l^a/$ la plus proche du plan médiosagittal des structures rigides recoupées (traits pointillés) et déformables tracées (traits continus) ainsi que de la grille pour cette section.

cervicales autour de C1 sous l’effet d’un changement d’inclinaison. Une mesure de l’inclinaison de la tête peut donc être obtenue sur les vertèbres cervicales. La figure 3.7 suggère également que la paroi du pharynx peut constituer une mesure similaire de l’inclinaison de la tête. La mesure de l’inclinaison ainsi effectuée, dont le principe est détaillé ci-dessous, pourra alors être considérée comme un *prédicteur* d’inclinaison de la tête et sa composante pourra être calculée par régression linéaire sur chaque organe. Il suffira finalement de retirer des données de chaque organe la contribution apportée par sa composante d’inclinaison.

Les vertèbres cervicales définissent la colonne vertébrale dans la région de la tête et du cou. Étudier sa position dans le plan médiosagittal revient à étudier la ligne que forme le ligament longitudinal postérieur dans ce plan. Il est représenté sur la figure 3.7 droite par la ligne tracée la plus postérieure et peut être considéré comme un bon marqueur de l’inclinaison de la tête. L’inclinaison étant liée à un mouvement involontaire du sujet, les variations de positions enregistrées sont relativement faibles. De manière à orienter d’une part puis à valider d’autre part les analyses effectuées autour de ce problème, un jeu de données supplémentaires dédié à l’étude de l’inclinaison a été enregistré. Il s’agit, pour chaque articulation d’un sous-ensemble de 38 phonèmes sur les 46 du corpus principal, de trois images médiosagittales pour chacune des trois positions d’inclinaison envisageables exagérées : avant, arrière et normal. Les caractéristiques des images sont similaires à celles du tableau 2.2. La figure 3.6 montre trois de ces images pour la configuration $/i/$ dans chacune des positions. Sur chaque image sont tracés la paroi du pharynx

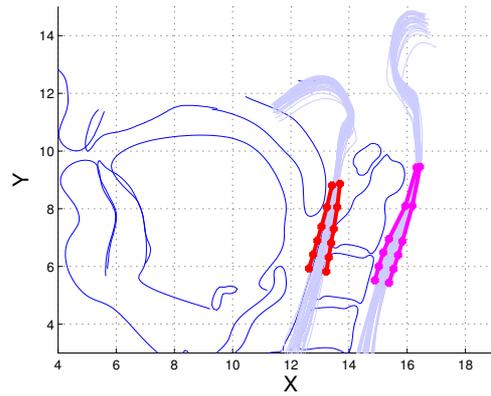


FIG. 3.8 – Contours estimés de la paroi du pharynx et du ligament postérieur (traits épais) pour des valeurs de ± 3 de leur paramètre respectif. Ils sont superposés sur les contours de la tête dans une configuration de repos et sur ceux de tous les tracés médiosagittaux du ligament longitudinal postérieur et de la paroi du pharynx pour l'inclinaison normale seule.

et le ligament longitudinal postérieur (le contour plus postérieur de la figure 3.7, droite) qui représente la position des vertèbres. On montre sur ce corpus de 3×38 articulations qu'une ACP appliquée aux points d'intersection entre les droites de la grille et chacun des deux contours tracés (voir par exemple [Beautemps *et al.*, 2001] pour la méthodologie) conduit à l'émergence d'un même paramètre qui explique 98% du mouvement total. La paroi du pharynx apparaît donc comme un bon marqueur de l'inclinaison de la tête sur le corpus incluant de larges variations d'inclinaison.

La même procédure est ensuite appliquée à la restriction des données correspondant à l'inclinaison neutre seule. L'objectif est de tester l'approche proposée sur un corpus dont les variations d'inclinaison se limitent à des mouvements involontaires du sujet dans des conditions d'enregistrement normales. Une corrélation de coefficient 0.83 reste observée entre les premiers paramètres d'ACP sur chacun des deux contours, qui expliquent 94% de la variance de la paroi du pharynx et 82% de celle du ligament longitudinal postérieur. On vérifie que les contours prédits par ces paramètres restent bien liés à un mouvement d'inclinaison de la tête, comme le montre la figure 3.8. La paroi du pharynx reste donc un bon marqueur de l'inclinaison de la tête pour des petites variations. On vérifie enfin que la mesure de ce paramètre sur la partie de la paroi du pharynx située à hauteur de la pointe de la luette constitue également un bon marqueur de l'inclinaison.

Suite à ces résultats, le *Paramètre d'Inclinaison de la Tête HT* (*Head Tilt* en anglais) est mesuré dans nos données sur les contours médiosagittaux de la paroi du pharynx, obtenus par la recoupe du maillage générique pour les 46 articulation du corpus, à la hauteur de la pointe de la luette. La contribution de ce paramètre à l'explication de la variance totale des données a été calculée par régression linéaire de la partie centrale du voile du palais et de la paroi du

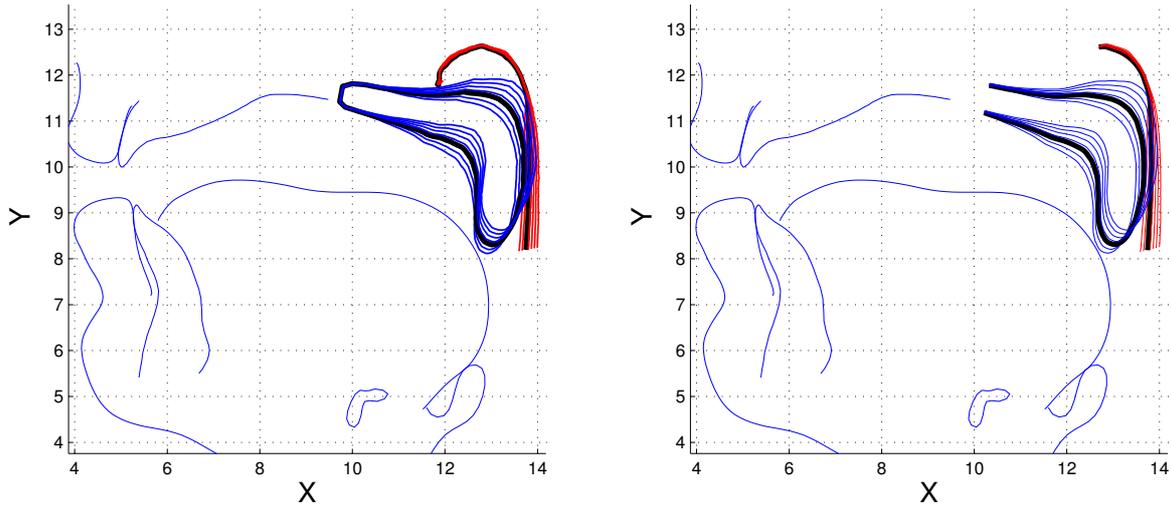


FIG. 3.9 – Nomogrammes du voile du palais et de la paroi du pharynx 3D recoupés dans le plan médiosagittal (gauche) et nomogrammes des contours médiosagittaux (droite) pour des variations de -1.5 à 2.5 du paramètre HT, pour un conduit vocal articulant un /a/. Les traits épais représentent les valeurs moyennes.

pharynx sur ce paramètre. Il ressort que 17% de la variance cumulée des 2812×3 variables du voile du palais et 31% de celle des 2110×3 variables de la paroi du pharynx sont expliquées par le paramètre HT. Le modèle ainsi calculé permet de prédire la variation des formes 3D du voile du palais et de la paroi du pharynx en fonction de l'inclinaison globale de la tête. Les recoupes médiosagittales des nomogrammes 3D du voile du palais et de la paroi du pharynx pour des variations de -1.5 à 2.5 du paramètre d'entrée du modèle d'inclinaison ainsi déterminé sont visibles à la figure 3.9 (gauche). Les valeurs extrêmes -1.5 et 2.5 correspondent aux deux valeurs effectivement extrêmes mesurées sur les 46 observations. Le mouvement de la paroi du pharynx paraît cohérent avec un pur mouvement d'inclinaison de tête comme celui de la figure 3.7 : il décrit un déplacement selon une direction horizontale, maximal dans la région inférieure et fixe dans la région supérieure. Le mouvement du voile du palais correspond à un déplacement simultané le long des axes horizontal et vertical d'une amplitude qui peut paraître disproportionnée au vu des variations de la paroi du pharynx. En outre, le mouvement observé suggère une corrélation avec le mouvement principal de voile du palais mesuré par l'intermédiaire de la bobine EMA (voir la figure 2.10 page 52). En réalité, la forme générale du conduit vocal se trouvant modifiée en lien avec l'inclinaison de la tête, et donc sans rapport avec la tâche de production de parole, il se peut que le sujet introduise une action de compensation pour corriger l'influence de l'inclinaison et atteindre ainsi la cible phonétique ; ceci expliquerait à la fois la forte amplitude et l'orientation des nomogrammes observés. La distance médiosagittale relativement constante entre le voile du palais et la paroi du pharynx (figure 3.9, gauche) suggère d'ailleurs une compensation active par le sujet, logiquement liée au mouvement principal du voile du palais et contrôlé par le sujet.

Enfin, de la même manière, le paramètre HT contribue à l'explication de la variance des contours médiosagittaux du voile du palais à hauteur de 14% et de la paroi du pharynx à hau-

teur de 35%. Les nomogrammes de ces deux contours médiosagittaux pour ce paramètre sont visibles à la figure 3.9 (droite). Le même comportement que celui des maillages 3D recoupés dans le plan médiosagittal est observé.

Les résultats présentés dans cette section ont confirmé la présence d'une variabilité des mouvements des articulateurs du conduit vocal introduite par un mouvement d'inclinaison de la tête. Une mesure de ce mouvement a été déduite des données et sa contribution aux mouvement du voile du palais et de la paroi du pharynx a été évaluée. Dans la suite de l'étude, pour s'affranchir de ce mouvement non lié à la tâche de parole, les modèles proposés sont construits à partir des 46 positions des maillages du voile du palais et de la paroi du pharynx ainsi que leurs correspondants médiosagittaux auxquels ont été soustraits les contributions estimées dues à l'inclinaison de la tête.

3.4 Modèles préliminaires

Deux modèles articulatoires préliminaires du voile du palais ont été développés à l'ICP avant cette étude. Le premier, développé dans le plan médiosagittal, a révélé une première estimation du mouvement du voile du palais. Le second, en 3D, concerne le comportement tridimensionnel du voile du palais et sa relation avec la langue.

3.4.1 Modèle médiosagittal

Le premier modèle articulatoire de voile du palais développé à l'ICP et basé sur les données, en accord avec l'approche présentée dans cette étude, est le modèle médiosagittal de [Rossato, 2000]. Le contour du voile du palais tracé sur 46 images considérées comme médiosagittales a été échantillonné selon une grille mobile attachée à la pointe de la luette comme le montre la figure 3.10 pour les configurations $/n^a/$ et $/t^a/$.

L'échantillonnage d'un contour selon une grille médiosagittale pour appailler ses points

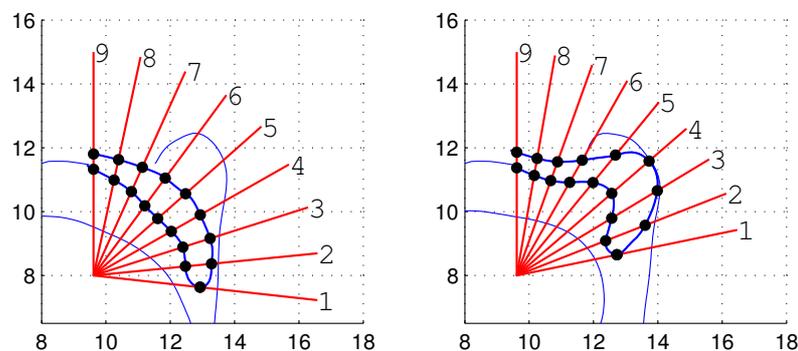


FIG. 3.10 – Intersections des contours médiosagittaux du voile du palais en position $/n^a/$ (gauche) et $/t^a/$ (droite) par les 9 droites de la grille affichée.

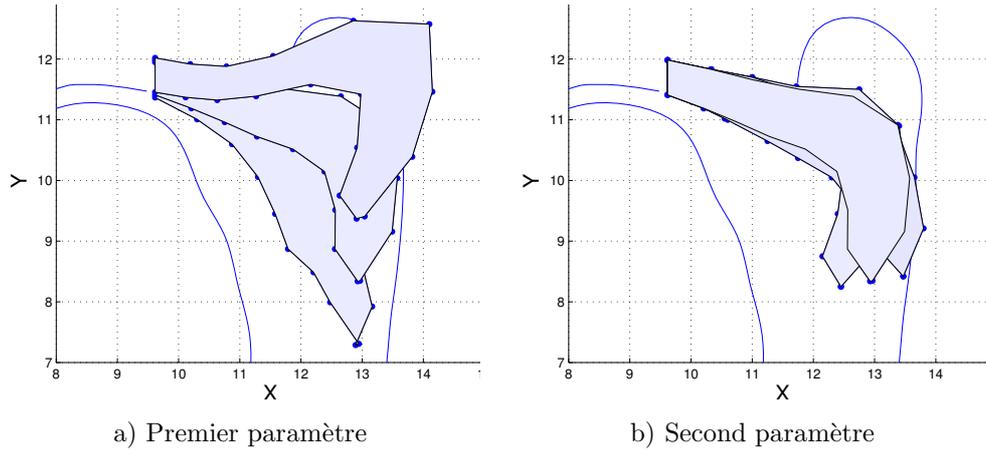


FIG. 3.11 – Nomogrammes des contours médiosagittaux du voile du palais, échantillonnés selon la grille mobile de la figure 3.10, pour des valeurs extrêmes ± 3 des deux premiers paramètres d'ACP du modèle de [Rossato, 2000].

entre plusieurs configurations est un procédé classique en modélisation articulatoire (voir par exemple [Lindblom et Sundberg, 1971], [Maeda, 1979], [Maeda, 1990], [Engwall et Badin, 1999] ou [Beautemps *et al.*, 2001]). L'approche proposée avec une grille mobile permet de *suivre* le mouvement du voile du palais pour appareiller les points des différentes articulations. Une ACP sur ces points révèle qu'une très grande partie de la variance totale des données, 94%, est expliquée par seulement deux paramètres : 88% pour le seul premier paramètre, 8% pour le second. La figure 3.11 montre les nomogrammes du voile du palais associés aux valeurs extrêmes ± 3 de chacun des deux paramètres.

L'analyse des résultats de l'ACP et des nomogrammes de chacun des deux paramètres permet de mettre en lumière deux comportements importants du voile du palais dans le plan médiosagittal :

1. Son mouvement est principalement dû à l'action d'une seule composante qui explique presque 90% de sa variance.
2. Un léger mouvement avant-arrière de luette est observé.

Ces observations nécessitent de plus amples analyses, notamment tridimensionnelles, pour valider ce comportement, mais constituent néanmoins une première estimation très intéressante.

3.4.2 Modèle tridimensionnel

Après un premier modèle médiosagittal détaillé ci-dessus, ce paragraphe présente une première ébauche datée de 2002 / 2003 pour établir un modèle articulatoire tridimensionnel de voile du palais basé sur des données similaires à celles décrites au chapitre 2. Une première estimation des mouvements tridimensionnels du voile du palais est alors observée. Développée dans un cadre plutôt informel pour guider une future modélisation plus complète, l'objet de ce document, sa construction diffère de l'approche présentée jusqu'ici sur un certain nombre de points, dont voici

les principaux :

- Le corpus d'étude est limité à un sous-ensemble équilibré oral-nasal de 22 articulations parmi les 46 du corpus de notre étude.
- Les éventuels décalages de tête du sujet dans des orientations autres que sagittale sont négligés, de même que l'inclinaison.
- Les tracés du voile du palais et de la paroi du pharynx sur les images sagittales latérales sont extrapolés à partir des tracés des images proches du plan médiosagittal pour garder une forme semblables entre les régions latérales et médiosagittale.
- Les maillages génériques sont obtenus par seuillage du bloc d'images CT pour obtenir les volumes du voile du palais et de la paroi du pharynx.

Une ACP est appliquée aux coordonnées 3D de la restriction des points du maillage du voile du palais à ceux compris entre deux plans sagittaux maximale-ment distants de ± 1.5 cm du plan médiosagittal. Comme dans l'analyse médiosagittale précédente et malgré une erreur globalement plus importante, ce tableau montre que 88% de la variance cumulée des points du voile du palais est expliquée par seulement deux paramètres de contrôle articulatoires, le premier expliquant 75% à lui seul ; le second paramètre explique quant à lui 13% de la variance globale. En cohérence avec l'analyse médiosagittale précédente, on observe que l'action du premier paramètre, largement dominant également, correspond à un mouvement du voile du palais simultanément le long des directions verticale et horizontale, qui conforte les observations médiosagittales. Le second paramètre à l'inverse est plutôt lié à un mouvement horizontal qui viendrait compléter la fermeture du port vélopharyngé au lieu de correspondre à un mouvement de l'uvule.

Par ailleurs, ces données ont permis de mettre en lumière une corrélation entre la hauteur du voile du palais et celle de la mâchoire à laquelle il est anatomiquement indirectement relié par l'intermédiaire du muscle palatoglosse.

Ces deux modèles ont mis en lumière des caractéristiques importantes du comportement du voile du palais en parole :

1. Son mouvement est principalement dû à l'action d'un unique paramètre articulatoire expliquant 75% de sa variance en 3D et jusqu'à 88% en 2D.
2. Un second paramètre, expliquant entre 8% de la variance en 2D et 13% en 3D est observé. Son action est différente entre les deux modèles : un mouvement avant-arrière de la l'uvule en 2D, et un mouvement avant-arrière beaucoup plus global en 3D.
3. Une relation entre le voile du palais et la langue a été observée. Les analyses autour d'un tel lien seront discutés plus loin dans le document (voir le paragraphe 3.8).

3.5 Modélisation linéaire tridimensionnelle

Le processus d'extraction des données largement détaillé dans le chapitre 2 a permis de construire un jeu de formes bidimensionnelles et tridimensionnelles du port vélopharyngé utili-

sables pour la construction de modèles articulatoires. Dans cette optique, un premier paramètre HT lié à l'inclinaison de la tête a été calculé et sa contribution soustraite pour limiter la variance de chaque organe à leurs mouvements propres à la parole. Sur la base des données tridimensionnelles ainsi corrigées, nous développons dans cette section un modèle articulatoire linéaire tridimensionnel du port vélopharyngé, c'est-à-dire du voile du palais et de la paroi du pharynx.

3.5.1 Le voile du palais

Pour déterminer les principaux mouvements du voile du palais en trois dimensions, nous avons appliqué une ACP aux 46 observations, corrigées de l'inclinaison de la tête, des 2812×3 coordonnées des points du maillage générique du voile du palais situés entre deux plans sagittaux distants de ± 1.5 cm du plan médiosagittal. Les résultats sont résumés dans le tableau 3.1.

<i>Paramètre</i>	<i>Variance expliquée</i>	<i>Variance expliquée cumulée</i>	<i>Erreur RMS de reconstruction cumulée</i>
1 ^{er} paramètre ACP <i>VL</i>	83 %	83 %	0.08 cm
2 ^e paramètre ACP <i>VS</i>	6 %	89 %	0.06 cm

TAB. 3.1 – Variances expliquées et erreur RMS de reconstruction du voile du palais pour les deux premiers paramètres du modèle 3D de la région du voile du palais comprise entre ± 1.5 cm du plan médiosagittal.

Le premier paramètre d'ACP *VL* explique 83% de la variance cumulée de tous les points du voile du palais distants de moins de 1.5 cm du plan médiosagittal tandis que l'erreur RMS de reconstruction associée est de 0.08 cm. Ce paramètre peut être utilisé pour prédire par régression linéaire l'ensemble des points du voile du palais en incluant les régions latérales non utilisées pour la construction du modèle (voir paragraphe 2.5.2), de manière à obtenir un modèle complet de voile du palais. En d'autres termes, les régions latérales sont prédites à partir de la région centrale. L'effet du paramètre VL sur le voile du palais complet 3D est illustré à la figure 3.12 pour les deux valeurs extrêmes de VL trouvées dans les données. De manière à visualiser l'effet de ce paramètre dans toutes les dimensions et comprendre le mouvement associé dans sa globalité, les figures 3.13 et 3.14 montrent des recoupes dans des plans sagittaux et transverses de nomogrammes correspondants à la variation linéaire de VL entre les deux valeurs extrêmes trouvées dans les données (les nomogrammes de la paroi du pharynx visibles sur l'image correspondent au modèle conjoint de cet organe détaillé au paragraphe 3.5.2).

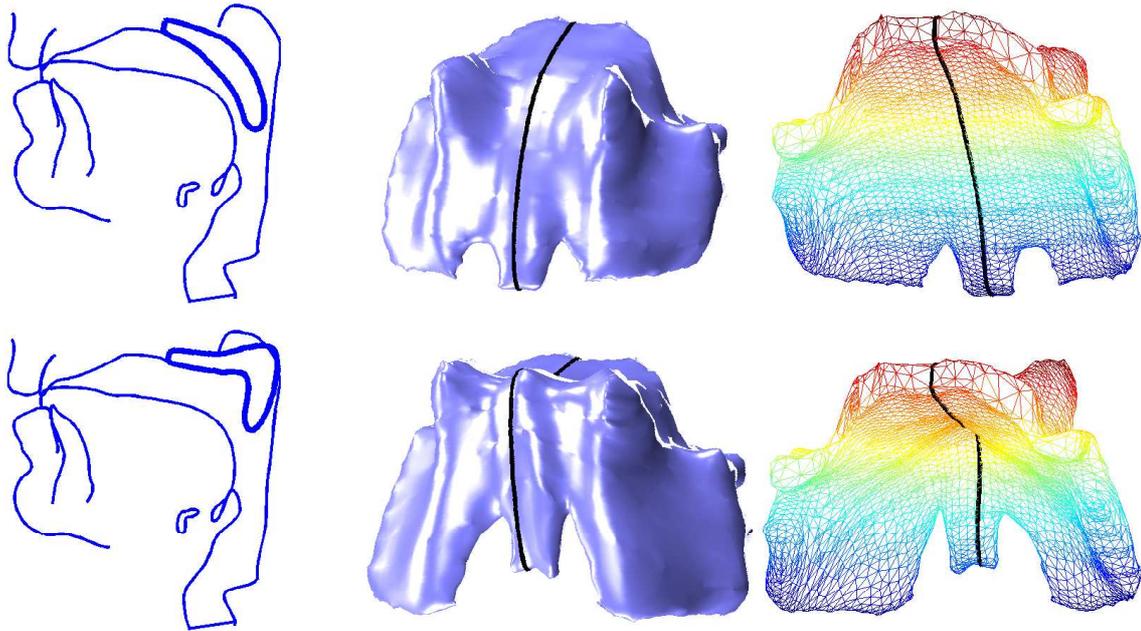


FIG. 3.12 – Maillage tridimensionnel du voile du palais et recoupes médiosagittales pour deux valeurs extrêmes du paramètre VL issues des données (-2 en haut et +1.4 en bas) : vues postérieure (milieu) et antérieure (droite) de la forme 3D ; le trait épais représente la recoupe médiosagittale aussi affichée sur la figure de gauche. Le maillage central est texturé.

Le principal mouvement associé à VL se traduit par une déformation de la région centrale dans une direction oblique à 45° environ similaire à celle observée dans les données EMA (voir la figure 2.39 page 85). Plus précisément, le passage dans le plan médiosagittal (figure 3.13) d'une position basse du voile du palais de forme allongée à une position haute où il est plié à angle droit à mi-distance entre la palais dur et la luette suggère un mécanisme de traction oblique active du voile du palais en son milieu, le reste suivant passivement ce mouvement. La luette se déplace alors verticalement pour suivre la trajectoire oblique du milieu du voile du palais. Ces observations couplées à la description de l'action du muscle *levator veli palatini*, par ailleurs largement admis dans la littérature comme le muscle principal du voile du palais en parole (voir par exemple [Bell-Berti, 1993] pour une revue de la littérature), conduisent naturellement à relier à ce muscle le premier paramètre d'ACP, alors nommé *Velum Levator* (VL). Une illustration schématique de l'action de ce paramètre dans le plan médiosagittal et dans un plan transverse est visible à la figure 3.15. Outre les observations déjà formulées, on peut également noter que les régions latérales sont moins sensibles à l'action de ce paramètre que les régions centrales, ceci pour deux raisons : d'une part elles sont physiquement plus lointaines du lieu d'action du muscle *levator veli palatini*, ce qui atténue son effet, et d'autre part elles sont attachées à d'autres tissus mous formant une certaine résistance à l'action unilatérale d'un muscle.

Deux types d'erreur RMS peuvent être définies pour évaluer la reconstruction du maillage de voile du palais par le modèle :

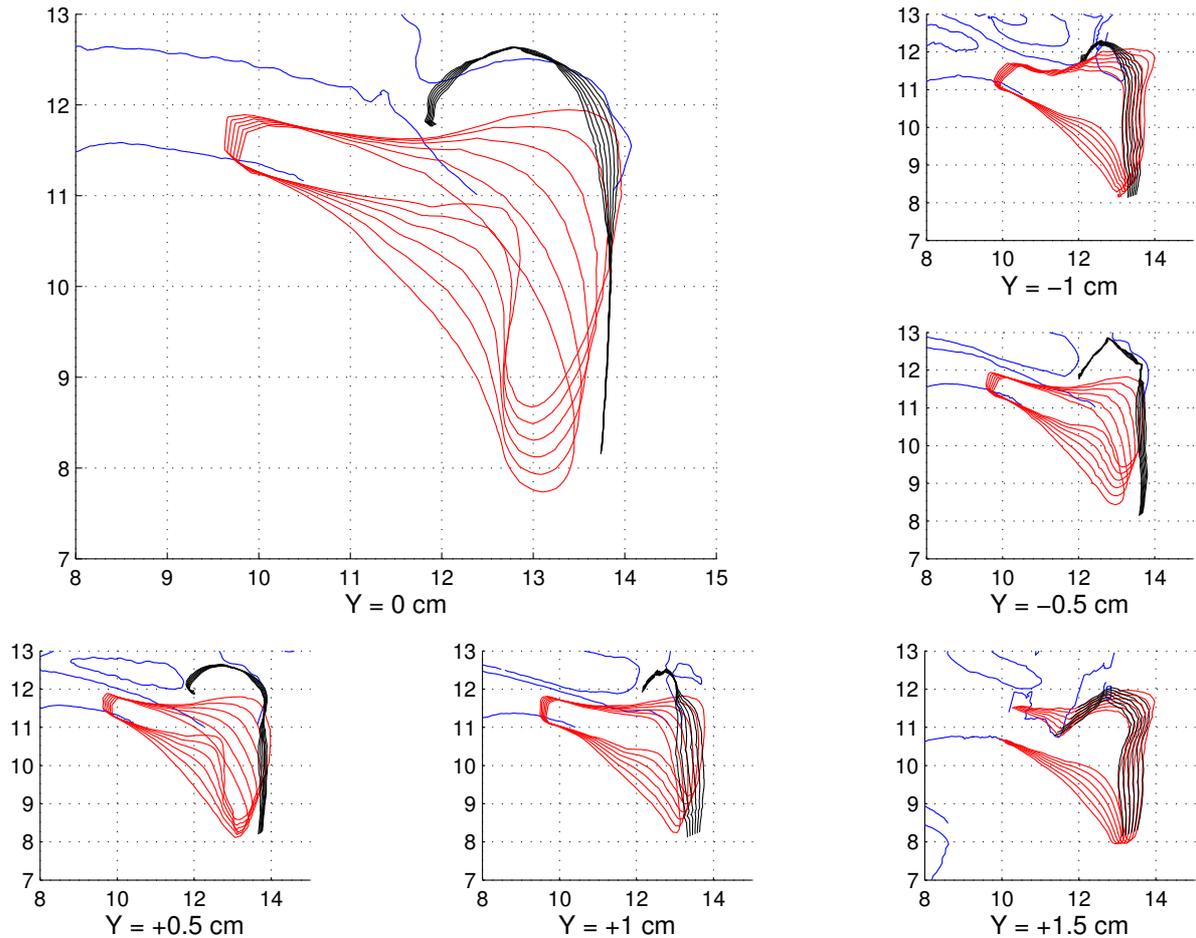


FIG. 3.13 – Nomogrammes du voile du palais et de la paroi du pharynx 3D pour une variation linéaire de VL entre les deux valeurs extrêmes trouvées dans les données : recoupes dans le plan médiosagittal (large figure) et dans des plans sagittaux situés respectivement à -1 cm, -0.5 cm, +0.5 cm, +1 cm et +1.5 cm du plan médiosagittal disposés de haut en bas et de gauche à droite. Les autres courbes sont les recoupes des structures rigides correspondantes dans ces plans.

1. Une erreur RMS *point à point* définie comme la distance RMS entre les points du maillage estimé par le modèle et leurs équivalents dans le maillage des données. C'est l'erreur classique utilisée jusqu'ici. Elle est représentée pour chacune des articulations sur la figure 3.16 par les barres les plus claires et vaut 0.08 cm sur l'ensemble des articulations.
2. Une erreur RMS *point à surface* définie comme la distance RMS entre les points du maillage estimé par le modèle et leurs projetés sur le maillage des données. Elle autorise donc le maillage estimé à *glisser* le long de la forme à obtenir théoriquement sans augmenter l'erreur, contrairement à l'erreur précédente; elle lui est donc forcément inférieure. Sa comparaison avec l'erreur précédente permet par ailleurs de se faire une idée du *glissement* du maillage introduit par le modèle. Elle est représentée par les barres les plus foncées sur la figure 3.16 et vaut 0.04 cm sur l'ensemble des articulations.

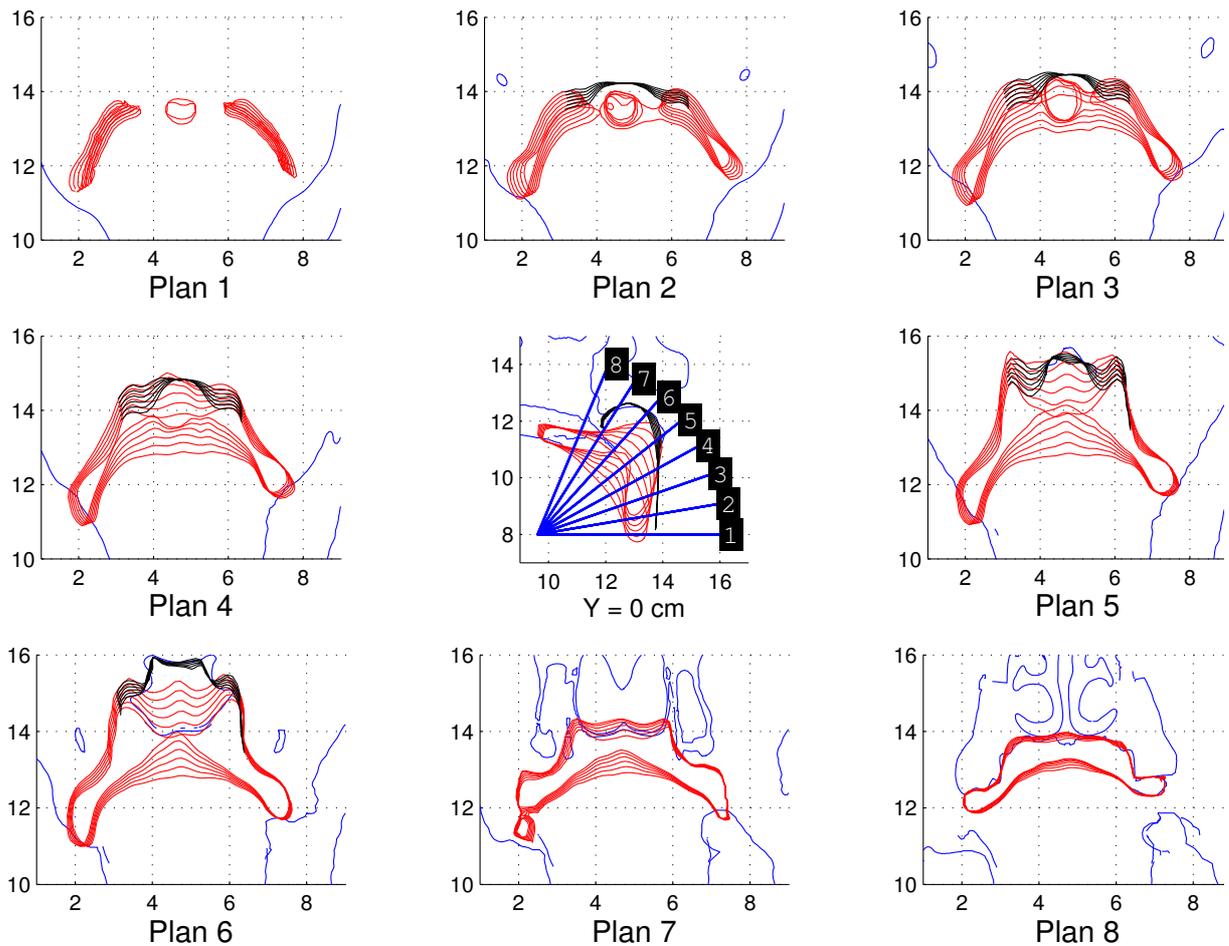


FIG. 3.14 – Nomogrammes du voile du palais et de la paroi du pharynx 3D pour une variation linéaire de VL entre les deux valeurs extrêmes trouvées dans les données : recoupes dans les plans transverses correspondant aux plans de grille visibles sur l'image médiosagittale centrale et disposés par ordre croissant de haut en bas et de gauche à droite. Les autres courbes sont les recoupes des structures rigides correspondantes dans ces plans.

À ces deux calculs d'erreur s'ajoute l'erreur RMS *point à surface* calculée entre les points des contours tracés manuellement sur les images IRM et le maillage 3D du voile du palais estimé par le modèle limité au paramètre VL. Elle permet d'obtenir une évaluation un peu plus poussée de la qualité de la reconstruction du modèle en comparant les résultats obtenus avec les tracés réels, qui constituent pour notre étude la référence absolue. Elle est représentée par les barres de niveau de gris intermédiaire sur la figure 3.16 et vaut 0.055 cm sur l'ensemble des articulations. En comparaison avec l'erreur *point à surface* précédente, logiquement inférieure, elle permet d'estimer l'erreur réelle *du modèle* plutôt que celle *de modélisation*.

Outre les remarques déjà formulées ci-dessus, il est intéressant de constater que toutes les erreurs sont largement inférieures au millimètre et que l'erreur *point à surface* calculée avec les tracés de référence reste inférieure à l'erreur *point à point* définie ci-dessus. On peut donc considérer le modèle assez proche de la réalité, c'est-à-dire des tracés manuels du voile du palais.

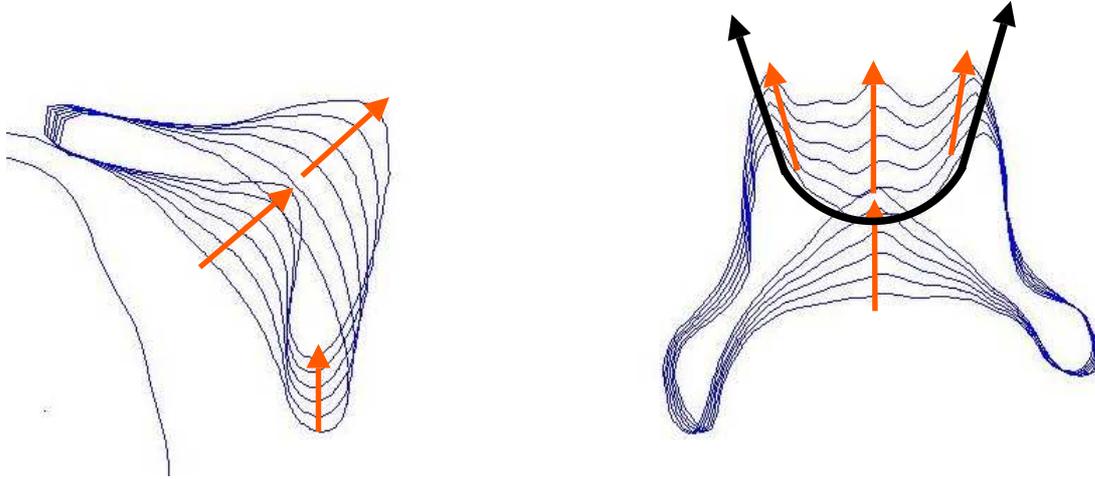


FIG. 3.15 – Interprétation de l'action du paramètre VL sur des nomogrammes du voile du palais 3D recoupé dans le plan médiosagittal (gauche) et sur le plan numéroté 5 de la figure 3.14 (droite). L'arc de cercle épais foncé sur la figure droite représente la traction exercée par le muscle *levator veli palatini* qui entraîne le voile du palais selon les flèches plus claires visibles dans les deux plans de coupe.

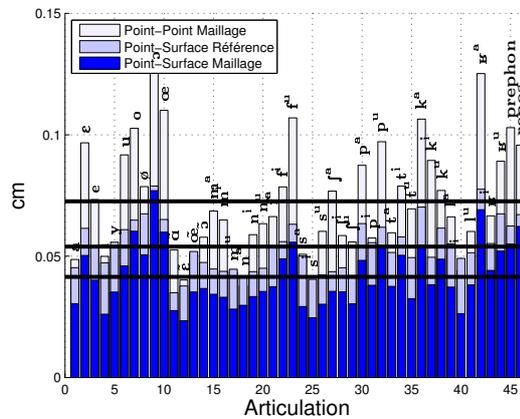


FIG. 3.16 – Erreurs RMS *point à point* et *point à surface* calculées entre les maillages de voile du palais des données et les maillages estimés par le modèle limité au paramètre VL, et erreurs RMS *point à surface* calculées entre les maillages estimés et les tracés de référence, pour chaque articulation et pour les points maximale-ment distants de 1.5 cm du plan médiosagittal. Les droites horizontales représentent les moyennes respectives des trois jeux d'erreurs.

Enfin, l'ultime étape de validation consiste à vérifier visuellement l'adéquation entre : (1) les contours planaires tracés manuellement, c'est-à-dire toujours la référence, (2) les contours recoupés du maillage obtenu par déformation du maillage générique, c'est-à-dire les données utilisées pour construire le modèle, et (3) les contours recoupés du maillage estimé du modèle

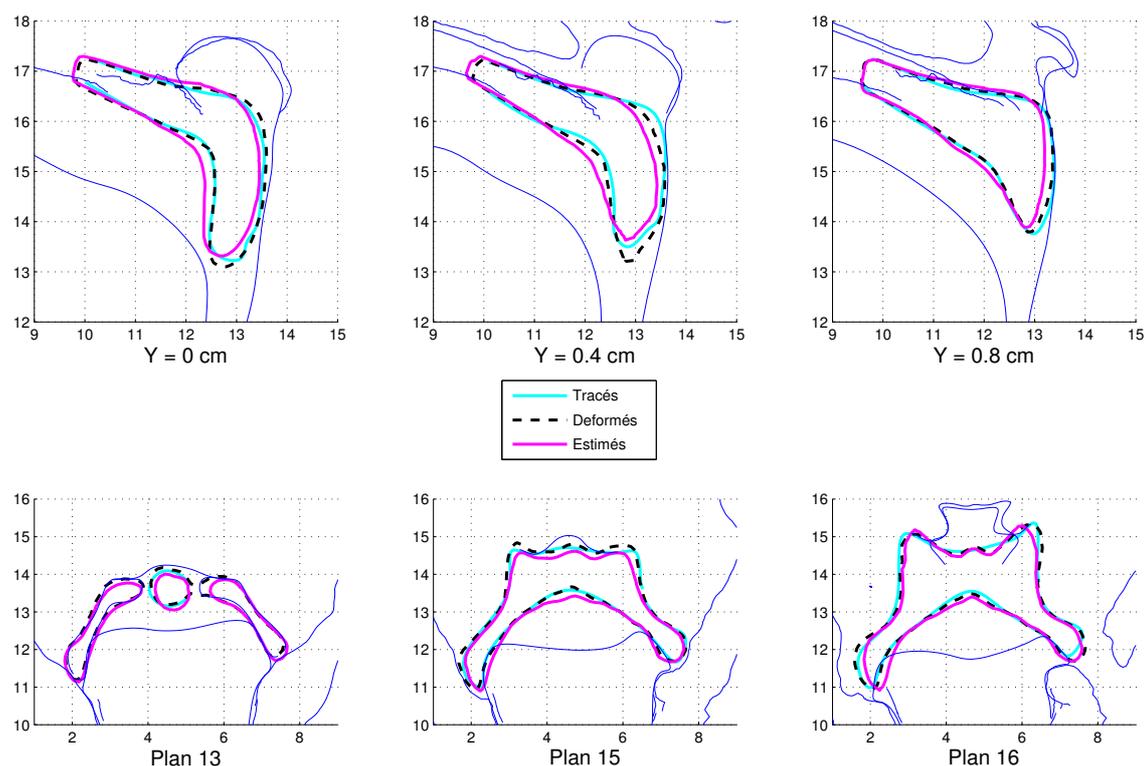


FIG. 3.17 – Superposition dans trois plans d’images sagittales (distants de 0 cm, 0.4 cm et 0.8 cm du plan médiosagittal, haut) et trois plans d’images transverses (bas) des contours tracés manuellement dans ces plans et des recoupes dans ces mêmes plans du maillage obtenu par déformation du maillage des données et du maillage estimé par le modèle limité au paramètre VL pour l’articulation /œ/. Les contours plus fins correspondent aux autres structures rigides et déformables recoupées ou tracées dans ces plans. Les trois plans transverses (bas) correspondent de gauche à droite aux plans numérotés 13, 15 et 16 de la grille de la figure 2.8 page 50.

limité au prédicteur VL. La figure 3.17 montre la superposition des ces contours sur trois plans d’images sagittales et trois d’images transverses pour l’articulation /œ/. On y observe, malgré quelques petites imperfections, notamment dans les plans sagittaux pour cette articulation, une bonne adéquation générale.

Le second paramètre d’ACP **VS** explique 6% de la variance totale des points de la région centrale du voile du palais. La variance cumulée expliquée par ces deux paramètres VL et VS atteint 89% tandis que l’erreur RMS de reconstruction s’abaisse à 0.06 cm. Les maillages estimés du voile du palais pour deux valeurs de VS correspondant aux deux valeurs extrêmes trouvées dans les données sont montrés à la figure 3.18. Les figures 3.19 et 3.20 montrent des recoupes des nomogrammes de maillage dans des plans sagittaux et transverses pour des valeurs de VS linéairement réparties entre les deux valeurs extrêmes trouvées dans les données.

L’action de ce paramètre VS est plutôt reliée à un déplacement horizontal couplé à une

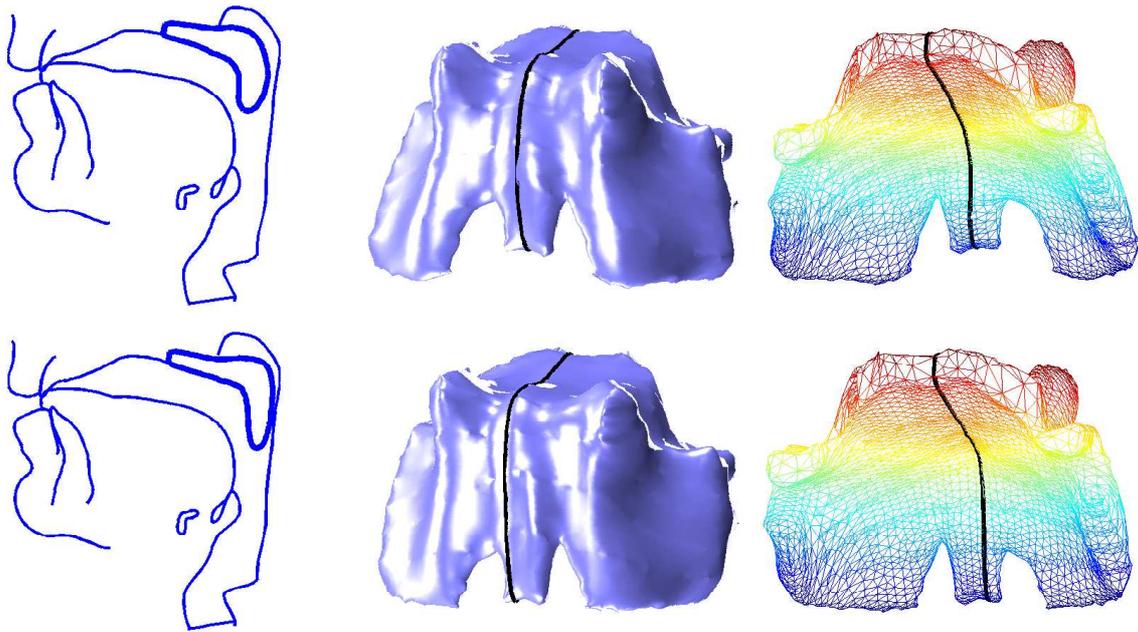


FIG. 3.18 – Maillage tridimensionnel du voile du palais et recoupes médiosagittales pour deux valeurs extrêmes du paramètre VS issues des données (-2 en haut et +2.2 en bas) : vues postérieure (milieu) et antérieure (droite) de la forme 3D ; le trait épais représente la recoupe médiosagittale aussi affichée sur la figure de gauche. Le maillage central est texturé.

élongation verticale du voile du palais dans la région inférieure, complétant la fermeture du port vélopharyngé par un mouvement avant-arrière qui pourrait modifier significativement l'aire de couplage du port vélopharyngé (voir le paragraphe 4.3.2). La figure 3.21 permet de mieux visualiser ce mouvement avant-arrière localement légèrement descendant. L'interprétation biomécanique de ce mouvement est moins évidente que pour VL. L'idée la plus intuitive, mais très improbable, est qu'il pourrait correspondre à un autre muscle du port vélopharyngé parfois impliqué en parole, le *constricteur supérieur* du pharynx (voir notamment [Bell-Berti, 1993]). Ce muscle entourant le port vélopharyngé pourrait éventuellement en se contractant tirer sur le voile du palais en direction horizontale vers le pharynx et conduirait à un effet semblable à celui observé à la figure 3.21. Néanmoins, comme son nom l'indique, les effets de la contraction de ce muscle sont généralement attribués dans la littérature à la paroi du pharynx qui se contracterait alors pour assurer conjointement avec la remontée du voile du palais une fermeture du port vélopharyngé selon un effet sphincter (voir encore [Bell-Berti, 1993]). Si un second effet de ce muscle sur le voile du palais est envisageable, il est à supposer qu'il serait corrélé avec la remontée de celui-ci, contrôlée par VL. Or VL et VS sont décorrélés par construction, ce qui rend cette interprétation difficile. Cette observation n'exclut pas cependant une autre interprétation de l'origine biomécanique de VS comme un second effet des *constricteurs supérieur* ou *médian* du pharynx impliqués dans d'autres tâches de parole que la fermeture du port vélopharyngé (voir à nouveau [Bell-Berti, 1993]). De manière plus générale, l'effet de ce paramètre peut résulter d'une combinaison complexe des divers muscles du voile du palais considérés comme secondaires

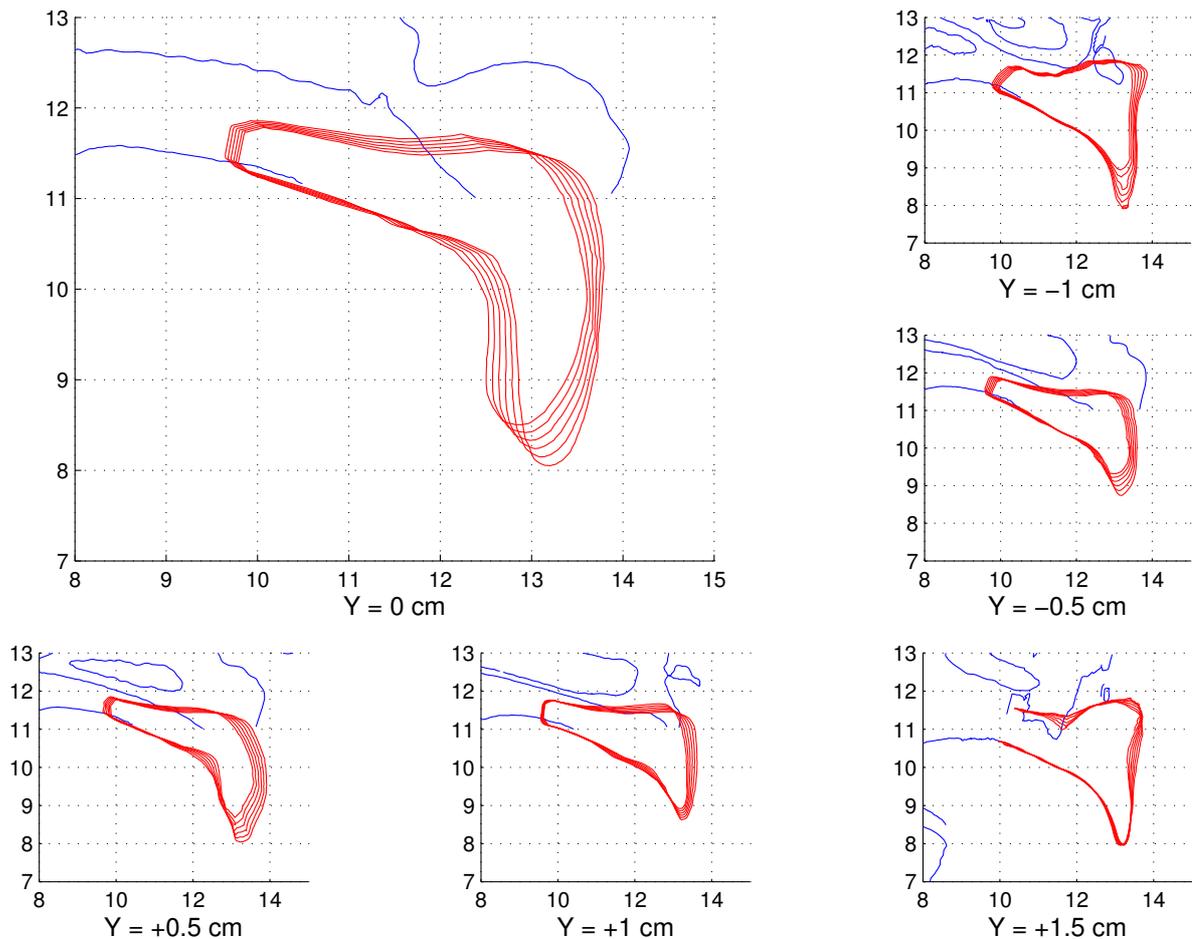


FIG. 3.19 – Nomogrammes du voile du palais 3D pour une variation linéaire de VS entre les deux valeurs extrêmes trouvées dans les données : recoupes dans le plan médiosagittal (large figure) et dans des plans sagittaux situés respectivement à -1 cm, - 0.5 cm, + 0.5 cm, + 1 cm et + 1.5 cm du plan médiosagittal disposés de haut en bas et de gauche à droite. Les autres courbes sont les recoupes des structures rigides correspondantes dans ces plans.

et qu'il est difficile d'identifier. Il se peut par ailleurs que ce mouvement ait une origine autre que musculaire ; les variations de pression dans le conduit vocal en fonction de l'articulation produite peuvent ainsi avoir une influence sur la forme du voile du palais dont ce paramètre peut être une traduction totale ou partielle. Quelle que soit son origine, ce mouvement effectivement constaté dans les données statiques par l'intermédiaire du modèle, et qui sera confirmé par la comparaison avec les données dynamiques (paragraphe 3.5.3), peut être vu comme une amélioration fine de la forme du voile du palais après l'action importante du premier paramètre VL. La question du contrôle de ce mouvement de plus faible amplitude reste néanmoins ouverte à l'issue de notre étude : ce mouvement correspond-il à un contrôle actif du sujet ? Si ce n'est pas le cas, le sujet compense-t-il ce mouvement par l'intermédiaire du paramètre VL pour atteindre l'aire de couplage du port vélopharyngé requise pour le phonème considéré ?

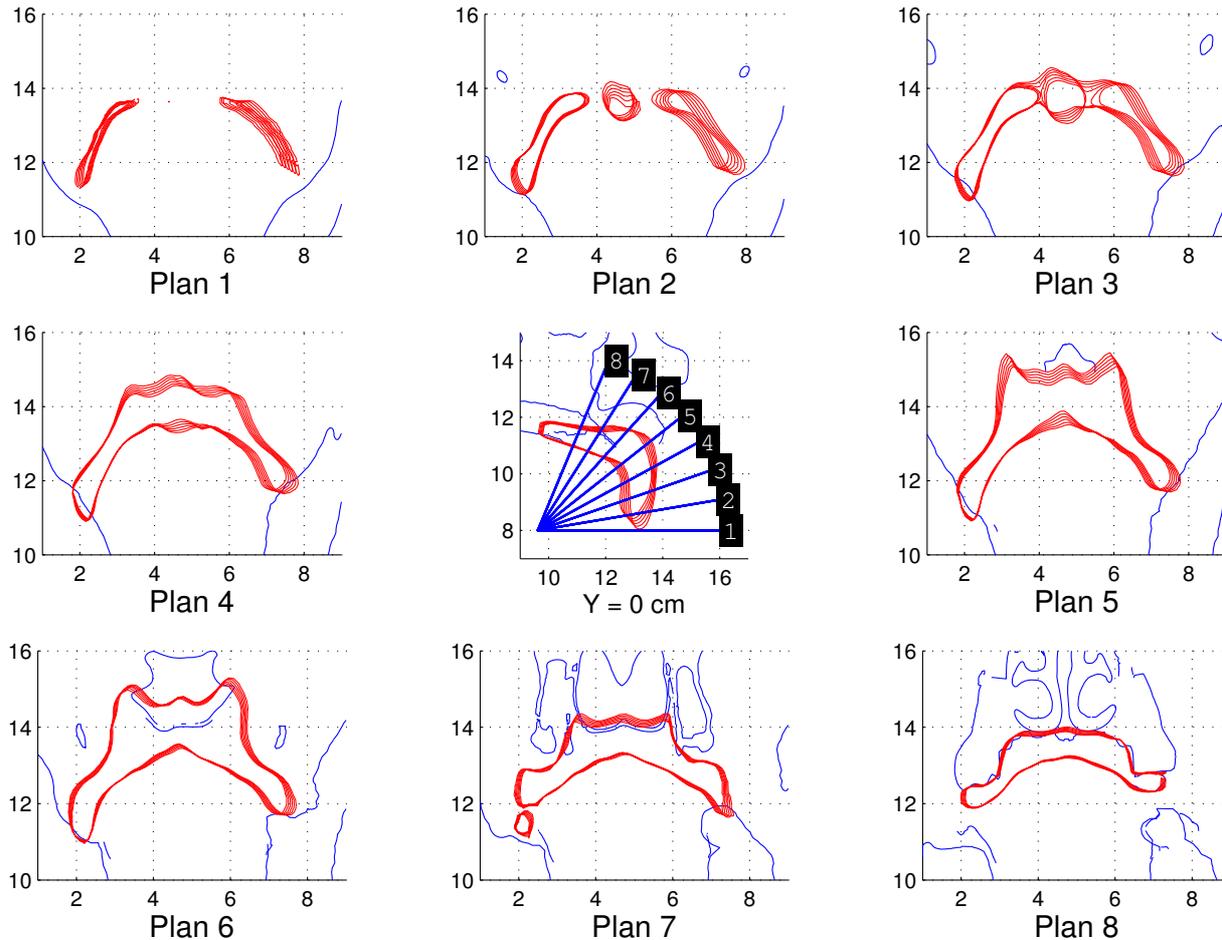


FIG. 3.20 – Nomogrammes du voile du palais 3D pour une variation linéaire de VS entre les deux valeurs extrêmes trouvées dans les données : recoupes dans les plans transverses correspondant aux plans de grille visibles sur l'image médiosagittale centrale et disposés par ordre croissant de haut en bas et de gauche à droite. Les autres courbes sont les recoupes des structures rigides correspondantes dans ces plans.

L'évaluation quantitative du modèle par le calcul des erreurs RMS *point à point* et *point à surface* entre les maillages des données et ceux estimés par le modèle à deux paramètres VL et VS et l'erreur RMS *point à surface* entre les maillages estimés et les tracés de référence est montrée pour les 46 articulations à la figure 3.22. On observe que la baisse la plus importante par rapport au modèle limité au paramètre VL est celle de l'erreur RMS *point à point*, passant de 0.08 cm à 0.06 cm. Les deux autres erreurs baissent de manière négligeable. Néanmoins, on observe que les différents types d'erreur conservent le même ordre que précédemment (comme sur la figure 3.16). Signalons que le faible apport de VS en termes de baisse d'erreur de reconstruction ne remet néanmoins pas en cause son éventuelle contribution significative à une variation de l'aire du port vélopharyngé (voir le paragraphe 4.3.2 à ce sujet).

Pour éclairer le rôle de chacun des deux paramètres, nous avons positionné sur la figure 3.23

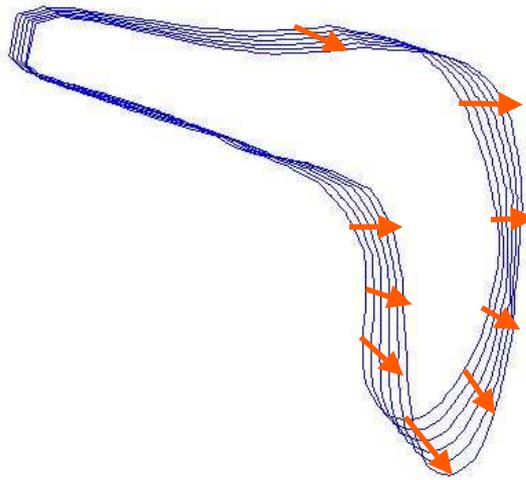


FIG. 3.21 – Description schématique de l'action du paramètre VS sur des nomogrammes du voile du palais 3D recoupé dans le plan médiosagittal. Les flèches représentent le déplacement du voile du palais dans le plan médiosagittal sous l'action de VS.

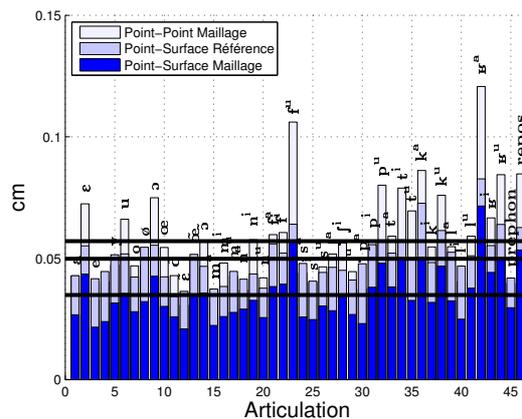


FIG. 3.22 – Erreurs RMS *point à point* et *point à surface* calculées entre les maillages de voile du palais des données et les maillages estimés par le modèle à deux paramètres VL et VS, et erreurs RMS *point à surface* calculées entre les maillages estimés et les tracés de référence, pour chaque articulation et pour les points maximale-ment distants de 1.5 cm du plan médiosagittal. Les droites horizontales représentent les moyennes respectives des trois jeux d'erreurs.

les 46 articulations artificiellement soutenues en fonction de leur valeurs respectives de VL et VS. Le premier paramètre VL apparaît ainsi logiquement directement lié au degré de nasalité puisque les voyelles et consonnes nasales sont concentrées dans une zone inférieure à -1 tandis que les orales le sont dans une zone supérieure à -0.5 ; seules deux articulations intermédiaires /a/ et /repos/ sont situées entre ces deux valeurs. Il est à l'inverse difficile de trouver une

explication phonétique à la répartition des articulations en fonction des valeurs de VS, malgré l'observation de fortes valeurs de VS pour les voyelles orales ; rappelons d'ailleurs que l'amplitude des variations de VS est bien inférieure à celle de VL et que les valeurs de ces paramètres sont normalisées sur la figure.

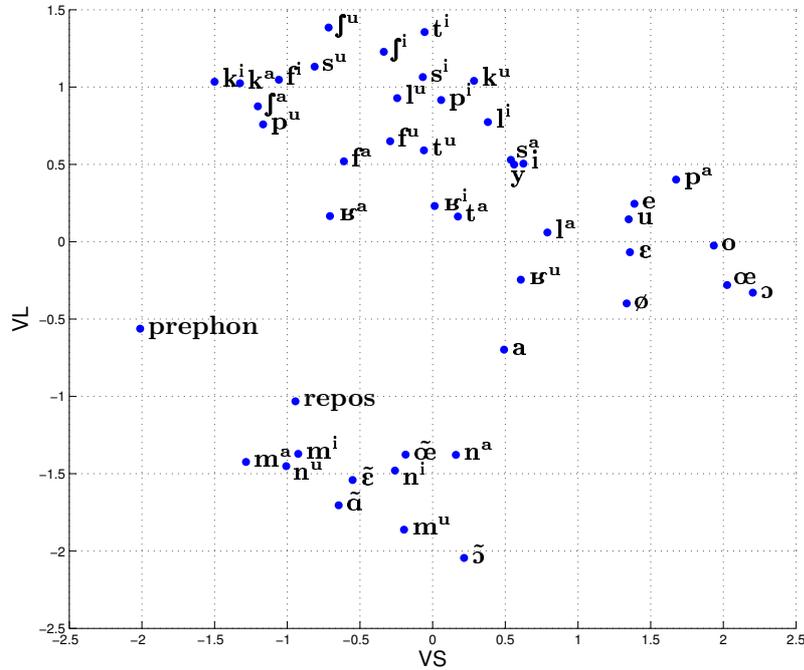


FIG. 3.23 – Répartition des 46 articulations artificiellement soutenues en fonction de leurs valeurs de VL et VS.

Le troisième paramètre d'ACP n'explique que 2.5 % de la variance et semble plus correspondre à du bruit et à une dissymétrie gauche-droite qu'à un réel mouvement du voile du palais. Il n'a donc pas été pris en compte dans la modélisation.

3.5.2 La paroi du pharynx

De manière similaire à l'analyse du voile du palais, nous avons effectué une ACP sur les 46 observations des 2110×3 variables de la paroi du pharynx. Le premier paramètre explique 50 % de la variance et permet une erreur de reconstruction de 0.06 cm. Comme souligné dans le premier chapitre (au paragraphe 1.2.2), il est largement admis dans la littérature que la fermeture du port vélopharyngé est réalisée sous l'action combinée de l'élévation du voile du palais et du resserrement des parois du pharynx fonctionnant comme un sphincter (voir par exemple [Kent, 1997]). On peut donc s'attendre à une forte corrélation entre la paroi du pharynx et le voile du palais. Une régression linéaire des 46 observations des 2110×3 variables du pharynx sur le paramètre VL extrait par ACP sur le voile du palais est effectuée. Il apparaît que VL explique 47 % de la variance totale et que l'erreur de reconstruction avec ce paramètre est de 0.07 cm. Un coefficient de corrélation très fort (0.96) entre VL et le premier paramètre d'ACP de la paroi

du pharynx confirme la forte corrélation avec le voile du palais et justifie un modèle combiné voile du palais/ paroi du pharynx, c'est-à-dire un modèle de port vélopharyngé, contrôlé par le paramètre VL.

Il faut néanmoins noter que ce modèle de paroi du pharynx est biaisé par les choix effectués pour les tracés de cet organe sur les images IRM (voir paragraphe 2.5.2) : dans les régions latérales au-delà de l'ouverture du port vélopharyngé, la paroi est tracée en contact avec le voile du palais pour assurer une fermeture du conduit nasal. Elle est donc construite de manière à épouser les courbes du voile du palais dans les régions latérales. Or une analyse détaillée de la variance des points de la paroi du pharynx révèle que la majeure partie de sa variance, en moyenne quatre fois inférieure à celle du voile du palais tout de même, se situe dans ces régions latérales. Il paraît donc logique de trouver un premier prédicteur d'ACP de la paroi du pharynx (qui explique par définition la majeure partie de la variance, qui correspond aux régions latérales en contact avec le voile du palais) fortement corrélé à celui du voile du palais. Une analyse sur la restriction du maillage de la paroi du pharynx à une région comprise entre ± 0.5 cm du plan médiosagittal, c'est-à-dire sans prendre en compte les régions latérales, n'a rien révélé. La trop

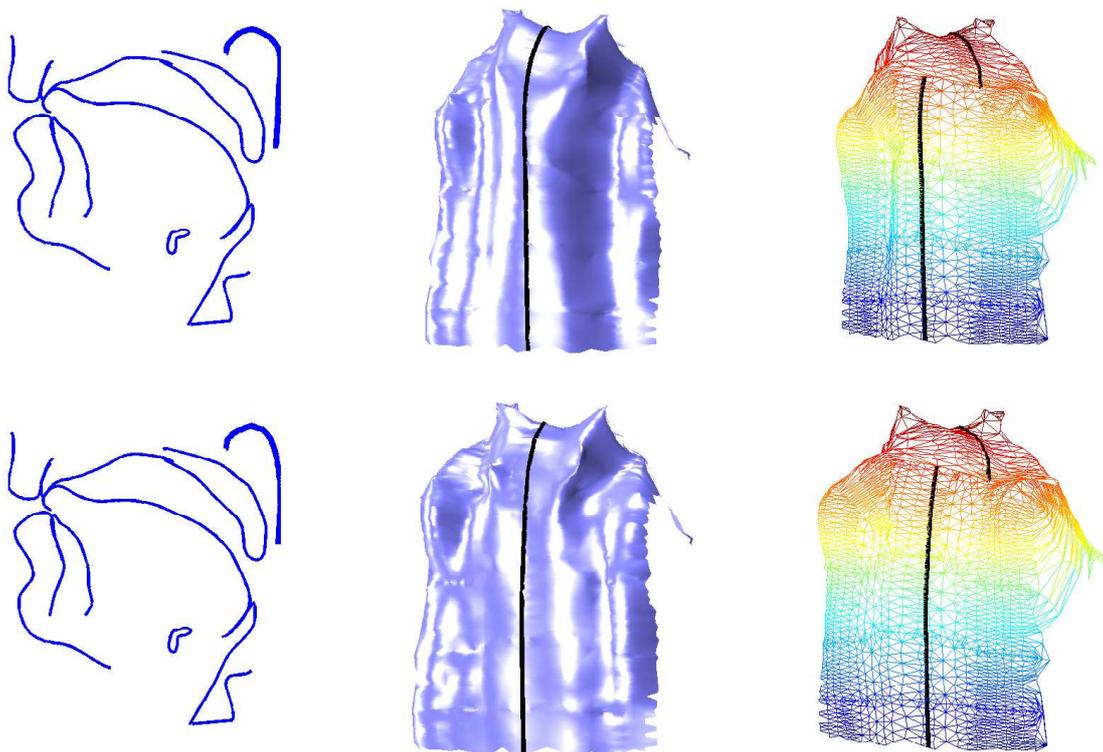


FIG. 3.24 – Nomogrammes du maillage tridimensionnel de la paroi du pharynx et recoupes médiosagittales pour deux valeurs extrêmes du paramètre VL issues des données (-2 en haut et +1.4 en bas) : vues postérieure texturée (milieu) et antérieure non texturée (droite) de la forme 3D ; le trait épais représente la recoupe médiosagittale aussi affichée sur la figure de gauche.

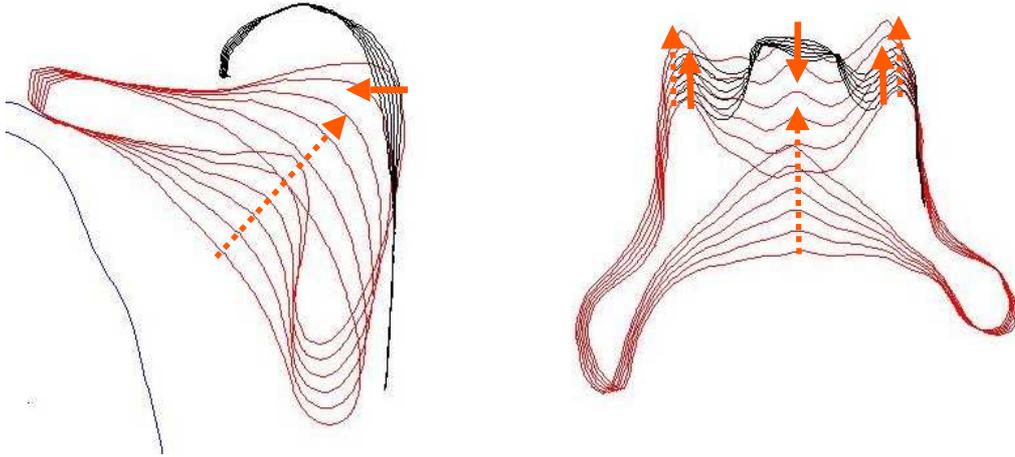


FIG. 3.25 – Description schématique de l’action du paramètre VL sur des nomogrammes de la paroi du pharynx 3D et du voile du palais pour rappel recoupé dans le plan médiosagittal (gauche) et sur le plan numéroté 5 de la figure 3.14 (droite). Les flèches représentent le déplacement conjoint de la paroi du pharynx (flèche continue) et du voile du palais rappelé ici (flèche pointillée) dans chacun des deux plans sous l’action de VL.

faible variance des points, sept fois inférieure en moyenne à celle du voile du palais, avec un écart-type moyen de 0.07 cm, ne permet pas de faire ressortir de composante significative.

L’effet du paramètre VL sur la paroi du pharynx est visible en 3D sur la figure 3.24. Il est également montré en parallèle de son effet sur le voile du palais sur les coupes sagittales à la figure 3.13 et perpendiculaires à la figure 3.14. Conformément aux précédents rappels sur les tracés de la paroi du pharynx qui épousent dans les régions latérales ceux du voile du palais, on y observe un mouvement latéral en cohérence avec celui du voile du palais. De manière plus intéressante, la déformation dans le plan médiosagittal (voir la grande image de la figure 3.13) se produit dans la région supérieure. Elle agit le long du principal axe de déformation du voile du palais, mais de manière antagoniste, c’est-à-dire dans une direction arrière quand le voile du palais baisse et dans une direction avant quand il remonte. Ce mouvement vient ainsi compléter la fermeture du port vélopharyngé. Il correspond à une variation d’épaisseur du *Bourrelet de Passavant* ([Zemlin, 1968]) qui participe à l’effet sphincter global du port vélopharyngé (voir le paragraphe 1.2.2) et qui justifie l’utilisation de VL comme prédicteur de la paroi du pharynx. Ces observations sont schématisées sur la figure 3.25.

L’évaluation quantitative du modèle par le calcul des erreurs RMS *point à point* et *point à surface* entre les maillages des données et ceux estimés par le modèle à deux paramètres VL et VS et l’erreur RMS *point à surface* entre les maillages estimés et les tracés de référence est montrée pour les 46 articulations à la figure 3.26. On note anecdotiquement que seules les articulations /f^h/ et /repos/ sont prédites avec une erreur légèrement supérieure à 1 mm.

3.5.3 Comparaison aux données dynamiques

La comparaison des données dynamiques EMA avec les formes de voile du palais issus des images IRM a permis d'étiqueter un point du maillage générique comme *point équivalent EMA* (voir le paragraphe 2.5.6). Ce point simule sur le maillage du voile du palais 3D le point d'attache de la bobine EMA collée sur le voile du palais réel. Ceci permet de comparer les enregistrements des déplacements de la bobine EMA du voile du palais avec ceux simulés sur le *point équivalent EMA* du maillage de voile du palais. En d'autres termes, la déformation du voile du palais 3D sous l'action de VL ou VS entraîne un déplacement du *point équivalent EMA* (assimilable à un point de chair) comparable aux déplacements enregistrés de la bobine EMA collée sur le voile du palais. Rappelons que l'espace de réalisation de cette dernière s'inscrit globalement dans un parallélogramme visible à la figure 3.27 dont l'axe principal d'orientation oblique mesure approximativement 1.5 cm de long et l'axe horizontal secondaire 4-5 mm.

Étant donné la nature linéaire du modèle de voile du palais, chaque paramètre de contrôle pilote le déplacement de chaque point du maillage selon une ligne droite. La figure 3.27 affiche les lignes décrites par le *point équivalent EMA* sous l'action indépendante de VL et VS dans le plan médiosagittal sur l'espace couvert par la bobine EMA du voile du palais. On observe que l'espace total couvert par le *point équivalent EMA* pour le modèle à deux paramètres VL et VS correspond à l'espace de réalisation du point de chair sur lequel est attaché la bobine EMA dans le plan médiosagittal. Plus précisément, le déplacement du *point équivalent EMA* sous l'action du paramètre principal VL correspond exactement à l'axe de déplacement principal de la bobine EMA. Plus intéressant, le déplacement sous l'action du second paramètre VS correspond lui

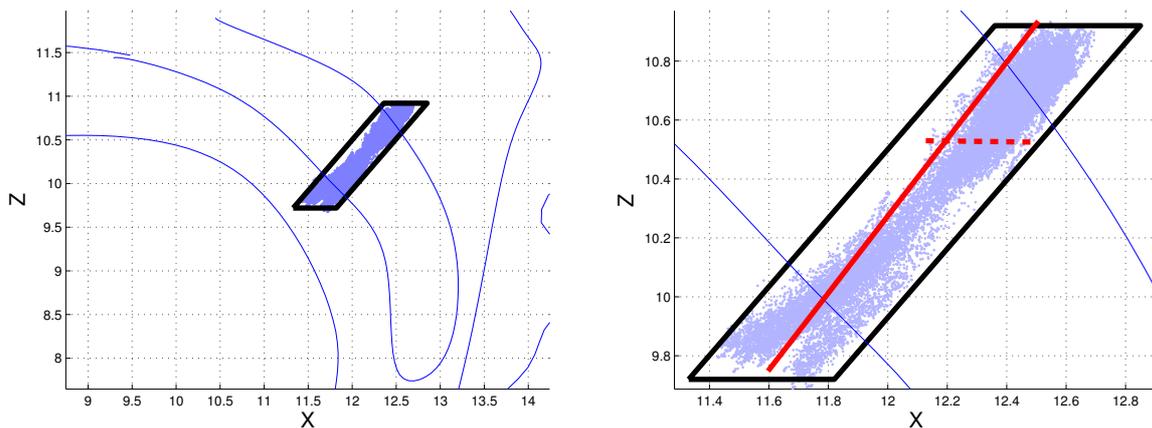


FIG. 3.27 – Espace médiosagittal couvert par la bobine EMA du voile du palais pour le corpus complet des séquences VCV ; un parallélogramme dessiné manuellement encadre cet espace. Les coordonnées X-Z du *point équivalent EMA* du maillage du voile du palais (droite) dans un mouvement associé à VL (ligne continue) et à VS (ligne pointillée) sont représentées pour des valeurs de VL et VS comprises entre les valeurs extrêmes trouvées dans les données.

aussi exactement au second axe de déplacement de la bobine EMA, l'axe horizontal, dans des ordres de grandeur semblables.

Ces observations valident donc les deux paramètres du voile du palais proposés pour le modèle. En particulier, VS, malgré un gain en explication de variance relativement faible, apparaît comme un degré de liberté articulaire complémentaire cohérent du voile du palais :

- il réduit l'erreur RMS de reconstruction de 25%
- il peut modifier sensiblement l'aire de couplage vélopharyngé
- il contribue à la couverture de l'espace médiosagittal des données EMA réelles par le *point équivalent EMA* de manière cohérente.

En conclusion, ce modèle linéaire tridimensionnel a révélé deux degrés de liberté articulaires indépendants et complémentaires pour le port vélopharyngé. Le premier, correspondant au paramètre VL, est associé au mouvement principal de montée-descente du voile du palais selon un axe oblique globalement orienté à 45 ° dans le plan médiosagittal. Il est attribué au muscle *levator veli palatini*. Dans le même temps, il correspond à un mouvement avant-arrière de la région supérieure de la paroi du pharynx autour du plan médiosagittal, dont la proéminence résultante est le *Bourrelet de Passavant*. Il vient compléter le mouvement du voile du palais pour obtenir une meilleure fermeture ou ouverture du port vélopharyngé selon un mécanisme de sphincter. Le second degré de liberté articulaire, correspondant au paramètre VS, est associé à un léger mouvement descendant avant-arrière du voile du palais qui peut modifier sensiblement l'aire de couplage vélopharyngé. Son origine biomécanique, qui pourrait être liée aux muscles constricteurs du pharynx, semble plus incertaine. Ces deux degrés de liberté articulaires permettent au *point équivalent EMA* du maillage de voile du palais de couvrir l'espace médiosagittal des données dynamiques EMA du voile du palais pour l'ensemble des séquences VCV. L'approche présentée dans cette section est reprise dans la section suivante pour construire un modèle médiosagittal tiré des mêmes données.

3.6 Modélisation linéaire médiosagittale

Basée sur une approche similaire à celle adoptée pour la modélisation tridimensionnelle, nous présentons dans cette section un modèle médiosagittal du port vélopharyngé construit à partir des données 2D obtenues des images IRM (voir le paragraphe 2.5.5) et corrigées de l'inclinaison de la tête (voir la section 3.3). Il s'appuie sur les résultats obtenus en modélisation 3D pour développer un modèle similaire. Il permet de tester l'hypothèse de déformation élastique uniforme dans la plan médiosagittal et d'observer le lien entre les formes médiosagittales et tridimensionnelles des organes.

3.6.1 Le voile du palais

De la même manière que pour le modèle 3D, une ACP est effectuée sur les 46 observations des 202×2 variables du voile du palais médiosagittal corrigé de l'inclinaison de la tête. Le schéma de

l'ACP tridimensionnelle se reproduit dans le plan médiosagittal : un paramètre principal, appelé $VL\ 2D$, en référence au paramètre tridimensionnel VL , explique la majeure partie la variance, plus de 86%. Un second paramètre, $VS\ 2D$ encore en référence à son correspondant 3D, explique une petite part, un peu plus de 7%, pour porter l'ensemble de la variance expliquée à 94%. L'erreur RMS globale de reconstruction est de l'ordre du millimètre pour $VL\ 2D$, légèrement plus grand que pour son correspondant 3D, et tombe à 0.06 cm pour les deux paramètres. Les effets de ces derniers sur le voile du palais sont visibles à la figure 3.28.

L'action du premier paramètre $VL\ 2D$ sur le voile du palais est logiquement très similaire à celle de son correspondant en 3D : un mouvement de grande amplitude selon un axe oblique autour de 45° au niveau du milieu du voile du palais entre son attache au palais dur et la pointe de la luette. Cette dernière suit un mouvement vertical quand le milieu suit un mouvement oblique. Ce paramètre est corrélé à son correspondant 3D avec un coefficient de 0.96 ; les paramètres 2D et 3D peuvent alors être considérés comme équivalents. Conformément aux observations du mouvement du modèle 3D, le mouvement observé ici est *a priori* imputable au muscle *levator veli palatini* et assure en grande partie le mécanisme d'ouverture/fermeture du port vélopharyngé. Une description schématique de ce mouvement est proposée à la figure 3.29 (gauche). L'action du second paramètre $VS\ 2D$ est globalement similaire à son correspondant 3D malgré quelques légères différences. D'une part l'amplitude du mouvement est plus grande, à l'image de la plus forte baisse d'erreur RMS de reconstruction qu'il introduit, 0.04 cm. D'autre part, si le mouvement global avant-arrière se retrouve, il est accompagné d'une distorsion plus importante de la forme générale du voile du palais et en particulier de sa face inférieure. La décomposition proposée ici caractérise plus exactement la déformation du voile du palais dans le plan médiosagittal mais l'absence de la dimension latérale peut conduire à des mouvements plus statistiques que biomécaniques. Ainsi la distorsion observée traduit probablement plus le

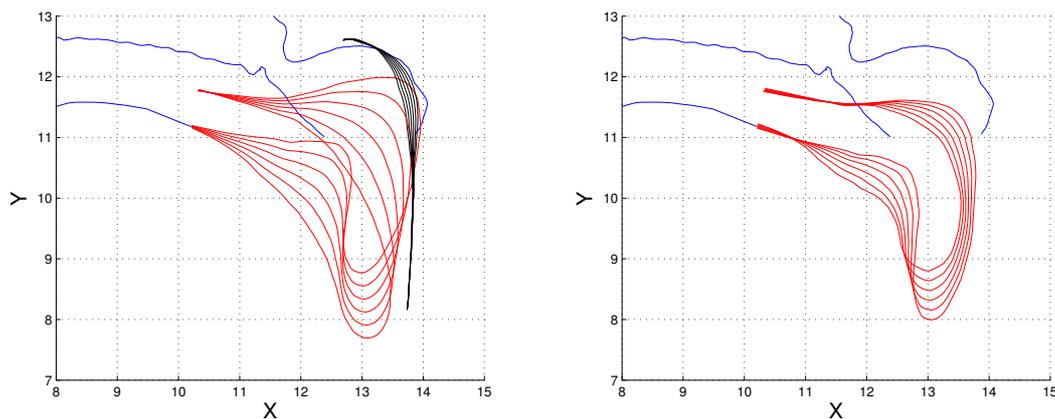


FIG. 3.28 – Nomogrammes médiosagittaux du voile du palais et de la paroi du pharynx pour des variations linéaires de $VL\ 2D$ (gauche) et $VS\ 2D$ (droite) entre les deux valeurs extrêmes trouvées dans les données. Les autres courbes représentent les recoupes des structures rigides dans ce plan.

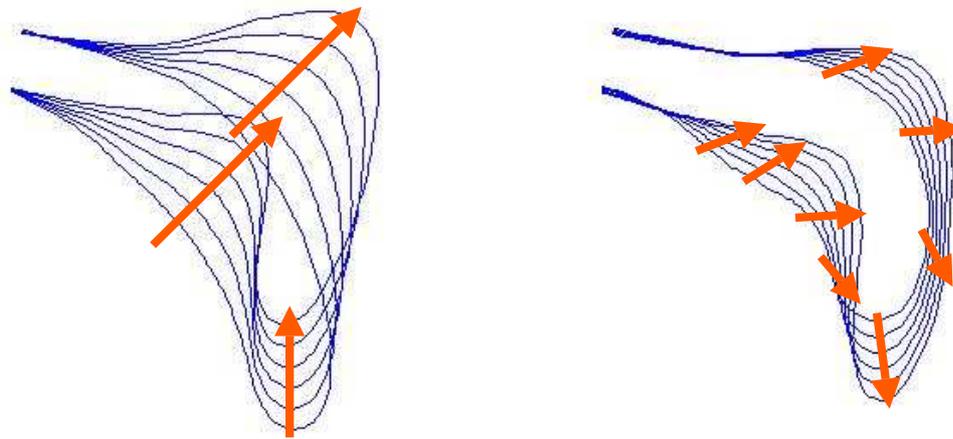


FIG. 3.29 – Description schématique de l’action des paramètres VL 2D (gauche) VS 2D (droite) sur des nomogrammes du voile du palais médiosagittal. Les flèches représentent le déplacement du voile du palais dans le plan médiosagittal sous l’action de ces paramètres.

comportement atypique du voile du palais pour certaines articulations (le $/b^a/$ par exemple) que le mouvement général plus fidèle à la réalité trouvé en 3D par une ACP plus globale. Par ailleurs, il est important de rappeler que l’hypothèse de déformation du voile du palais formulée pour le plan médiosagittal s’était avérée moins fiable que son équivalent 3D dans la région médiane entre le palais dur et la pointe de la luette (voir paragraphe 2.5.6), obligeant toute interprétation biomécanique des mouvements issus du modèle à une certaine prudence. Enfin, notons que les paramètres VS 2D et VS ne sont corrélés qu’à hauteur de 0.65. Une description schématique de la déformation du voile du palais sous l’effet du second paramètre VS 2D décrite ci-dessus est visible à la figure 3.29 (droite).

De manière intéressante, le troisième paramètre d’ACP se traduit par un mouvement avant-arrière de la luette similaire à celui observé dans le modèle médiosagittal de [Rossato, 2000]. N’expliquant que 3% de la variance totale, il n’est pas pris en compte pour le modèle 2D. Néanmoins les mouvements de la luette sous l’action de la langue sont discutés au paragraphe 3.8.2.

3.6.2 La paroi du pharynx

Dans la continuation de l’approche de modélisation conjointe voile du palais/ paroi du pharynx, présentée dans la section précédente, dans le but de construire un modèle de port vélopharyngé basé sur une ouverture-fermeture par effet sphincter, une régression linéaire des 101×2 variables de la paroi du pharynx est effectuée sur le paramètre VL 2D obtenu sur le voile du palais.

Même si l’erreur RMS de reconstruction de 0.06 cm montre une modélisation fidèle, le pourcentage d’explication de la variance, de 14 %, reste faible. Ceci s’explique par une très faible

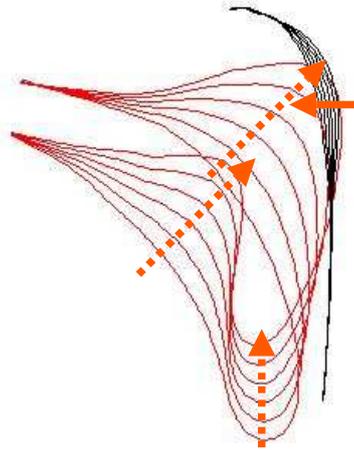


FIG. 3.30 – Description schématique de l'action du paramètre VL sur des nomogrammes de la paroi du pharynx médiosagittale et du voile du palais pour rappel. Les flèches représentent le déplacement conjoint de la paroi du pharynx (flèche continue) et du voile du palais rappelé ici (flèches pointillées) sous l'action de VL.

variance des données, 17 fois inférieure à celle du voile du palais médiosagittal, avec un écart-type moyen de l'ordre 0.065 cm ! Il apparaît alors difficile de séparer du bruit les composantes ayant réellement un sens. Une ACP sur tous les points de la paroi du pharynx ne permet effectivement pas de révéler de paramètre réaliste. Les nomogrammes obtenus par VL 2D sont visibles à la figure 3.28 (gauche) en parallèle de ceux du voile du palais pour le même paramètre. De manière très intéressante, on retrouve tout de même corrélé au mouvement de remontée du voile du palais un mouvement d'avancement horizontal de la région supérieure de la paroi du pharynx, identifié à la section précédente comme le *Bourrelet de Passavant* et contribuant à l'ouverture et à la fermeture du port vélopharyngé selon un fonctionnement de sphincter. La description de ce mouvement est schématisée à la figure 3.30. Au vu des ordres de grandeur de la variance de la paroi du pharynx révélés ici, sa modélisation avec le paramètre VL 2D seul, malgré sa faible participation à la variance globale, paraît un bon compromis : le mouvement associé représente un mouvement réaliste qui vient compléter l'ouverture et la fermeture du port vélopharyngé en grande partie assurées par le voile du palais et l'approche adoptée dans le plan médiosagittal est ainsi similaire à celle proposée en 3D dans la section précédente.

Le modèle articulatoire linéaire médiosagittal du port vélopharyngé proposé ici est similaire à celui proposé en 3D. Un paramètre principal, VL 2D, corrélé à hauteur de 0.96 avec le paramètre VL obtenu en 3D, contrôle le principal mouvement de déformation du voile du palais : un mouvement selon un axe oblique orienté environ à 45° actif au niveau médian du voile du palais entre son attache au palais dur et la pointe de la luvette, cette dernière se déplaçant alors verticalement ; ce mouvement assure conjointement avec un mouvement avant-arrière horizontal de la région supérieure de la paroi du pharynx l'ouverture et la fermeture du port vélopharyngé. Un second paramètre, VS 2D, corrélé avec VS à hauteur de 0.65 seulement, contrôle un mouve-

ment avant-arrière du voile du palais, auquel s'ajoutent quelques distorsions, et qui peut modifier significativement la distance sagittale de couplage du port vélopharyngé. Il est intéressant de souligner le lien fort entre les formes médiosagittales et tridimensionnelles des organes du port vélopharyngé.

La première manipulation de ce modèle consiste comme pour le modèle tridimensionnel à comparer ses performances aux données dynamiques EMA.

3.6.3 Comparaison aux données dynamiques

De la même manière que pour le modèle tridimensionnel, les droites de déplacement du *point équivalent EMA* du contour médiosagittal du voile du palais sous l'action de VL et VS 2D pour observer l'espace de réalisation de ce point sont affichées à la figure 3.31.

On observe de prime abord un décalage non négligeable entre les espaces couverts par la vraie bobine EMA et par le *point équivalent EMA*, contrairement au schéma équivalent en 3D (voir la figure 3.27 page 120). Le décalage de la droite de déplacement du *point équivalent EMA* sous l'action de VL 2D (droite continue) avec l'axe principal des enregistrements EMA souligne la relative inadaptation à la réalité de l'hypothèse de déformation du voile du palais dans le plan médiosagittal, inadaptation déjà observée au paragraphe 2.5.6 lors de l'étude des contours bruts rééchantillonnés. Le décalage observé entre la droite de déplacement sous l'action de VS 2D (droite pointillée) et l'orientation horizontale attendue trouve a priori sa double origine dans l'inadaptation de l'hypothèse de déformation rappelée ci-dessus et les distorsions propres du contour du voile du palais médiosagittal soulignées lors de la description du modèle 2D. Ces

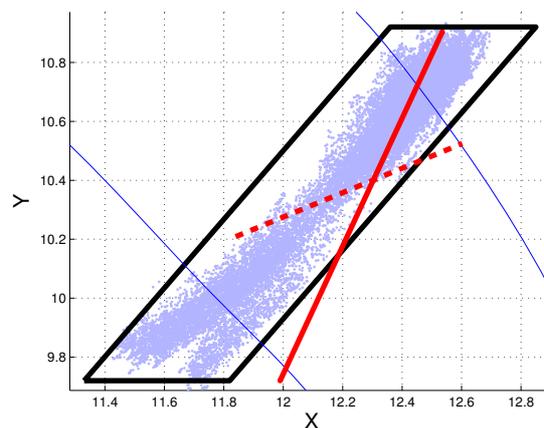


FIG. 3.31 – Superposition des coordonnées du *point équivalent EMA* du contour médiosagittal du voile du palais dans un mouvement associé à VL (ligne continue) et à VS (ligne pointillée) sur l'espace médiosagittal couvert par la bobine EMA du voile du palais pour le corpus complet des séquences VCV ; un parallélogramme dessiné manuellement encadre cet espace. Les valeurs de VL et VS sont comprises entre les valeurs extrêmes trouvées dans les données.

observations un peu décevantes de la modélisation médiosagittale par rapport à celle tridimensionnelle ne remettent malgré tout pas en cause la décomposition statistique du modèle, tout à fait valide, mais incitent plutôt à la prudence lors de l'interprétation des mouvements associés en termes biomécaniques. Elles confirment en particulier l'inadéquation déjà observée au paragraphe 2.5.6 entre l'espace couvert par la vraie bobine EMA dans le plan médiosagittal et celui couvert par le *point équivalent EMA*, réel ou estimé par le modèle.

Le modèle médiosagittal présenté dans cette section apporte deux éléments importants à la compréhension de la modélisation articuloire du port vélopharyngé :

- Le choix de l'hypothèse de déformation des contours médiosagittaux et des maillages tridimensionnels est importante pour l'interprétation biomécanique des mouvements issus du modèle.
- Un lien fort entre les contours médiosagittaux et les formes 3D existe, suggérant la possibilité de prédire les formes 3D à partir des contours médiosagittaux.

Après avoir présenté les modèles articuloires linéaires tridimensionnel et médiosagittal du port vélopharyngé, nous nous intéressons dans la section suivante aux liens entre 2D et 3D en terme d'*inversion articuloire* : comment retrouver la forme globale tridimensionnelle des organes à partir de mesures dans le plan médiosagittal ?

3.7 Du 2D vers le 3D

La modélisation tridimensionnelle permet la reconstruction de la forme complète du port vélopharyngé à partir d'un nombre limité de paramètres de contrôle, en l'occurrence deux, VL et VS. Cette section vise à évaluer la fiabilité de la reconstruction 3D du port vélopharyngé, et en particulier du voile du palais, obtenue pour des valeurs de VL et VS estimées par des *mesures articuloires* du conduit vocal. L'objectif sous-jacent à cette évaluation est de pouvoir obtenir une estimation robuste du port vélopharyngé en 3D grâce à des enregistrements simples dans le conduit vocal. La démarche consiste dans un premier temps à évaluer les valeurs de VL et VS et dans un deuxième temps l'erreur introduite dans la reconstruction tridimensionnelle. La longue tradition de l'étude du conduit vocal dans le plan médiosagittal et la relative simplicité des mesures associées incitent à estimer VL et VS à partir de données médiosagittales. Ce choix s'appuie sur le lien fort constaté entre les modèles 2D et 3D présentés dans les sections précédente, à l'image des modèles de langue, de lèvres et de visage proposés par [Badin *et al.*, 2002]. En partant de cette observation, nous présentons dans cette section l'estimation des formes 3D à partir de deux types de *mesures articuloires* : les enregistrements EMA d'une part et les contours médiosagittaux du voile du palais d'autre part.

3.7.1 Des enregistrements EMA vers la forme 3D

L'articulographe électromagnétique permet d'enregistrer la position de points de chair du conduit vocal au cours du temps à des fréquences élevées. La connaissance du passage de la

position des points de chair aux formes 3D associées des organes concernés permettrait une estimation 3D du conduit vocal au cours du temps. Les analyses présentées dans ce paragraphe s'attachent donc à obtenir la forme tridimensionnelle du port vélopharyngé à partir de la position enregistrée par une bobine collée sur le voile du palais, comme dans nos données (voir paragraphe 2.2.6).

L'analyse du maillage tridimensionnel du voile du palais a permis d'identifier un point du maillage, le *point équivalent EMA*, comme le point de maillage qui correspondrait dans la réalité au point de chair sur lequel est attaché la bobine EMA (voir paragraphe 2.5.6). Les 46 positions prises par ce point de maillage dans les données couvrent approximativement l'espace global pris par toutes les positions de la bobine réelle dans le plan médiosagittal (figure 2.39, gauche, page 85). Par ailleurs, le déplacement de ce point selon chacun des deux paramètres du modèle 3D de voile du palais permet également de couvrir l'espace de réalisation de la bobine réelle dans le plan médiosagittal (figure 3.27, droite, page 120). L'estimation des valeurs des deux paramètres VL et VS à partir de la position dans le plan X-Z du *point équivalent EMA* est présentée dans la suite du paragraphe.

Pour deux valeurs VL et VS données, le modèle linéaire prédit les coordonnées (X_{eqEMA}, Z_{eqEMA}) du *point équivalent EMA* comme suit :

$$\begin{pmatrix} X_{eqEMA} \\ Z_{eqEMA} \end{pmatrix} = C \begin{pmatrix} VL \\ VS \end{pmatrix} + \varepsilon$$

où C , de taille 2×2 , est la restriction de la matrice du modèle aux variables concernées et ε l'erreur de modélisation pour les coordonnées X-Z du *point équivalent EMA*. En supposant (X_{eqEMA}, Z_{eqEMA}) connu, VL et VS peuvent être estimés à partir de l'équation précédente en annulant l'erreur de modélisation pour les coordonnées X-Z de ce point :

$$\begin{pmatrix} \widehat{VL} \\ \widehat{VS} \end{pmatrix} = C^{-1} \begin{pmatrix} X_{eqEMA} \\ Z_{eqEMA} \end{pmatrix}$$

Notons que les valeurs ainsi estimées de VL et VS permettent une erreur de reconstruction nulle sur les coordonnées X-Z du *point équivalent EMA* mais peuvent conduire à une modélisation non optimale pour l'ensemble des autres points du maillage. Ceci revient en d'autres termes à propager l'erreur de modélisation ε du *point équivalent EMA* vers les autres points. Géométriquement, cette inversion revient à projeter comme sur la figure 3.32 (gauche) le *point équivalent EMA* dans le plan X-Z sur les droites d'équation $VS = 0$ et $VL = 0$ obtenues par le déplacement de ce point à l'aide du modèle.

La majeure partie de la variance du voile du palais étant expliquée par le paramètre VL, se pose la question de sa reconstruction 3D par ce paramètre seul. Il suffit alors de ne conserver que les valeurs de \widehat{VL} dans l'inversion articulaire proposée ci-dessus. Néanmoins, cette estimation seule ne correspond alors plus à la reconstruction optimale pour le *point équivalent EMA* mais plutôt à la reconstruction optimale de sa coordonnée Z, comme le montre la figure 3.32

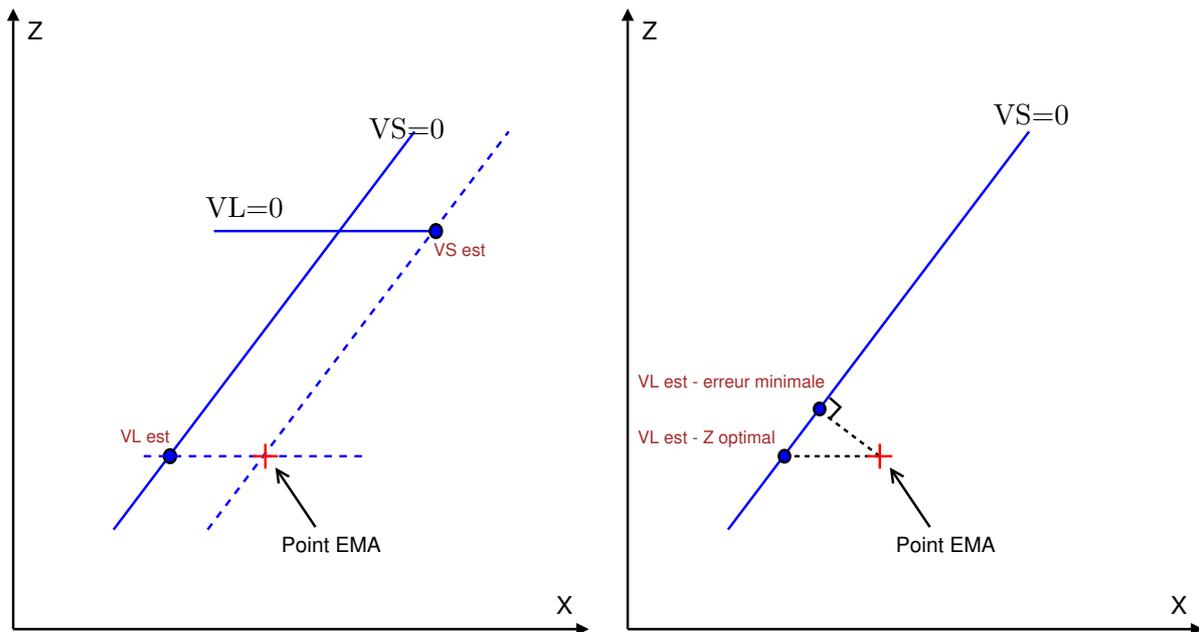


FIG. 3.32 – Estimations des valeurs de VL et VS combinés (gauche) et VL seul (droite) à partir des coordonnées d'un point EMA dans le plan médiosagittal. Les droites d'équations $VS = 0$ et $VL = 0$ correspondent au déplacement du *point équivalent EMA* par le modèle dans ce plan avec VL ou VS nul.

(droite). Minimiser l'erreur de modélisation ε pour ce point revient géométriquement à chercher la valeur de VL qui amène le *point équivalent EMA* modélisé sur la droite d'équation $VS = 0$ le plus proche possible du *point équivalent EMA* mesuré. La figure 3.32 (droite) montre les deux valeurs possibles pour estimer VL seul. Les analyses sur l'ensemble de la forme 3D pour ces deux valeurs révéleront que minimiser l'erreur de modélisation sur le *point équivalent EMA* apparaît légèrement sous-optimal par rapport à la première approche qui minimise l'erreur sur la coordonnée Z. Cette dernière sera donc préférée dans le cas d'une inversion articulaire à un seul paramètre.

Les valeurs de VL et VS ainsi estimés présentent de fortes corrélations avec leurs correspondant optimaux sur les 46 articulations, de coefficients 0.98 pour VL et 0.95 pour VS. Une reconstruction assez fidèle de la surface 3D du voile du palais est par conséquent attendue. L'erreur RMS de reconstruction du voile du palais 3D avec ces valeurs estimées des paramètres VL et VS est évaluée par rapport à deux références :

- d'une part par rapport au voile du palais 3D estimé avec les valeurs optimales des deux paramètres VL plus VS : l'erreur serait donc de zéro si les valeurs des paramètres étaient estimées toutes les deux avec exactitude ; on l'appelle *erreur RMS relative*.
- d'autre part par rapport au voile du palais 3D constituant les données ayant servi à la construction du modèle, de manière à toujours rester rattaché à la réalité ; on l'appelle *erreur RMS totale*.

Notons que ces deux erreurs sont des erreurs RMS *point à point* au sens défini au para-

graphe 3.5.1. Les résultats des reconstructions sont fournis dans le tableau 3.2.

<i>Paramètre</i>	<i>Erreur RMS de reconstruction relative cumulée</i>	<i>Erreur RMS de reconstruction totale cumulée</i>	<i>Erreur RMS par rapport à VL</i>	<i>Erreur RMS par rapport à VS</i>
\widehat{VL}	0.055 cm	0.08 cm	4 %	
\widehat{VS}	0.035 cm	0.07 cm		5 %

TAB. 3.2 – Erreurs RMS totale et relative de reconstruction du voile du palais 3D par les valeurs estimées des paramètres VL et VS à partir du *point équivalent EMA* et erreurs RMS de l'estimation de ces valeurs par rapport aux valeurs optimales des paramètres VL et VS.

Conformément aux attentes formulées, on retrouve presque exactement les erreur RMS *totales* de reconstruction obtenues grâce aux valeurs optimales des paramètres (voir le tableau 3.1 page 106). On constate une erreur d'estimation des deux paramètres assez faible, malgré une proportion de bruit plus importante à attendre lors de l'estimation de la valeur de VS.

Ce paragraphe montre qu'il est possible d'estimer la forme tridimensionnelle du voile du palais, et par extension du port vélopharyngé, à partir d'enregistrements EMA avec un niveau de reconstruction presque optimal pour le modèle. Une approche analogue, détaillée au paragraphe suivant, consiste à l'estimer à partir de contours médiosagittaux.

3.7.2 Des contours médiosagittaux vers la forme 3D

Comme souligné par [Narayanan *et al.*, 2004], on assiste ces dernières années à un développement significatif de l'acquisition de données articulatoires dynamiques du conduit vocal par imagerie IRM temps-réel ou assimilé. En dépit des contraintes techniques rencontrées (fréquence temporelle d'enregistrement difficile à augmenter et qualité variable des images en particulier), les résultats récemment obtenus (voir par exemple [Shadle *et al.*, 1999], [Demolin *et al.*, 2000], [Mády *et al.*, 2002] ou [Narayanan *et al.*, 2004]) et les larges perspectives visibles laissent envisager un développement conséquent de cette technique. Elle permet un suivi dynamique par imagerie IRM du conduit vocal dans le plan médiosagittal. Par traitement d'image il est donc possible de suivre dynamiquement les contours médiosagittaux du conduit vocal, et en particulier du voile du palais, pour des tâches de parole réelle. La connaissance du passage du contour médiosagittal du voile du palais vers sa forme 3D, et par extension vers celle du port vélopharyngé, autoriserait donc l'obtention synchrone de la forme du port vélopharyngé en 3D à partir d'IRM dynamiques médiosagittales. Dans la continuité des résultats précédents, ce paragraphe évalue la qualité de la reconstruction 3D du voile du palais à partir de son contour médiosagittal.

Les valeurs des paramètres VL et VS sont estimées par minimisation d’une fonction d’erreur définie comme l’erreur RMS *point à surface* entre le contour cible médiosagittal et le contour obtenu par la recoupe médiosagittale du maillage du voile du palais 3D correspondant aux valeurs de VL et VS en entrée. Ces dernières sont ainsi estimées pour toutes les articulations à l’aide de la fonction de minimisation sans contrainte *fminunc* de MATLAB. Les résultats sont présentés dans le tableau 3.3. On y observe une excellente estimation du paramètre VL, dont

<i>Paramètre</i>	<i>Erreur RMS de reconstruction relative cumulée</i>	<i>Erreur RMS de reconstruction totale cumulée</i>	<i>Erreur RMS par rapport à VL</i>	<i>Erreur RMS par rapport à VS</i>
\widehat{VL}	0.05 cm	0.08 cm	2 %	
\widehat{VS}	0.04 cm	0.07 cm		10 %

TAB. 3.3 – Erreurs RMS totale et relative de reconstruction du voile du palais 3D par les valeurs estimées des paramètres VL et VS à partir des contours médiosagittaux et erreurs RMS de l’estimation de ces valeurs par rapport aux valeurs optimales des paramètres VL et VS.

le coefficient de corrélation avec le paramètre optimal s’élève à 0.99. Le second paramètre, de coefficient de corrélation 0.81 avec le paramètre optimal, s’avère moins bien prédit qu’avec le *point équivalent EMA*, bien que les conséquences en terme d’erreur RMS de reconstruction soient faibles en raison d’une faible contribution intrinsèque de ce paramètre au mouvement total.

Les résultats présentés dans ce paragraphe montrent la possibilité de reconstruire la forme 3D du port vélopharyngé à partir du contour médiosagittal du voile du palais avec une erreur presque nulle sur VL et une estimation pour le moins correcte de VS. Un suivi temporel médiosagittal des contours du voile du palais, par exemple par IRM temps-réel, suffit pour obtenir une estimation temporelle tridimensionnelle fiable de la forme géométrique du port vélopharyngé.

Les résultats présentés dans cette section sont très satisfaisants en terme d’évaluation. La méthode décrite permet de relier la modélisation statique d’une part et le suivi dynamique d’autre part en combinant les avantages des deux techniques : suivi dynamique des articulateurs du conduit vocal dans un plan médiosagittal et récupération d’une forme tridimensionnelle précise. Le passage de la forme 3D aux fonctions d’aire, présenté dans le chapitre suivant, permettra alors d’obtenir dynamiquement des fonctions d’aire précises du conduit vocal par un suivi EMA ou IRM. Le couplage avec un modèle acoustique rendra possible la création de parole artificielle par *imitation* de la position des articulateurs du conduit vocal, avant l’étape de contrôle du modèle articulatoire pour faire de la synthèse proprement dite. Avant d’étudier le passage des formes tridimensionnelles des organes aux fonctions d’aire pour coupler avec un

modèle acoustique, les relations articulatoires du voile du palais avec la langue sont analysées dans la section suivante.

3.8 Le rôle de la langue

La position du voile du palais à la frontière entre les cavités pharyngée, buccale et nasale dans le conduit vocal soulève la question de son interaction avec les organes environnants. Le lien fort avec la paroi du pharynx a déjà été mis en lumière dans les sections précédentes. Qu'en est-il alors de ses relations avec l'organe central du conduit vocal, la langue ? Le muscle *palatoglosse* relie ces deux organes, même si son rôle en parole est encore controversé (voir par exemple [Fritzell, 1969], [Lubker *et al.*, 1972], [Bell-Berti, 1976] ou [Dixit *et al.*, 1987]). Le modèle tridimensionnel révèle à ce sujet quelques corrélations. Un contact mécanique est en outre observé dans le plan médiosagittal entre la luette et l'arrière ou le dos de la langue pour certaines configurations (voir [Demolin *et al.*, 2003]). Le mouvement du voile du palais lié au second paramètre de contrôle du modèle de [Rossato, 2000] (voir figure 3.11, droite, page 104) suggère un mouvement avant-arrière de la luette coordonnée à celui de la langue en cas de contact. La section présente analyse en ces termes la relation du voile du palais avec la langue : quel lien *moteur* ? quel lien *mécanique* ?

3.8.1 Un lien moteur ?

Le rôle en parole du muscle palatoglosse qui relie les racines latérales de la langue au voile du palais en formant les *piliers antérieurs* a donné lieu à des interprétations variées. Son action sur le voile du palais reste en particulier controversée. Si certaines interprétations liées au modèle *gate-pull* de [Lubker *et al.*, 1972] admettent l'idée que le palatoglosse joue un rôle dans le mouvement d'abaissement du voile du palais de manière antagoniste au *levator veli palatini* ([Fritzell, 1969], [Lubker *et al.*, 1970], [Moon *et al.*, 1994], [Wrench, 1999]), d'autres ne trouvent pas de rôle actif du palatoglosse dans le mouvement du voile du palais ([Bell-Berti et Hirose, 1973], [Bell-Berti, 1976]), tandis enfin que son rôle actif pourrait finalement dépendre du contexte vocalique ([Dixit *et al.*, 1987]). Le manque de certitude de l'une ou l'autre des théories mérite que le lien entre la hauteur de la racine de la langue (où est attaché le palatoglosse) et celle du voile du palais soit analysé dans les données présentes. Si la hauteur du voile du palais apparaît plus ou moins liée directement à celle de la racine de langue, alors cette dernière pourra être utilisée comme prédicteur pour augmenter l'explication de la variance du modèle. C'est donc une corrélation significative entre ces deux hauteurs qui est cherchée. Notons cependant que l'absence d'une telle corrélation générale ne signifierait pas pour autant que dans certains contextes bien définis le palatoglosse ne contribue pas à l'abaissement du voile du palais. Cette question encore partagée ne fait pas l'objet de l'étude de ce paragraphe.

Anatomiquement la racine de la langue est attachée à la mâchoire. On s'intéresse donc dans notre cas à la hauteur de la mâchoire. Dans les modèles de langue (voir par exemple [Maeda,

1990], [Beautemps *et al.*, 2001] ou [Badin *et al.*, 2002]), cette hauteur est classiquement considérée comme celle du point *Lower Teeth* (LT), équivalent pour la mâchoire inférieure du point UT défini dans ce document au paragraphe 2.4.1. Le positionnement de la mâchoire inférieure sur les différents blocs d'images IRM 3D dans notre étude (voir paragraphe 2.4.5), nous a permis d'obtenir la valeur de la hauteur du point LT pour les 46 articulations du corpus. La hauteur du point *équivalent EMA* est choisie comme marqueur de la hauteur du voile du palais.

Le coefficient de corrélation entre ces deux hauteurs s'élève à 0.53 sur les 46 articulations du corpus. Un faible effet de l'abaissement de la mâchoire est donc observable sur la hauteur du voile du palais. Pour quantifier ce phénomène en termes de composante du modèle, une régression linéaire des 46 observations des maillages du voile du palais original, c'est-à-dire les données brutes auxquelles la contribution de l'inclinaison de la tête a été soustraite, est calculée sur le paramètre de hauteur de la mâchoire. Il apparaît que ce paramètre explique jusqu'à 24% de la variance totale du voile du palais. À titre de comparaison, elle explique sur le même sujet 17% de la variance totale de la langue 3D pour le modèle de [Badin *et al.*, 2002] et 22% pour le modèle de [Badin et Serrurier, 2006]. Il paraît donc à ce point raisonnable d'utiliser la hauteur de la mâchoire comme un prédicteur de la forme tridimensionnelle du voile du palais : elle est un des articulateurs premiers du conduit vocal et sa mesure est facile. Par ailleurs le lien biomécanique avec le voile du palais par l'intermédiaire du palatoglosse semble logique.

Pour valider cette approche, nous avons mené des analyses similaires sur le corpus EMA, beaucoup plus large et plus représentatif de la parole réelle. Les hauteurs du voile du palais et de la mâchoire sont mesurées comme les coordonnées verticales des bobines collées sur le voile du palais et sur le point LT (voir la figure 2.10 page 52). Les espaces de déplacement des deux bobines pour l'ensemble des séquences VCV du corpus sont visibles figure 3.33 (gauche). La hauteur du voile du palais en fonction de la hauteur de la mâchoire pour l'ensemble des séquences VCV du corpus EMA est affiché sur la figure 3.33 (droite). On constate une couverture presque complète de l'espace défini par les deux hauteurs. Ceci signifie clairement qu'elles ne sont pas corrélées *en général* et que n'importe quelle hauteur du voile du palais peut être atteinte quelle que soit la hauteur de la mâchoire, et réciproquement.

Nous avons finalement mis en valeur dans nos données une corrélation de corpus et non pas une corrélation liée à la biomécanique. La hauteur de la mâchoire n'est donc pas utilisée comme prédicteur de notre modèle, en dépit de la forte quantité de variance expliquée.

Nous avons cependant observé une forte corrélation entre les hauteurs de la mâchoire et du voile du palais (0.88) pour un sous-corpus limité aux cibles vocaliques non-nasales. Cette observation ne remet donc pas en cause l'idée qu'une relation entre la hauteur de la langue et celle du voile du palais puisse exister par l'intermédiaire du palatoglosse pour certains contextes spécifiques, mais ne doit pas être prise en compte dans la construction d'un modèle général.

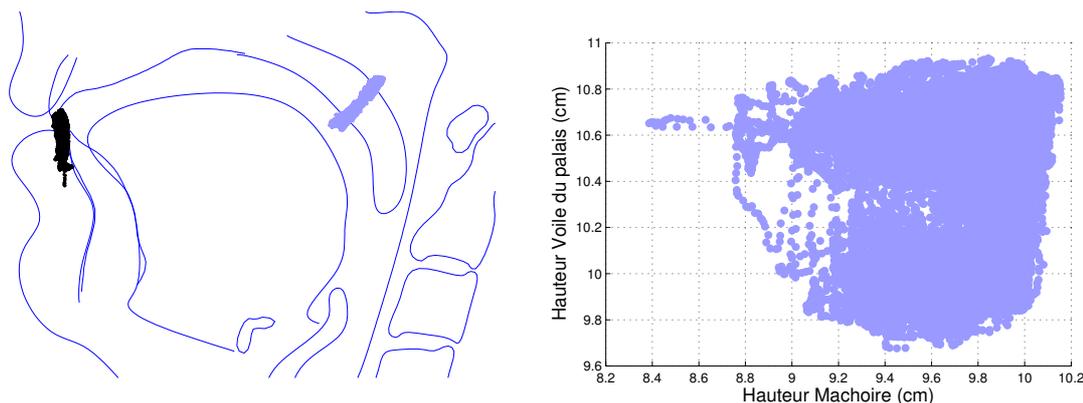


FIG. 3.33 – Espace de déplacement des coordonnées médiosagittales des bobines du voile du palais et de la mâchoire pour l’ensemble des séquences VCV superposé à un contour médiosagittal du conduit vocal en configuration repos (gauche) ; Coordonnée verticale de la bobine du voile du palais *vs.* coordonnée verticale de la bobine de la mâchoire pour l’ensemble des séquences VCV.

3.8.2 Un lien mécanique ?

L’observation de coupes médiosagittales du conduit vocal montre un contact entre le dos de la langue et la luette pour certaines configurations de langue reculée, déjà mentionné dans la littérature ([Demolin *et al.*, 2003]). Sur le corpus des 46 articulations, on en dénombre 18 avec un contact entre la langue et le voile du palais et 28 sans. Les coupes médiosagittales pour les quatre voyelles nasales du français, / \bar{a} /, / \bar{o} /, / \bar{e} / et / $\bar{\epsilon}$ /, sont visibles à la figure 3.34. Sur les coupes des voyelles avant / \bar{e} / et / $\bar{\epsilon}$ /, en dépit d’un contact très faible observé pour / \bar{e} /, le voile du palais prend sa forme courbée caractéristique dans le plan médiosagittal. Sur les coupes des voyelles arrières / \bar{a} / et / \bar{o} /, on observe un contact marqué entre le dos de la langue et la luette. L’étude des coupes médiosagittales suggère un recul de la luette provoqué par la position très arrière de la langue pour ces deux articulations. Le voile du palais prend alors une forme plus allongée et moins courbée que pour les cas sans contact. Ce phénomène est purement mécanique et ne semble pas avoir d’origine musculaire. Il est par ailleurs suggéré par les nomogrammes du second paramètre du modèle de [Rossato, 2000] (voir la figure 3.11, droite, page 104), qui expliquait 8% de la variance totale médiosagittale. Notons qu’un tel phénomène ne se retrouve pas dans le modèle tridimensionnel, si ce n’est, peut-être, un très léger mouvement de la luette dû au paramètre VS. La volonté de prendre en compte ce phénomène dans la modélisation repose sur deux motivations : d’une part pour augmenter le degré de précision du modèle 3D, et d’autre part pour tenir compte d’un mécanisme qui pourrait modifier significativement l’aire de couplage du port vélopharyngé, et donc avoir une influence acoustique non négligeable. Nous proposons deux approches distinctes pour tenter de la prendre en compte dans notre modèle.

La première approche envisagée fut de trouver un paramètre articulatoire classique, tel le second paramètre du modèle de [Rossato, 2000]. Considérant le lien avec le mouvement avant-arrière de la langue, voire le mouvement d’arrondi du dos de la langue, nous avons cherché une

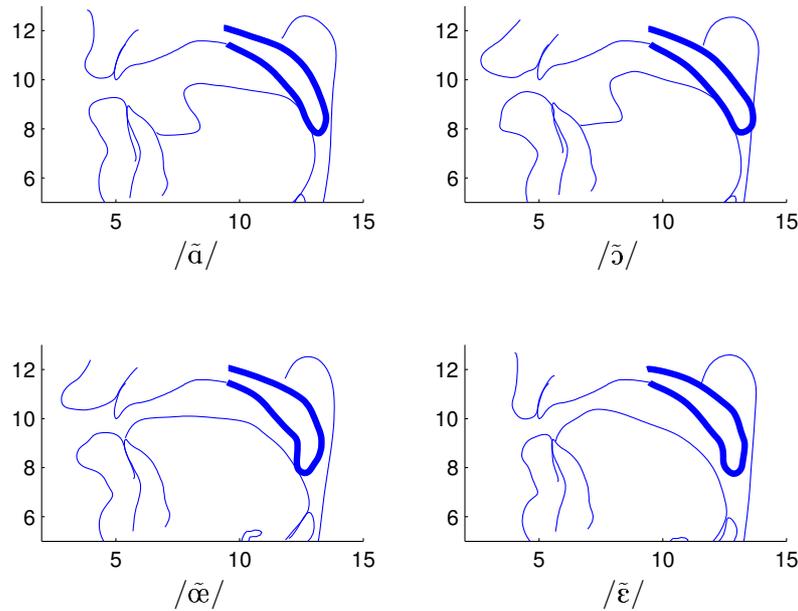


FIG. 3.34 – Contours médiosagittaux du conduit vocal pour les quatre voyelles nasales artificiellement soutenues du français pour le sujet en position allongé.

corrélation avec les paramètres associés TB et TD d'un modèle de langue construit sur le même corpus (voir par exemple [Badin *et al.*, 2002]). Nous avons construit pour cela un modèle de voile du palais similaire au modèle considéré jusqu'ici, mais seulement à partir des 28 articulations sans contact pour éliminer toute perturbation introduite par langue. Ce modèle est ensuite étendu aux 18 autres articulations par régression linéaire en calculant sur la région non uvulaire seulement, considérée comme non perturbée par la langue, la valeur des paramètres associés à ces articulations. Le résultat est une forme de voile du palais prédite pour les 46 articulations et libre de toute perturbation introduite par la langue. Naturellement les 28 articulations sans contact avec la langue sont globalement bien prédites tandis que les 18 avec contact le sont beaucoup moins, en particulier dans la région uvulaire. Le résidu des données qui n'a pas été expliqué par un tel modèle possède une variance forte dans la région de la luette. Nous avons cherché sur ce résidu une corrélation avec un mouvement de recul ou d'arrondissement du dos de la langue. Un modèle médiosagittal de langue similaire à celui de [Badin *et al.*, 2002] a été construit à partir des contours tracés sur nos images IRM pour les 46 articulations. Les paramètres TB et TD correspondent aux mouvements de langue considérés dans notre cas. Ils sont utilisés comme prédicteurs sur le résidu des données non expliqué à l'issue du modèle de voile du palais décrit ci-dessus. Aucune contribution supérieure à 5% de la variance totale des données et correspondant à un mouvement réaliste du voile du palais n'a pu être révélée. De même, une tentative d'explication de la variance du résidu en utilisant directement comme prédicteur une mesure des coordonnées du point de la langue en contact avec le voile du palais a conduit à la même conclusion. Nous avons donc conclu qu'aucune corrélation n'existait entre les paramètres

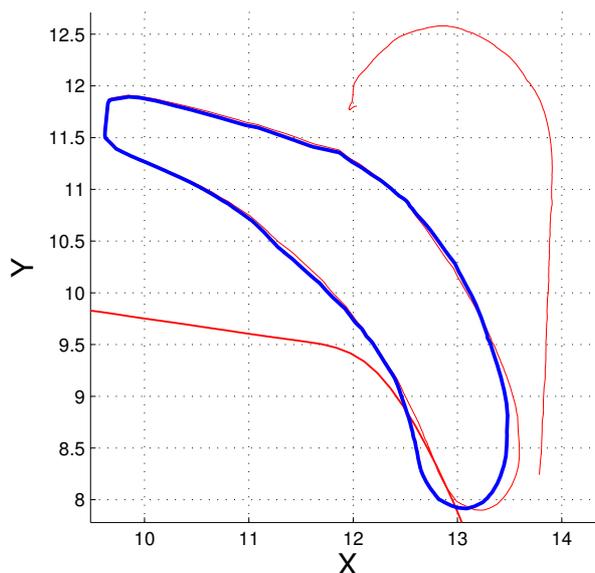


FIG. 3.35 – Contours médiosagittaux (ou assimilés) de la langue, du voile du palais et de la paroi du pharynx issus des données (traits fins) pour l’articulation / \tilde{a} / et contour médiosagittal du voile du palais 3D estimé par le modèle (trait épais).

du modèle de langue dans la région vélaire et le mouvement du voile du palais ; cette approche a donc été abandonnée.

La seconde approche, plus originale, se base sur la non-linéarité du phénomène et son caractère purement mécanique lié à la position de la langue. Le phénomène considéré n’est plus pris en compte par une composante articulatoire supplémentaire dans le modèle du voile du palais mais par une simulation géométrique de la perturbation mécanique du voile du palais par la langue. Le voile du palais est prédit par le modèle tridimensionnel classique (il aurait aussi pu l’être par le modèle libre de toute perturbation de langue présenté ci-dessus mais les résultats se sont avérés similaires) qui ne prend *a priori* pas en compte la perturbation introduite par la langue. Il arrive alors que le voile du palais prédit par ce modèle entre en collision avec la langue, comme on le voit pour le / \tilde{a} / sur la figure 3.35. L’influence de la langue est alors prise en compte en *repoussant* géométriquement de la distance de pénétration la région uvulaire du voile du palais en collision avec la langue. On fait donc l’hypothèse que la langue n’est absolument pas perturbée par le voile du palais, car beaucoup plus rigide, et que l’épaisseur de chair de ce dernier ne varie pas. La langue est à nouveau considérée comme un élément perturbateur du positionnement actif du voile du palais par le sujet. En pratique, dans le plan médiosagittal, une direction de déformation du voile du palais est donnée par les droites issus de la grille semi-polaire classique (voir l’exemple de / \tilde{a} / sur la figure 3.36). Chaque point du contour extérieur du voile du palais est alors reculé selon son orientation et la distance de pénétration sur la grille (figure 3.36, milieu). Le résultat est un contour médiosagittal reculé de manière à venir épouser celui de la langue (figure 3.36, droite). Le processus est facilement étendu à la troisième dimen-

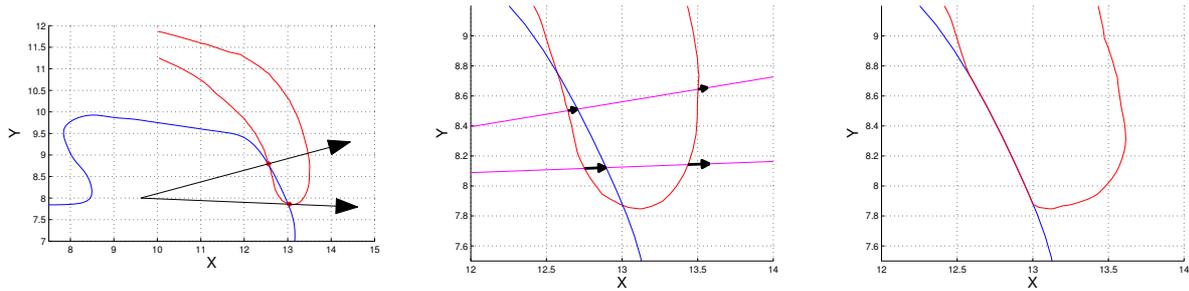


FIG. 3.36 – Processus géométrique de recul du voile du palais par la langue : définition d’une direction de déformation selon la grille (gauche), recul des points extérieurs du voile du palais selon leur orientation et la distance de pénétration dans la langue (milieu) et recul global obtenu par ce procédé (droite) pour une articulation $/\tilde{a}/$.

sion en reproduisant le décalage introduit dans le plan médiosagittal à tous les points de la luvette.

Comme le montre la figure 3.36 (droite), le contour obtenu, et par extension le maillage, n’est pas réaliste dans le sens où il est anguleux. Il est bien évident qu’un lissage est nécessaire pour se conformer de manière plus plausible à la réalité, mais les résultats présentés ici visent simplement à évaluer le principe d’une telle modélisation. La prise en compte de ce phénomène améliore sensiblement la reconstruction du voile du palais dans la région uvulaire, comme l’atteste la figure 3.37 pour l’articulation $/\tilde{a}/$. Cependant, sur l’ensemble des 18 articulations des données avec contact, seules 5 s’avèrent entrer effectivement en collision avec la langue à l’issue de la modélisation initiale : $[\tilde{a}, \tilde{o}, k^a, b^a \text{ et } v^u]$. Le poids joué par la petite correction apportée à quelques points (ceux de la luvette), parmi l’ensemble des points du voile du palais 3D, pour 5 articulations sur les 46 s’avère complètement négligeable dans le calcul de l’erreur RMS de reconstruction globale.

Néanmoins ces petits mouvements de luvette peuvent jouer un rôle localement important. Ils pourraient en effet modifier sensiblement l’aire de couplage, c’est-à-dire l’aire minimale du port vélopharyngé. Pour évaluer leur impact sur l’aire de couplage, nous avons calculé la fonction d’aire du port vélopharyngé (avec un peu d’avance sur la chapitre suivant qui détaille cette procédure) et déterminé l’aire de couplage du conduit nasal. Si l’erreur RMS globale de l’aire de couplage pour ces cinq articulations passe de 0.19 cm^2 à 0.11 cm^2 , c’est-à-dire proche de l’erreur globale sur les 46 articulations, la figure 3.38 montre encore de larges erreurs pour certaines de ces articulations. On observe malgré tout comme tendance que le recul mécanique du voile du palais par la langue comme nous l’avons proposé tend à faire diminuer l’aire de couplage du port vélopharyngé pour la rapprocher de celle des données.

La question de prendre en compte dans la modélisation ce phénomène se pose légitimement. D’une part une observation de la réalité et une modélisation qui rend compte, avec certaines erreurs, du phénomène observé, d’autre part une modélisation complexe qui n’améliore que de

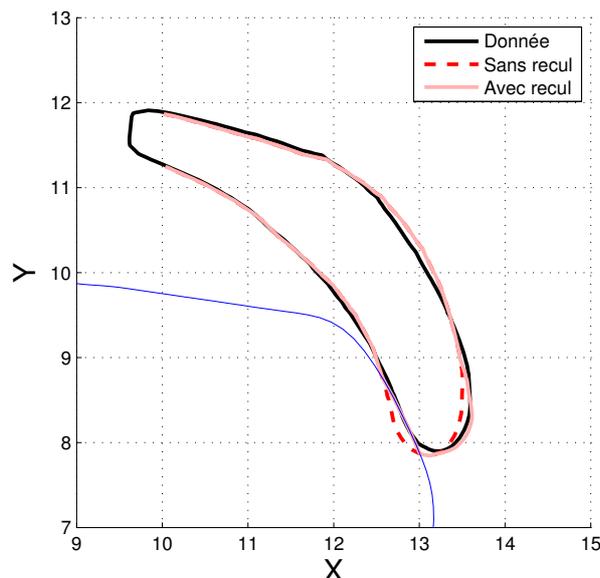


FIG. 3.37 – Contours médiosagittaux (ou assimilés) de la langue (trait fin) et du voile du palais issus des données ainsi que modélisé sans et avec prise en compte de la collision avec la langue (traits épais) pour l’articulation $/\tilde{a}/$.

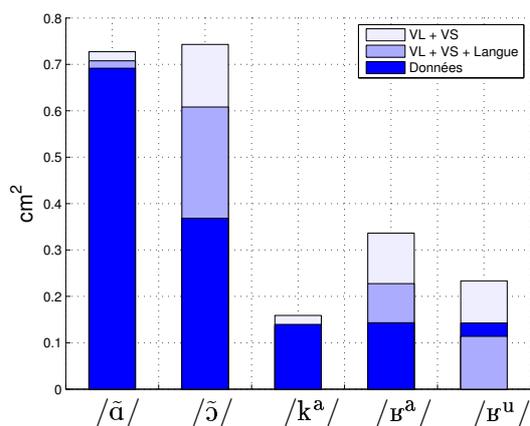


FIG. 3.38 – Aires de couplage vélopharyngé pour les 5 articulations modélisées avec la prise en compte d’une correction due à la collision avec la langue : aires de couplage comparées des données (barres foncées), du port vélopharyngé modélisé par les deux paramètres VL et VS du modèle (barres claires) et du port vélopharyngé ainsi modélisé avec un recul de la luvette par la langue (barres grisées).

manière infime les résultats quantitatifs du modèle global. Nous avons donc choisi de ne pas prendre en compte ce phénomène dans la suite de l’étude. Néanmoins l’analyse présentée dans ce paragraphe souligne un comportement du voile du palais, et de la luvette en particulier, en interaction avec la langue pour certaines articulations. Par ailleurs on note que la position de la langue peut indirectement modifier l’aire de couplage du port vélopharyngé. L’impact de l’in-

teraction clairement présente entre le voile du palais et la langue sur le comportement du voile du palais est donc un point qui renferme encore quelques secrets et qui mérite d’être encore analysé...

3.9 Conclusion

Nous avons présenté dans ce chapitre le cœur du travail de thèse. Ce travail a permis d’aboutir à un modèle complet du conduit nasal, de la pointe de la luvette jusqu’aux sorties narinaires, avec une erreur de reconstruction inférieure au millimètre. La géométrie complexe des fosses nasales largement détaillée a été comparée à la littérature, et s’est trouvée conforme aux études récentes. Après la mise en évidence d’un mouvement global d’inclinaison de la tête et la soustraction de sa contribution aux données, la partie déformable du conduit nasal, c’est-à-dire le port vélopharyngé, a été modélisée en deux et trois dimensions. Les deux modèles se sont révélés très proches l’un de l’autre : un degré de liberté articuloire principal, expliquant autour de 85% de l’ensemble des mouvements, contrôle l’ouverture et la fermeture du port vélopharyngé par un fort mouvement de remontée du voile du palais selon un axe oblique orienté vers 45° , attribué au muscle *levator veli palatini*. Conjointement, une avancée de la paroi du pharynx dans sa région supérieure, attribuée au comportement de sphincter du port vélopharyngé, vient compléter la fermeture vélaire. Un second paramètre du voile du palais, expliquant autour de 7% de l’ensemble des mouvements semble venir compléter finement ce comportement global. Le modèle 3D s’est révélé par ailleurs en totale cohérence avec les enregistrements dynamiques EMA, justifiant l’approche choisie. De très fortes corrélations ont été mises en évidence entre le modèle tridimensionnel et des mesures articuloires médiosagittales du voile du palais, de telle sorte qu’il est possible de prédire avec une erreur très faible, toujours inférieure au millimètre, la forme tridimensionnelle du port vélopharyngé à partir de ces mesures articuloires. Enfin, le voile du palais a été replacé dans son contexte d’articulateur du conduit vocal pour analyser quelques interactions biomécaniques avec la langue et mettre en lumière une interaction mécanique qui modifie sensiblement l’aire de couplage du port vélopharyngé.

Dans le cadre de l’étude des conséquences acoustiques des mouvements articuloires, l’étape suivante consiste à calculer les fonctions d’aire du conduit vocal. Le conduit vocal est alors vu comme un tuyau dont les propriétés géométriques sont obtenues à partir de la forme tridimensionnelle donnée par le modèle. Cette étude fait l’objet du chapitre suivant.

Chapitre 4

Fonctions d'aire

Le développement d'un modèle articulatoire tridimensionnel ouvre la voie à l'analyse du signal acoustique simulé à l'aide du modèle. Dans les basses fréquences, jusqu'à 5000 Hz environ, les modes de propagation transverses d'une onde dans le conduit vocal peuvent être négligées ([Fant, 1960]). On peut donc considérer que la propagation est plane. Il en résulte que la propagation d'onde est monodimensionnelle le long d'un axe central du conduit vocal. L'aire du conduit vocal varie le long de cet axe, c'est la *fonction d'aire*. Le modèle articulatoire 3D offre la possibilité de calculer des fonctions d'aire réalistes. Nous verrons dans ce chapitre la détermination des fonctions d'aire à partir du modèle et nous étudierons ensuite l'évolution de celles-ci en fonction des deux paramètres de contrôle du modèle.

4.1 Le calcul des fonctions d'aire : principe général

Cette section présente une méthode de calcul des fonctions d'aire à partir de coupes transverses du conduit vocal obtenues à l'aide du modèle tridimensionnel. Pour une propagation d'onde plane, comme on peut considérer que c'est le cas pour une onde acoustique dans le conduit vocal pour des fréquences inférieures à 5000 Hz, la fonction d'aire suffit à définir la géométrie du conduit vocal : c'est l'aire du conduit en fonction de la position le long de l'axe central de propagation. Le conduit peut alors être divisé en une succession de *tubes* cylindriques d'aire constante, dont le nombre dépend du degré de réalisme souhaité, et dans lesquels les équations de propagation d'une onde plane sont facilement calculables ([Fant, 1960], [Stevens, 1998]). Chaque tube est défini par sa longueur et son aire. Un exemple de fonction d'aire, pour les fosses nasales, est visible à la figure 4.5.

La première étape pour calculer la fonction d'aire du conduit vocal est d'échantillonner ce dernier en sections transverses régulièrement réparties entre la glotte et les lèvres ou les narines. La grille d'échantillonnage fixe définie pour cette opération est celle déjà utilisée dans cette étude (voir par exemple au paragraphe 2.2.5 pour les recoupes des images transverses) et visible sur la figure 4.1. Les contours du conduit oral sont alors tracés dans les images IRM correspondant aux plans de coupe pour obtenir des sections *a priori* transverses du conduit (voir un exemple de section transverse à la figure 4.1). Ce processus déjà utilisé depuis la généralisation de l'utilisation des images IRM en parole a montré son efficacité et sa précision (voir par exemple [Baer *et al.*, 1991], [Story *et al.*, 1996] ou [Demolin *et al.*, 1996b]). Dans la région du port vélopharyngée dont la géométrie est décrite par le modèle, les contours transverses sont obtenus par intersection des maillages selon une méthode décrite à la section 4.3. On peut donc considérer acquis pour cette partie l'échantillonnage du conduit vocal en sections transverses selon la grille visible à la figure 4.1. Un exemple complet d'échantillonnage pour le conduit oral articulant un /a/ est montré à la figure 4.2.

L'étape suivante consiste à obtenir en 3D la *ligne médiane* du conduit, c'est-à-dire l'axe central de propagation de l'onde dans le conduit vocal. Celle-ci est définie comme la ligne moyenne pour laquelle la normale au plan du front d'onde est toujours parallèle à la tangente à cette ligne. Sa longueur représente donc une estimation de la longueur totale du conduit entre l'entrée et la

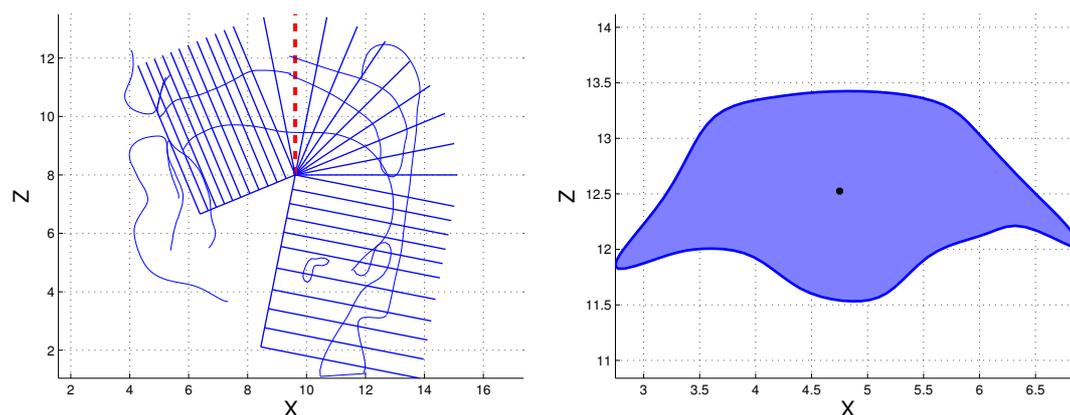


FIG. 4.1 – Superposition de la grille d'échantillonnage du conduit vocal sur les contours médiosagittaux pour une articulation /a/ (gauche) ; section transverse du conduit oral (droite) selon le plan indiqué en pointillé sur la figure gauche. Le point noir à droite est le centre de gravité de la section.

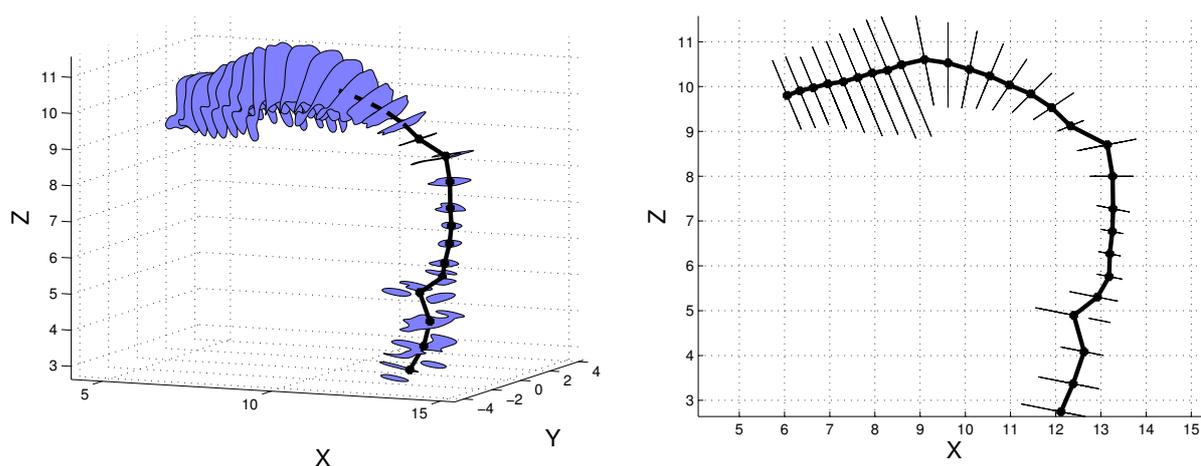


FIG. 4.2 – Vues 3/4 postérieure (gauche) et latérale (droite) des sections transverse qui échantillonnent le conduit oral pour une articulation /a/. Les centres de gravité des sections sont reliés entre eux par une ligne.

sortie ([Beautemps *et al.*, 1995]). Elle passe par les centres de gravité de chaque section transverse définie ci-dessus (ligne continue sur la figure 4.2). Nous considérons qu'un tube élémentaire du conduit est délimité par deux sections transverse. On détermine donc $(n - 1)$ tubes élémentaires pour n sections du conduit. Chaque segment reliant deux centres de gravité consécutifs définira alors la longueur du tube élémentaire calculé entre ses deux sections limitantes respectives. L'aire de chaque tube est calculée comme la moyenne des aires des deux sections qui le délimitent.

Ce calcul fait aussi l'hypothèse que les sections sont exactement perpendiculaires à la ligne

médiane du conduit et que les aires calculées sur ces sections correspondent donc exactement aux aires transverses du conduit aux lieux d'échantillonnage. Il est néanmoins possible que les plans ne soient pas exactement perpendiculaires et que les aires soient alors surévaluées. Comme l'illustre la figure 4.3 gauche, l'aire de la section d'un tube cylindrique coupé par un plan incliné par rapport au plan exactement perpendiculaire augmente de l'inverse du cosinus de son angle d'inclinaison. En trois dimensions cet angle est défini comme l'angle entre le vecteur normal au plan de coupe oblique et celui au plan de coupe exactement perpendiculaire. Admettant dans notre approche une possible inclinaison des plans de coupe déterminés *a priori* par les plans de grille, on cherche à évaluer l'angle défini ci-dessus pour corriger l'aire de la section de son cosinus. On cherche pour cela deux vecteurs : le vecteur normal au plan de coupe, connu, et le vecteur normal au plan de coupe idéalement perpendiculaire à la ligne médiane, qui est aussi égal au vecteur tangent à la ligne médiane du conduit. Pour calculer ce dernier, la direction du conduit est approchée localement sur chaque section par un arc de cercle ; celui-ci passe par les centres de gravité de la section et de ses deux voisins, comme le montre la figure 4.3 (droite). Le vecteur tangent à la ligne médiane pour la section considérée est alors calculé comme le montre l'exemple de la figure 4.3 (droite). L'angle d'inclinaison du plan de coupe est donc estimé de cette manière pour chaque section. Avant le calcul des fonctions d'aire suivant la méthode décrite ci-dessus, l'aire de chaque section est finalement corrigée d'un facteur égal au cosinus de l'angle calculé. Cette méthode permet donc de s'affranchir des variations d'aire introduites par des recoupes mal orientées et obtenir ainsi des fonctions d'aire plus proches de la réalité.

Lorsque le voile du palais est abaissé, le conduit principal se sépare donc en Y vers deux conduits, le premier débouchant au niveau des lèvres, l'autre au niveau des narines. De manière plus générale, la prise en compte dans le calcul des fonctions d'aire de la séparation d'un tuyau

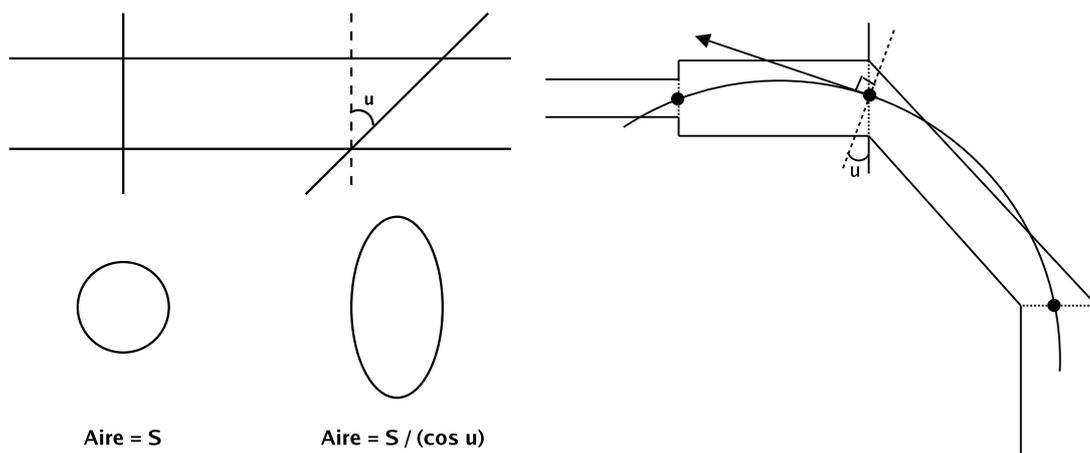


FIG. 4.3 – À gauche, coupes perpendiculaire et oblique d'un tube cylindrique : l'aire de la section oblique augmente de l'inverse du cosinus de son angle d'inclinaison ; à droite, tubes d'un conduit dont la direction au point d'origine du vecteur dessiné est estimé par l'arc de cercle passant aussi par les centres de gravité des deux sections voisines ; la flèche est le vecteur tangent à la direction du conduit pour la section considérée.

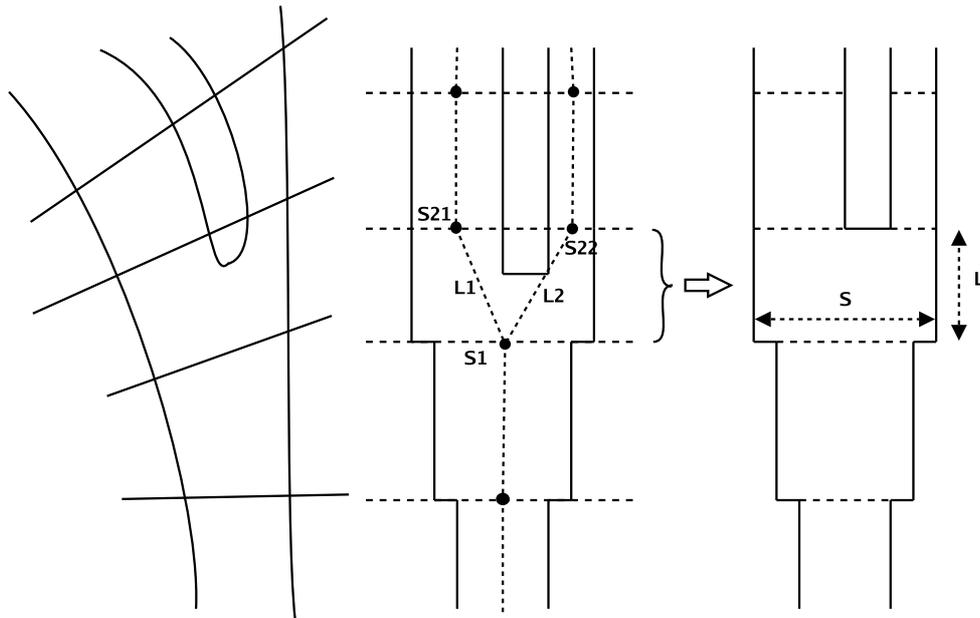


FIG. 4.4 – Modélisation en tubes d'une séparation d'un tuyau en deux : à gauche le tuyau et sa séparation en Y ainsi que les sections transverses correspondantes ; au milieu représentation du tuyau et de sa séparation en tubes élémentaires ; à droite modélisation finale du tuyau en tubes élémentaires.

principal en deux est résumé sur la figure 4.4 : n'ayant aucune connaissance *a priori* sur la position du point de séparation des tuyaux en elle-même hormis les aires obtenues dans les sections transverses des plans de coupe, nous considérons le tube au niveau duquel se produit la séparation comme un tube unique à la sortie duquel les deux tubes secondaires se connectent. Conformément aux notations de la figure 4.4, la longueur L et l'aire S de ce tube sont calculées comme des moyennes pondérées des aires S_1 , S_{21} et S_{22} des deux sections limitantes du tube et des des longueurs L_1 et L_2 associées :

$$S = \frac{(S_{21} + S_{22}) + S_1}{2} \quad \text{et} \quad L = \frac{L_1 S_{21} + L_2 S_{22}}{S_{21} + S_{22}}$$

La proportion de chacun des deux tuyaux est ainsi prise en compte dans la détermination du tube de connexion.

Nous avons présenté dans cette section les bases du calcul des fonctions d'aire d'un conduit à partir de ses sections transverses. Dans les sections suivantes, nous nous appuyons sur ces principes généraux pour déterminer les fonctions d'aire des fosses nasales et des conduits oral et nasal dans la région vélaire.

4.2 Les fosses nasales

Les sections transverses des fosses nasales, directement issues des données, ont été présentées en détail dans la section 3.1. En appliquant la méthode précédemment définie aux sections

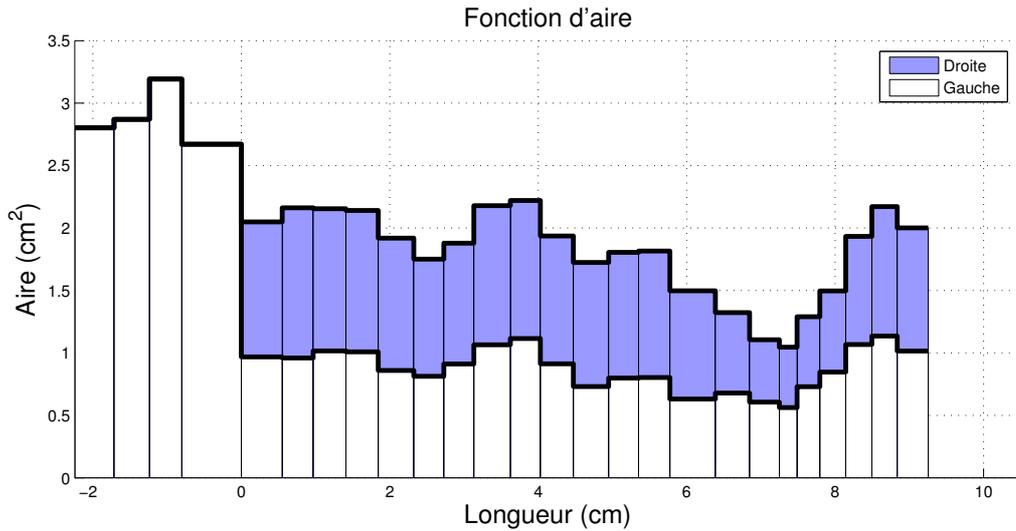


FIG. 4.5 – Fonction d'aire de la partie rigide du conduit nasal, du cavum (à gauche) aux narines (à droite). L'origine de l'axe des abscisses est arbitrairement fixée au point de séparation du conduit nasal en deux choanes.

transverses des fosses nasales, on obtient les fonctions d'aire gauche et droite visibles sur la figure 4.5. Ces fonctions d'aire sont naturellement calculées à partir des contours obtenus des tracés IRM qui constituent pour les fosses nasales une meilleure approximation de la réalité que ceux issus des images CT (voir section 3.1).

On observe une longueur d'environ 9.2 cm pour les fosses nasales (au sens large, c'est-à-dire les deux cavités du point de séparation jusqu'aux narines) et une constriction d'environ 0.6 cm de long et de section 1 cm^2 à 1.7 cm^2 de la sortie narinaire. En comparaison, la longueur des fosses nasales proposées par [House et Stevens, 1956] est approximativement de 8.5 cm ; leur aire décroît de 5.9 cm^2 à 0.23 cm^2 aux narines, considérées comme la constriction. La longueur proposée par [Bjuggren et Fant, 1964], autour de 7-8 cm, comme celle de [Dang *et al.*, 1994], est légèrement plus faible que la nôtre. Au contraire, comme nous l'avons vu dans le chapitre précédent, les aires proposées [Bjuggren et Fant, 1964] apparaissent surévaluées par rapport aux nôtres, même si l'allure est similaire : elles varient de 7 cm^2 dans la partie centrale jusqu'à 2 cm^2 aux narines, ce qui constitue la constriction. Comme nous l'avons déjà vu précédemment, [Dang *et al.*, 1994] expliquent cette différence par une rétraction des muqueuses des fosses nasales du cadavre ayant servi de référence, particulièrement dans la région centrale où elles occupent une place importante, ce qui conduit à une augmentation significative du volume mesuré. Nos mesures sont plus proches de celles rapportées par [Matsumura *et al.*, 1994], moins de 3 cm^2 d'aire pour les régions centrale et antérieure, avec une constriction inférieure à 1 cm^2 et située entre 0.8 cm et 2.8 cm de la sortie narinaire suivant les sujets. De même, comme souligné antérieurement dans ce document pour les sections transverses, nos résultats sont tout à fait cohérents avec ceux de [Dang *et al.*, 1994], pour lesquels l'aire varie entre 1 cm^2 et 4 cm^2 dans la région centrale, 1 cm^2 et 2 cm^2 dans la région narinaire, et une constriction autour de 1 cm^2 située entre 1.5 cm et 3 cm de la sortie narinaire.

L'exploitation de la géométrie des fosses nasales sous forme de fonction d'aire permet d'appliquer la théorie acoustique générale d'analogie électrique pour calculer une fonction de transfert dans les basses fréquences ([Stevens, 1998]). Néanmoins, au vu de la forme tourmentée des cavités nasales, il se peut que les modes de propagation non plan dans les plus hautes fréquences jouent un rôle important. C'est donc en première approximation que la théorie acoustique générale est appliquée, bien qu'une analyse ciblée des fosses nasales devrait être envisagée.

4.3 Le port vélopharyngé

Nous possédons à ce stade un modèle complet tridimensionnel du port vélopharyngé, et plus précisément du voile du palais et de la paroi supérieure du pharynx. En s'appuyant sur le processus de détermination des fonctions d'aire décrit dans le paragraphe 4.1, nous nous intéressons dans cette partie à la fonction d'aire du port vélopharyngé. Celle-ci est calculée à partir des sections transverses obtenues en recoupant les maillages du voile du palais et de la paroi du pharynx selon une grille similaire à celle de la figure 4.1. Un maillage de langue dans la région vélaire est par ailleurs ajouté aux deux organes précédents pour mesurer les sections orales dans la région vélaire. Ce maillage de langue est déterminé à partir des mêmes données IRM que celles de cette étude, selon le même procédé, dans le but de construire un modèle tridimensionnel de langue dont un prototype est proposé par [Badin et Serrurier, 2006]. On dispose ainsi du maillage de la forme de la langue pour les 46 articulations du corpus.

Les trois organes sont recoupés dans la région vélaire selon neuf plans visibles dans le plan médiosagittal sur la figure 4.6. Un exemple de recoupe des trois maillages dans chacun de ces plans est montré à la figure 4.7. Les sections transverses du conduit oral sont alors clairement délimitées par la langue et le voile du palais, voire la paroi du pharynx pour les plans les plus

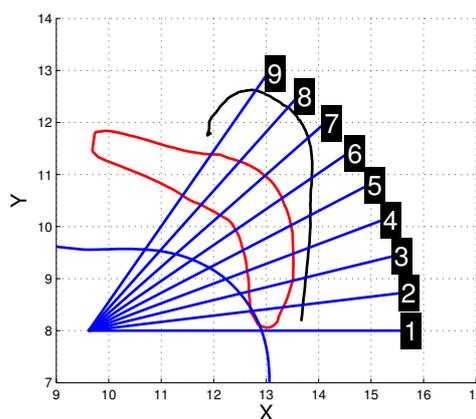


FIG. 4.6 – Grille de recoupe du conduit vocal dans la région vélaire superposée aux contours médiosagittaux de la langue, du voile du palais et de la paroi du pharynx pour l'articulation /a/.

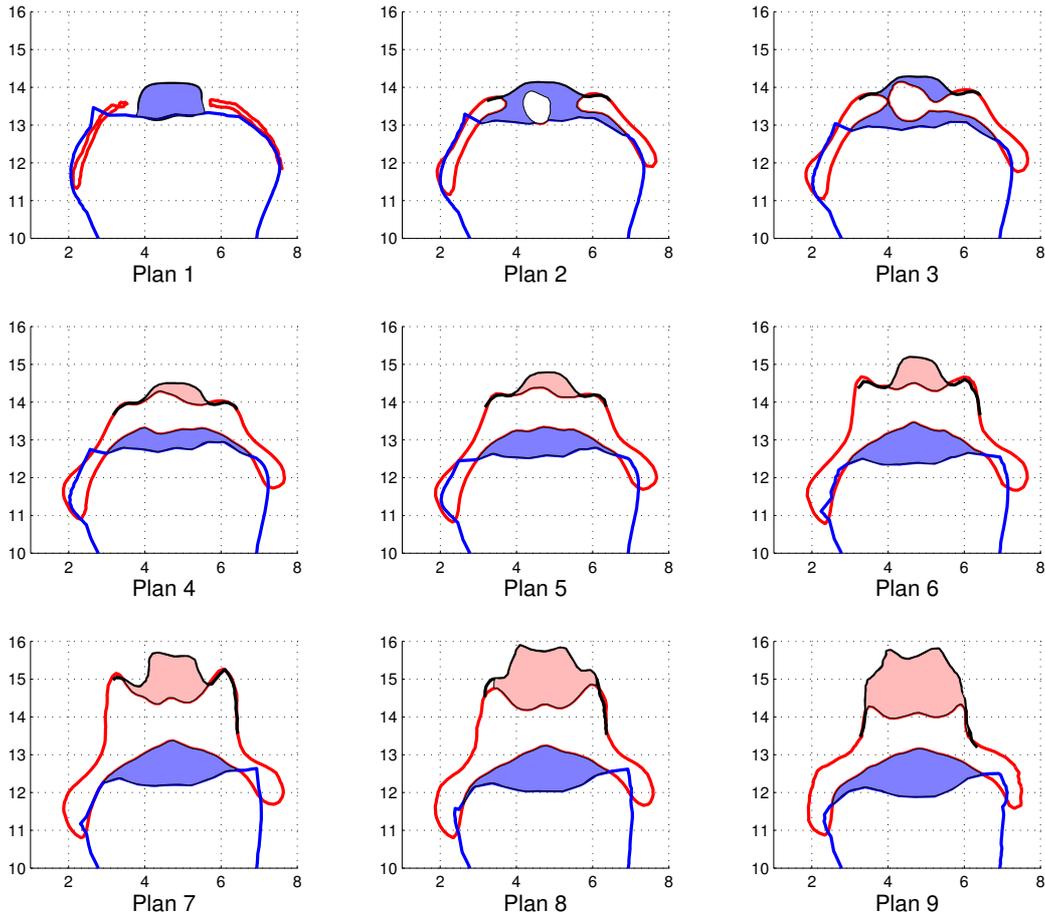


FIG. 4.7 – Affichage plan par plan des sections transverses obtenues par recoupe des maillages de la langue, du voile du palais et de la paroi du pharynx dans la région vélaire selon les plans de grille de la figure 4.6 pour l’articulation /a/ du conduit vocal. Les aires transverses du conduit oral (resp. nasal) sont les aires affichées en couleur plus foncée (resp. plus claire). Les intersections du voile du palais sur les plans 2, voire 3, dessinent une surface pleine isolée au milieu du conduit oral qui représente la luette.

bas où le conduit nasal n’est pas encore formé ; et par le voile du palais et la paroi du pharynx pour le conduit nasal. De manière pratique, ces sections sont déterminées automatiquement en considérant en contact des contours d’organes distants de moins d’un tiers de millimètre, pour éviter des cas de fuite et obtenir des sections fermées. Les sections ainsi déterminées sont affichées plan par plan sur la figure 4.7. De manière intéressante, on note sur le plan 2, et dans une moindre mesure sur le plan 3, que l’intersection avec le voile du palais a révélé un îlot isolé à l’intérieur du conduit oral : il s’agit en réalité de la luette. Sa surface dans ce plan est retranchée à celle de la section orale pour obtenir l’aire réelle du conduit oral dans ce plan. Ces sections sont également représentées en trois dimensions à la figure 4.8 sur laquelle on peut deviner les surfaces du voile du palais, de la langue et de la paroi du pharynx qui viennent respectivement épouser les formes des sections en dessous, au milieu et en arrière des deux conduits.

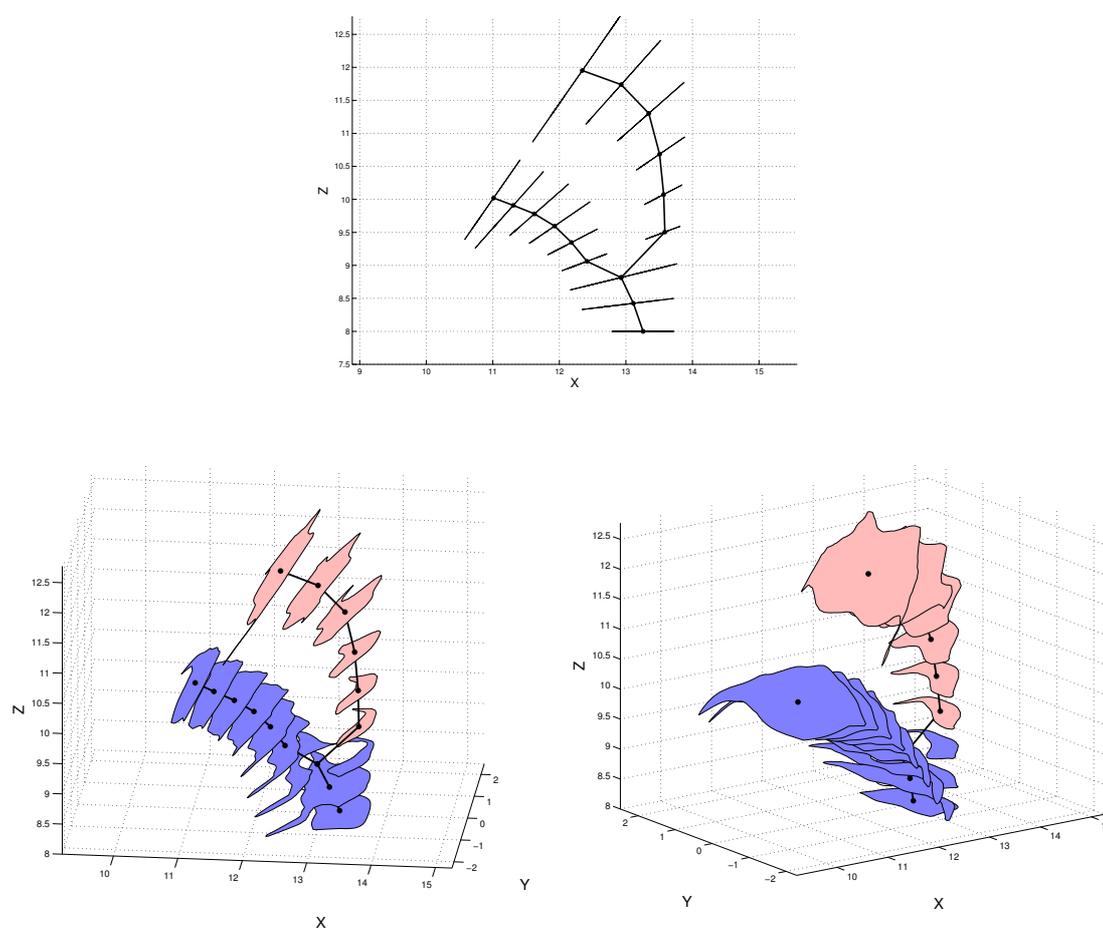


FIG. 4.8 – Vues latérale (haut), latérale postérieur (bas gauche) et latérale antérieure (bas droite) tridimensionnelles des sections transverses qui échantillonnent le conduit vocal dans la région vélaire pour une articulation /a/. Les centres de gravité des sections sont reliés entre eux par une ligne.

Les sections ainsi calculées sont utilisées pour déterminer les fonctions d'aire des conduits oral et nasal dans la région vélaire. La figure 4.9 montre les fonctions d'aire déterminées à partir des sections transverses présentées ci-dessus, c'est-à-dire pour un conduit vocal en articulation /a/. On note au passage une *fuite* vélaire légèrement supérieure à 0.4 cm^2 pour cette articulation du /a/ artificiellement soutenue, confirmant ainsi les résultats de la littérature sur les voyelles basses, en articulations soutenues ou en parole libre (voir par exemple [Clumeck, 1976] pour une revue).

De manière plus générale, il est possible de déterminer l'*aire de couplage* du conduit nasal pour les 46 articulations du corpus. Cette aire est définie comme l'aire minimum de la fonction d'aire du port vélopharyngé. Elle correspond donc à une *constriction* du port vélopharyngé. Cette aire est calculée pour les 46 articulations sur les fonctions d'aire obtenues à partir des

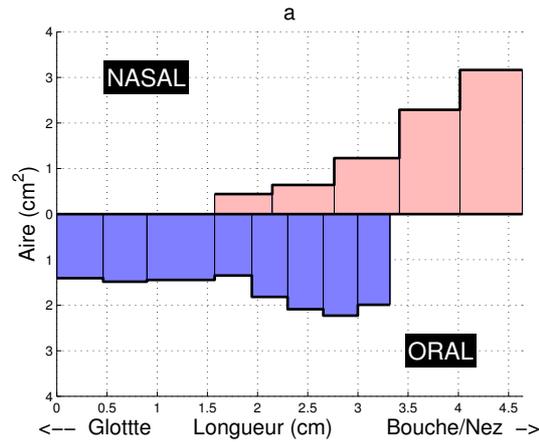


FIG. 4.9 – Fonctions d'aire orale et nasale dans la région vélaire pour l'articulation /a/.

maillages ayant servi à construire le modèle. Les 46 phonèmes artificiellement soutenus peuvent ainsi être classés selon leur degré d'ouverture du port vélopharyngé : voir la figure 4.10 pour un aperçu. Conformément aux résultats de [Rossato *et al.*, 2003] pour le même sujet en dynamique (EMA), on distingue quelques groupes de phonèmes qui se détachent et peuvent être classés selon leur degré de fermeture du port vélopharyngé : les voyelles nasales ont une plus grande ouverture que les voyelles et les consonnes orales ; les consonnes nasales apparaissent par contre mélangées aux voyelles nasales. Aucune logique claire ne permet de distinguer les consonnes orales entre elles. Les voyelles orales semblent quant à elles également mélangées ; seule la voyelle basse /a/ se distingue par son ouverture plus importante que les autres (en cohérence avec [Rossato *et al.*, 2003]). Au niveau des voyelles nasales, on constate une plus grande ouverture pour la voyelle /ã/, suivie ensuite des voyelles /ẽ/ et /œ/ puis de la voyelle /õ/ qui a la plus faible ouverture. Ceci apparaît cohérent avec les observations habituelles que l'on peut trouver dans la littérature qui classent systématiquement /õ/ comme la plus fermée des quatre et généralement /ã/ comme la plus ouverte : [Delvaux *et al.*, 2002b] sur des enregistrements IRM temps-réels, [Amelot *et al.*, 2002] sur des enregistrements fibroscopiques dynamiques, [Demolin *et al.*, 2003] sur des enregistrements IRM statiques et [Rossato *et al.*, 2003] sur des enregistrements EMA dynamiques du même sujet. La plus faible ouverture pour l'articulation /õ/ est très probablement due en partie

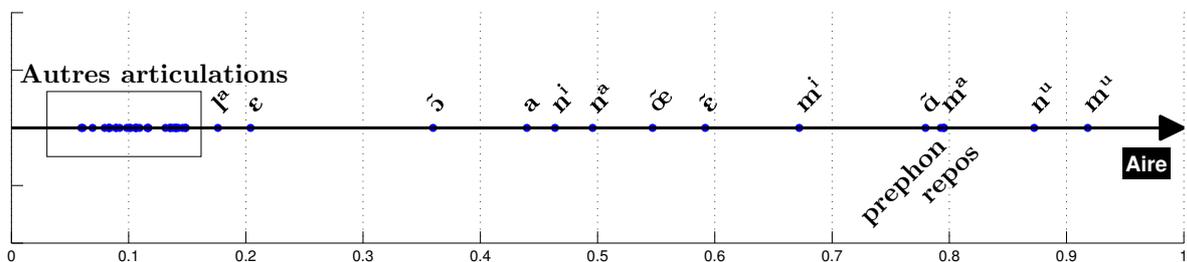


FIG. 4.10 – Répartition des 46 phonèmes artificiellement soutenus du corpus en fonction de l'aire de couplage du port vélopharyngé.

au recul du voile du palais provoqué par une langue très en arrière.

Nous étudions dans la suite plus en détail la fonction d'aire du port vélopharyngé, ainsi que l'effet des deux prédicteurs VL et VS sur celle-ci ; la fonction d'aire orale est étudiée ensuite.

4.3.1 Effets de VL sur la fonction d'aire

Nous avons calculé la fonction d'aire du port vélopharyngé pour des variations linéaires de VL entre -3 et +3. L'évolution de cette dernière est visible sur la figure 4.11. Conformément aux attentes, les variations de VL induisent des variations importantes de la fonction d'aire : le volume du port vélopharyngé du point de séparation du conduit oral au point de séparation en deux choanes varie de 9.70 cm^3 pour un voile du palais maximalement bas dans les données à 4.25 cm^3 à la fermeture. Lorsque le port vélopharyngé se ferme, le voile du palais remonte selon un axe oblique de 45° environ : la fonction d'aire suit logiquement ce mouvement avec une réduction de l'aire qui s'accompagne d'un déplacement en direction du cavum (figure 4.11, droite). L'allure générale de la courbe de la fonction d'aire du port vélopharyngé entre le point de départ du conduit nasal et le cavum est conforme aux observations des aires transverses faites par [Demolin *et al.*, 2003] sur quatre sujets francophones. L'aire est minimale au début du conduit nasal puis augmente progressivement en direction du cavum, jusqu'à obtenir une aire à peu près stable, autour de 2.5 cm^2 , juste avant la séparation du conduit en deux choanes. Notons que cet élargissement progressif de l'embranchement du conduit nasal jusqu'au cavum n'est pas tout à fait linéaire, comme supposé dans la littérature ([Maeda, 1982], [Feng et Castelli, 1996]), mais suit une courbe légèrement convexe, particulièrement pour des positions basses du voile du palais.

On observe sur la figure 4.11 que le point le plus bas du conduit nasal en direction de la

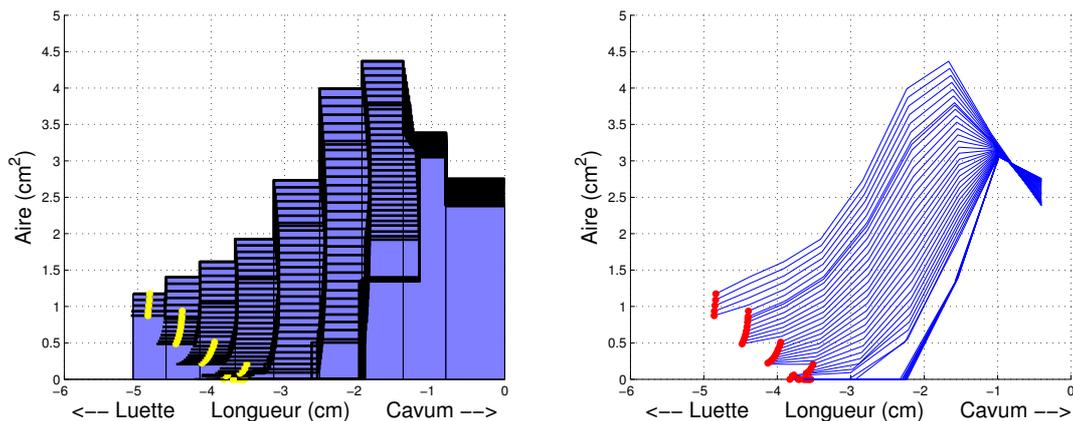


FIG. 4.11 – Variations de la fonction d'aire du port vélopharyngé pour VL évoluant linéairement entre -3 et +3 ; à droite, les milieux des tubes sont reliés entre eux par un trait. Les points des deux figures représentent la constriction. L'origine est arbitrairement fixée au point de séparation du conduit nasal en deux choanes.

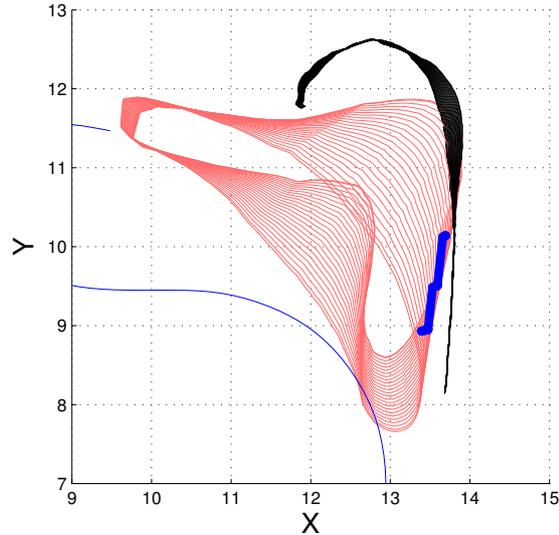


FIG. 4.12 – Représentation de la position du point de séparation du conduit nasal dans le plan médiosagittal pour des valeurs de VL balayant le domaine de variation constaté dans les données.

glotte se déplace en fonction de VL. Ceci signifie en d'autres termes que le point de couplage du conduit nasal avec le conduit oral se déplace avec les mouvements du voile du palais en fonction de VL : lorsque le voile du palais s'abaisse, le point de couplage se rapproche de la glotte. Nous avons représenté ce point de couplage dans le plan médiosagittal pour situer celui-ci par rapport à la pointe du voile du palais (figure 4.12). On observe que le conduit nasal se sépare clairement au-dessus de la pointe de la luvette, ce qui confirme bien qu'une partie de la luvette se situe au milieu du conduit oral et que l'air à ce niveau s'écoule de part et d'autre de celle-ci.

La *constriction* est une notion importante pour la propagation des ondes acoustiques dans le conduit vocal. Elle désigne le point du conduit avec une aire minimum. Comme nous l'avons déjà mentionné, dans le cas du conduit nasal, l'aire de constriction correspond également à l'aire de couplage du conduit nasal avec le conduit oral. C'est sous cette dénomination que nous parlerons de l'aire de constriction du port vélopharyngé dans la suite. L'aire de couplage est déterminée comme le milieu du tube du port vélopharyngé ayant une aire minimale. Les points de la figure 4.11 montrent l'évolution de l'aire de couplage ainsi définie sous l'action de VL. Le point de constriction est localisée non loin de la limite inférieure du port vélopharyngé et suit l'évolution globale de la fonction d'aire : lorsque le port vélopharyngé se ferme, le voile du palais remonte, et la constriction diminue et remonte en direction du cavum dans le même temps. La figure 4.13 détaille l'évolution du point de constriction sous l'action de VL : d'une part la valeur de l'aire de couplage en fonction de VL (droite), et d'autre part la variation mutuelle de l'aire de couplage et de la position de cette constriction avec VL (gauche). L'axe des abscisses est croissant de droite à gauche et l'axe des ordonnées de bas en haut pour *imiter* l'évolution du port vélopharyngé : lorsque le voile du palais remonte, l'aire de couplage du conduit nasal

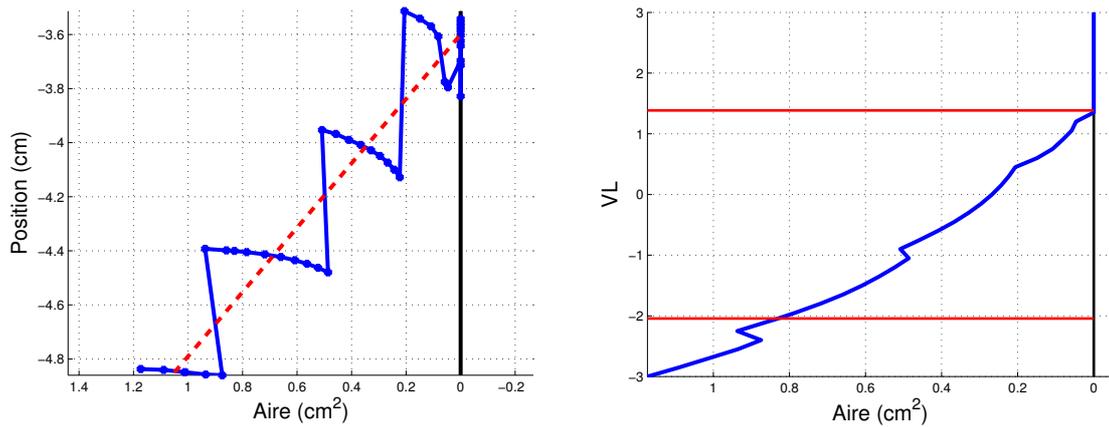


FIG. 4.13 – Variations de l'aire de couplage du port vélopharyngé pour VL évoluant linéairement entre -3 et +3. À gauche, l'aire de la constriction en fonction de sa position par rapport à l'origine (point de séparation du conduit nasal en deux choanes); le trait pointillé montre la direction générale d'évolution de la constriction. À droite, l'aire de couplage en fonction de la valeur de VL; les deux traits horizontaux marquent le domaine de variation de VL constaté dans les données.

diminue; le voile du palais en fin de trajectoire vient buter contre la paroi du pharynx à la fermeture, symbolisée par un trait plein vertical sur la droite. Conformément aux descriptions précédentes, on observe une remontée de la position de l'aire de couplage avec celle du voile du palais en même temps qu'une diminution de son aire. La direction globale est donnée par l'axe en pointillé sur la figure 4.13 (gauche). L'aspect en escalier est dû à l'échantillonnage du port vélopharyngé en plans de grilles : une augmentation de la densité des plans de grille (neuf dans notre cas dans la région vélaire, voir la figure 4.6) réduirait la taille des marches mais ne changerait pas l'orientation globale. La figure droite montre une relation presque linéaire de l'aire de couplage en fonction de VL. On note au passage une variation de cette aire dans les données par l'intermédiaire du modèle de 0.8 cm^2 pour un voile du palais maximale bas à 0 cm^2 à la fermeture, cohérent avec les résultats de [Demolin *et al.*, 2003] : ils trouvent une aire transverse minimale variant de 0.66 cm^2 à 1.93 cm^2 suivant les sujets, pour un voile du palais maximale bas. [Björk *et al.*, 1961] proposent par contre une aire possible jusqu'à 2.5 cm^2 à partir de mesures effectuées sur des images rayons-X. La variation de 0.8 cm^2 dans nos données correspond à un déplacement vertical de 0.9 cm environ dans notre modèle.

La relation presque linéaire entre VL et l'aire de couplage du port vélopharyngé, qui mesure implicitement son ouverture, incite à considérer VL comme une mesure du *degré d'ouverture* du port vélopharyngé. On observe que ce paramètre suffit pour passer d'une fermeture du conduit nasal à une ouverture maximale. On note cependant sur la figure 4.13 (droite) que la fermeture complète du port vélopharyngé au sens strict du terme, c'est-à-dire l'aire de couplage nulle, n'est presque jamais atteinte, ce qui ne veut pas dire pour autant qu'il y ait effectivement nasalité. Une comparaison entre les aires de couplage des données et celles estimées par le modèle avec

le seul paramètre VL a d'ailleurs révélé une erreur RMS de 0.15 cm^2 sur les 46 articulations, ce qui donne un ordre de grandeur de précision du modèle pour l'aire de couplage.

Notons enfin comme l'ont notamment souligné [Demolin *et al.*, 2003], que l'aire nasale doit également être étudiée en parallèle de l'aire orale dans cette région pour calculer le degré de couplage du conduit nasal (voir la section 4.5). Les aires de couplage présentées ci-dessus sont en effet très dépendantes du sujet considéré.

4.3.2 Effets de VS sur la fonction d'aire

Les analyses faites précédemment pour le paramètre VL sont dupliquées ici pour le paramètre VS. Les figures 4.14, 4.15 et 4.16 pour VS sont l'équivalent des figures obtenues pour VL. On observe logiquement de plus petites variations de la fonction d'aire que pour VL. Contrairement aux effets de VL, et particulièrement pour l'aire de couplage, la réduction des aires s'accompagne d'un petit mouvement descendant général conforme au mouvement global du voile du palais sous l'action de VS (figure 4.14). Une réduction de l'aire de couplage s'accompagne donc d'un allongement de la fonction d'aire du port vélopharyngé.

On observe sur la figure 4.15 que le paramètre VS a également un petit effet sur le point de couplage du conduit nasal avec le conduit oral : ce point descend légèrement vers la glotte lorsque le voile du palais recule et reste assez au-dessus de la pointe du voile du palais dans le plan médiosagittal.

La constriction vélopharyngée se situe également dans la partie inférieure, non loin de la connexion au conduit oral. Conformément à la description faite ci-dessus, une baisse de l'aire de couplage s'accompagne d'une baisse de sa position dans le port vélopharyngé, selon l'axe en

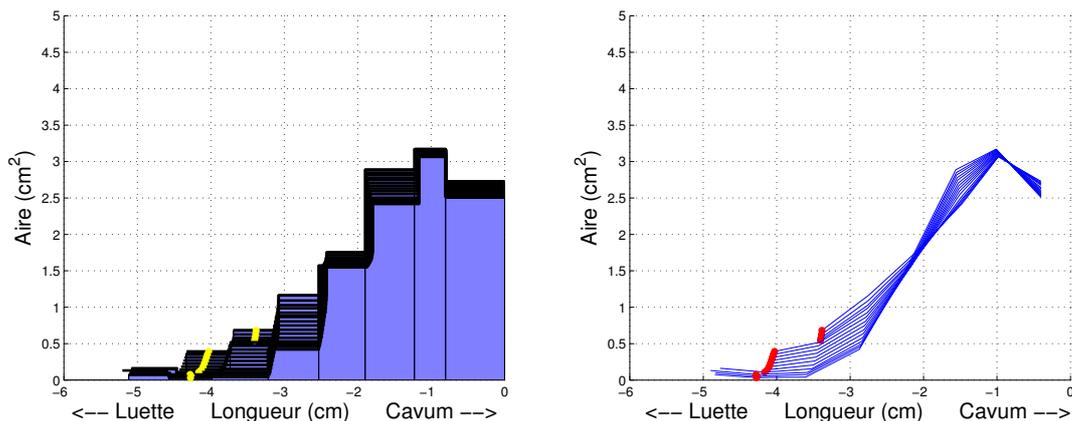


FIG. 4.14 – Variations de la fonction d'aire du port vélopharyngé pour VS évoluant linéairement entre -3 et +3 ; à droite, les milieux des tubes sont reliés entre eux par un trait. Les points des deux figures représentent la constriction. L'origine est arbitrairement fixée au point de séparation du conduit nasal en deux choanes.

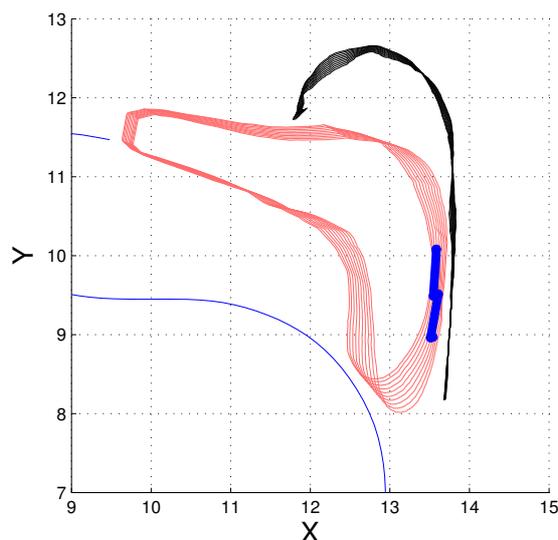


FIG. 4.15 – Représentation de la position du point de séparation du conduit nasal dans le plan médiosagittal pour des valeurs de VS balayant le domaine de variation constaté dans les données.

pointillés de la figure 4.16 (gauche). L'amplitude globale de cette variation verticale du point de constriction est de 0.6 cm pour la variation de VS maximale trouvée dans les données. L'aire de couplage elle-même, visible sur la figure 4.16 (droite), suit une relation presque linéaire avec VS. Dans la limite des variations de VS observées dans les données (schématisée par les deux traits horizontaux), l'aire de couplage varie fortement : elle passe d'une valeur inférieure à 0.1 cm^2 (presque une fermeture) à une valeur proche de 0.6 cm^2 , enregistrant une variation d'environ 0.5 cm^2 , soit une multiplication par 5. Par ailleurs, l'ajout de ce paramètre au modèle réduit significativement l'erreur RMS entre aire de couplage estimée et réelle de 0.15 cm^2 à 0.1 cm^2 . Il est à noter cependant qu'aucune explication phonétique n'a été trouvée à la répartition des 46 phonèmes selon ce paramètre.

Nous avons présenté dans cette section de nombreux résultats sur la fonction d'aire vélopharyngée, par définition difficile à estimer. Néanmoins, si les mouvements du voile du palais entraînent des variations du conduit nasal comme nous venons de le voir, ils entraînent également des variations sur le conduit oral qu'il convient de prendre en compte. La section suivante s'attache à mesurer les effets de chacun des deux paramètres de contrôle du modèle sur la fonction d'aire orale avant d'étudier les covariations des deux fonctions d'aire dans la section qui suivra.

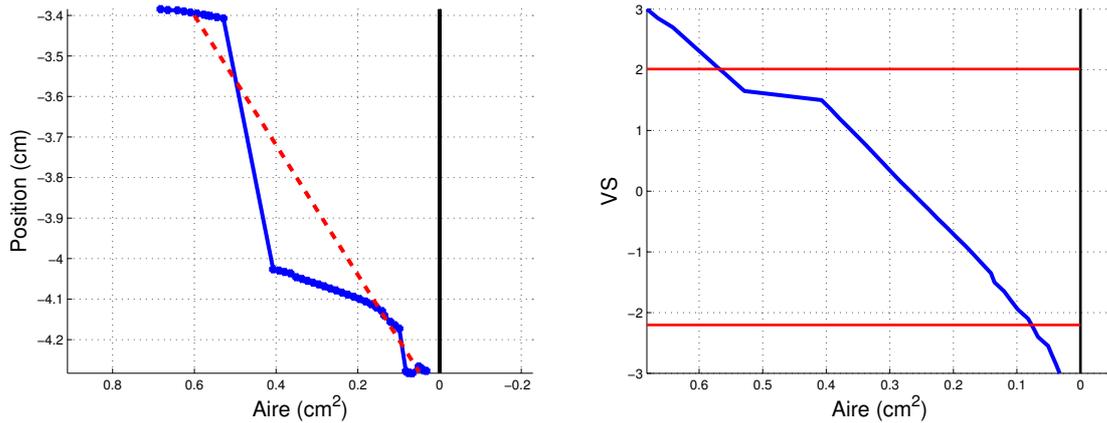


FIG. 4.16 – Variations de l'aire de couplage du port vélopharyngé pour VS évoluant linéairement entre -3 et +3. À gauche, l'aire de la constriction en fonction de sa position par rapport à l'origine (point de séparation du conduit nasal en deux choanes) ; le trait pointillé montre la direction générale d'évolution de la constriction. À droite, l'aire de couplage en fonction de la valeur de VS ; les deux traits horizontaux marquent le domaine de variation de VS constaté dans les données.

4.4 Influence de la position du voile du palais sur le conduit oral

Les effets acoustiques de la nasalité sont souvent considérés sous l'angle du couplage entre deux conduits. Les mouvements du voile du palais qui permettent ce couplage par l'ouverture du port vélopharyngé modifient par la même occasion la géométrie du conduit oral. Ces modifications peuvent avoir des conséquences acoustiques importantes (voir par exemple [Feng et Kotenkoff, 2004]). Le modèle développé ici permet d'étudier précisément l'influence de la position du voile du palais sur les fonctions d'aire orales. Pour cela, nous avons calculé les aires orales des sections transverses dans la région vélaire pour des variations de la position du voile du palais correspondant à des variations de VL et VS. Afin d'inscrire ces variations de la région vélaire au coeur du conduit oral complet de la glotte aux lèvres, nous avons également déterminé pour quelques articulations données les sections transverses fixes des régions pharyngée et buccale. Nous les avons extraites manuellement à partir des images IRM transverses du conduit vocal dont l'orientation est visible sur la figure 2.8 page 50. Nous n'avons traité que quelques cibles représentatives des articulations possibles du conduit vocal : les trois articulations extrêmes du triangle vocalique /a/, /i/ et /u/ et les quatre articulations des voyelles nasales /ã/, /õ/, /ẽ/ et /œ/. Puisque nous ne possédons pas encore de modèle articulatoire complet, les aires transverses du conduit oral dans la région vélaire ont été calculées avec un maillage de langue dans une position fixe pour une articulation donnée. La position du maillage du voile du palais est quant à elle variable selon les paramètres VL et VS. Les sections transverses orales ainsi obtenues dans la région vélaire sont intercalées à leur place parmi les sections transverses complètes du conduit oral. La fonction d'aire complète de la glotte aux lèvres est ensuite calculée suivant le protocole de la partie 4.1 pour chacune des sept articulations concernées. Les figures 4.17 et 4.18 montrent

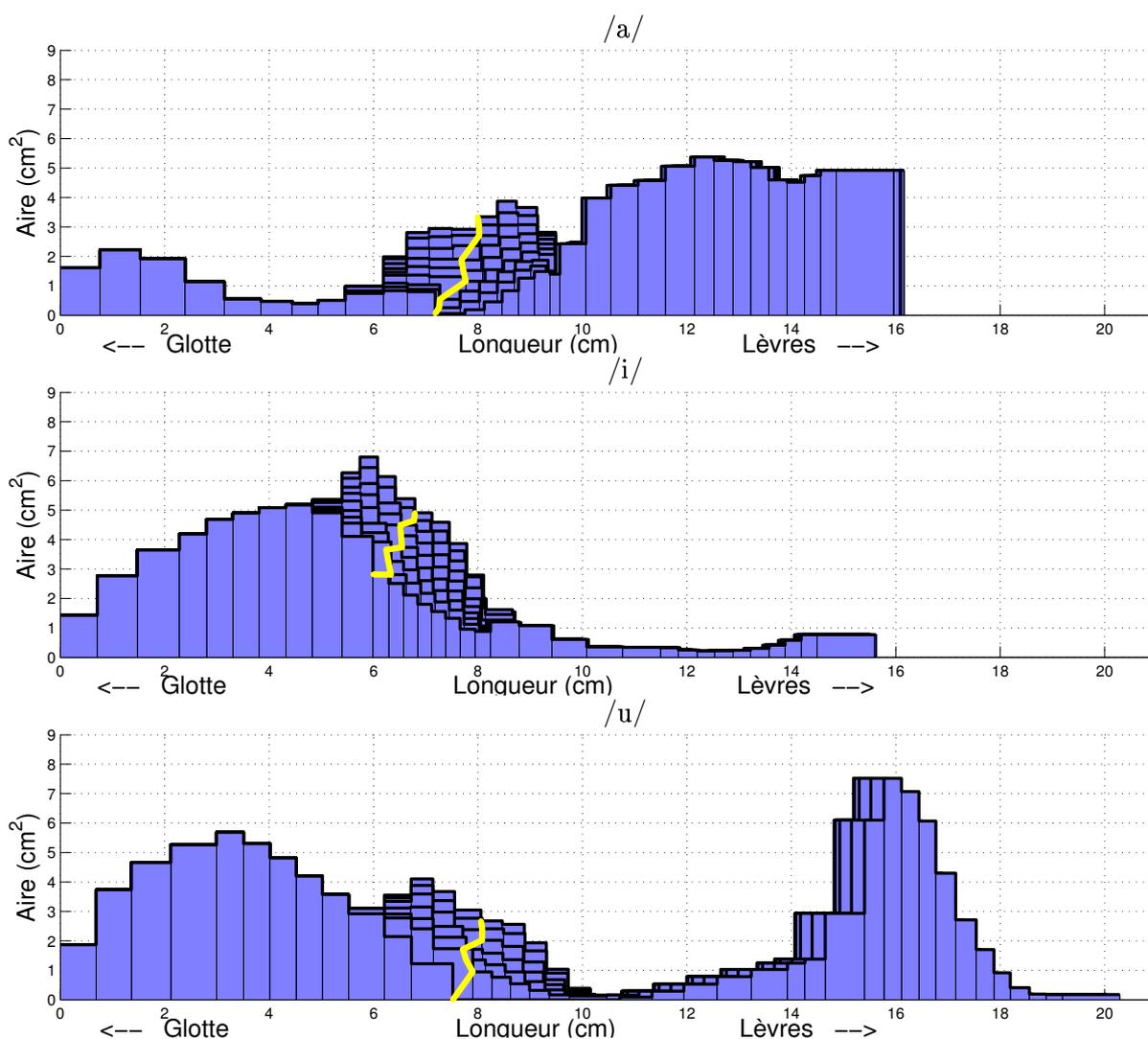


FIG. 4.17 – Fonctions d’aire du conduit oral des articulations /a/, /i/ et /u/ pour des variations du voile du palais selon le paramètre VL compris entre les valeurs extrêmes trouvées dans les données. L’origine est arbitrairement fixée à la glotte. La ligne claire au milieu des fonctions d’aires indique la position de branchement du conduit nasal.

les variations des fonctions d’aire mesurées pour les trois articulations /a/, /i/ et /u/ pour des mouvements du voile du palais associés aux paramètres VL et VS du modèle.

On observe naturellement une forte variation de l’aire dans la région vélaire pour des variations liées à VL, pouvant aller jusqu’à l’occlusion dans le cas du /a/ et du /u/ pour le voile du palais en position maximale basse (également dans le cas du /ã/ et du /õ/ non représentés ici). Afin de donner un ordre de grandeur des fortes variations observées pour les sept articulations analysées, on mesure pour chacune des articulations une variation d’aire supérieure à 3 cm² pour le tube qui enregistre la plus forte variation dans la région vélaire. Le second point à noter sur la figure 4.17 est l’étendue des modifications introduites par la variation du voile du

palais seul. On constate en effet que la fonction d'aire varie en moyenne sur les sept articulation sur une longueur de 4.4 cm. Il est par ailleurs parfois considéré dans la littérature que l'influence des mouvements du voile du palais sur la fonction d'aire orale ne se produit qu'après la point de séparation du conduit en deux. On observe cependant que la connexion du conduit nasal comme défini à la section 4.1 se trouve environ *au milieu* de la zone de variations de la fonction d'aire orale. Ceci confirme que la position du voile du palais, et de la luette en particulier, joue un rôle important pour la géométrie du conduit oral en amont de la séparation effective du conduit nasal.

Le second paramètre VS modifie les fonctions d'aire du conduit oral d'une amplitude plus faible, comme le montre la figure 4.18 pour les articulations /a/, /i/ et /u/. Une variation de la

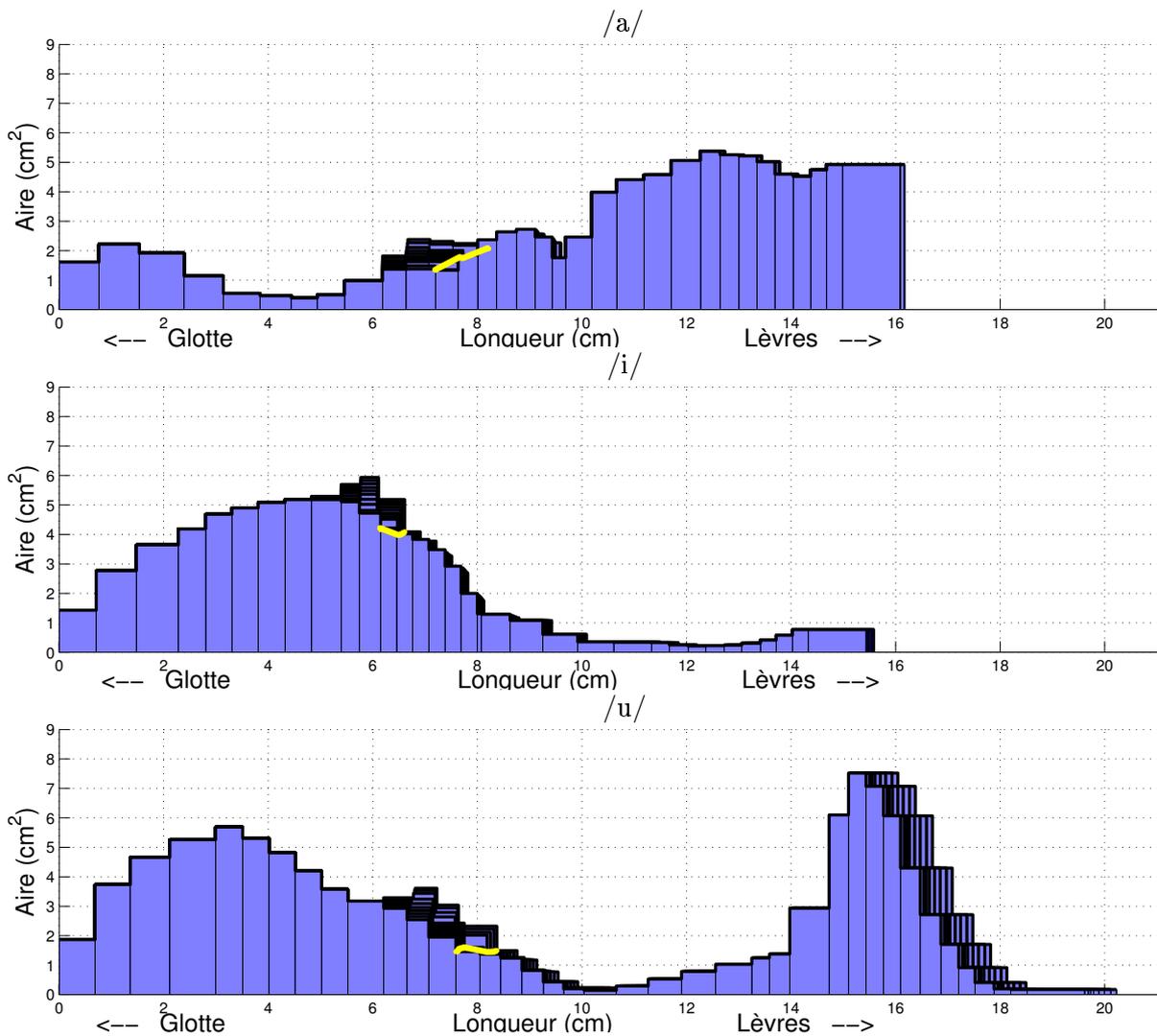


FIG. 4.18 – Fonctions d'aire du conduit oral des articulations /a/, /i/ et /u/ pour des variations du voile du palais selon le paramètre VS compris entre les valeurs extrêmes trouvées dans les données. L'origine est arbitrairement fixée à la glotte. La ligne claire au milieu des fonctions d'aires indique la position de branchement du conduit nasal.

forme du voile du palais selon le paramètre VS peut cependant modifier certaines aires jusqu'à plus d'1 cm², soit presque 50% de l'aire du tube concerné pour certaines articulations. L'amplitude de variation constatée ne peut pas conduire à une occlusion du conduit oral dans la région vélaire.

Nous avons présenté dans cette section l'influence des mouvements du voile du palais sur la fonction d'aire orale seule. De fortes variations apparaissent en fonction des mouvements du voile du palais, en particulier suivant la valeur de VL, quelle que soit la voyelle considérée. On observe donc que les mouvements du voile du palais font varier de manière importante à la fois les fonctions d'aire orale et nasale. Ces deux variations sont parfois considérées dans la littérature comme complémentaires pour garder une aire totale fixe ([Maeda, 1982]). Nous décidons d'étudier ce point dans la section suivante à la lumière du modèle articulatoire.

4.5 La covariation des aires de couplage orale et nasale

La proportion de couplage nasal peut être définie comme le rapport entre l'aire de couplage du conduit nasal sur la somme des aires de couplage des conduits oral plus nasal. Le rôle de cette valeur est une question importante dans l'étude de la nasalité. Comme nous l'avons déjà largement souligné dans ce mémoire, la difficulté à mesurer les dimensions du port vélopharyngé a conduit à la formulation d'hypothèses plus ou moins réalistes pour simuler cette proportion de couplage. Il est logique, et les analyses précédentes le confirment, que l'abaissement du voile du palais provoque conjointement une ouverture du port vélopharyngé et un rétrécissement de la cavité orale. La section 4.3 a permis de montrer une relation quasi-linéaire entre la valeur du paramètre VL, qui représente la hauteur du voile du palais, et l'aire de couplage vélopharyngé. Nous avons souligné dans la section 4.4 deux effets importants de la hauteur du voile du palais sur les fonctions d'aire :

- Le point de séparation du conduit vocal en deux conduits, l'un buccal et l'autre nasal, descend d'environ 1 cm quand le voile du palais passe d'une position maximale haute à une position maximale basse (voir la figure 4.17).
- La position du voile du palais influence fortement la fonction d'aire orale en amont de la séparation effective du conduit oral en deux.

Cette section s'attache à mesurer la variation conjointe des aires orales et nasales à partir de la séparation du conduit en deux en fonction de la hauteur du voile du palais.

On s'intéresse en premier lieu à la proportion de couplage du conduit nasal avec le conduit oral, c'est-à-dire la proportion d'aire nasale par rapport à l'aire totale au point de séparation du conduit. La figure 4.19 montre les variations conjointes des aires orales et nasales du premier tube après la séparation du conduit en deux en fonction de VL, qui contrôle la hauteur du voile du palais, pour les sept articulations /a, i, u, ā, ã, õ, ē/. On observe logiquement une diminution de l'aire nasale parallèlement à une augmentation des aires orales pour les sept articulations lorsque le voile du palais remonte (c'est-à-dire que la valeur de VL augmente). L'aire orale croît clairement plus vite que l'aire nasale ne décroît. [Maeda, 1982] et [Feng et Castelli, 1996] sup-

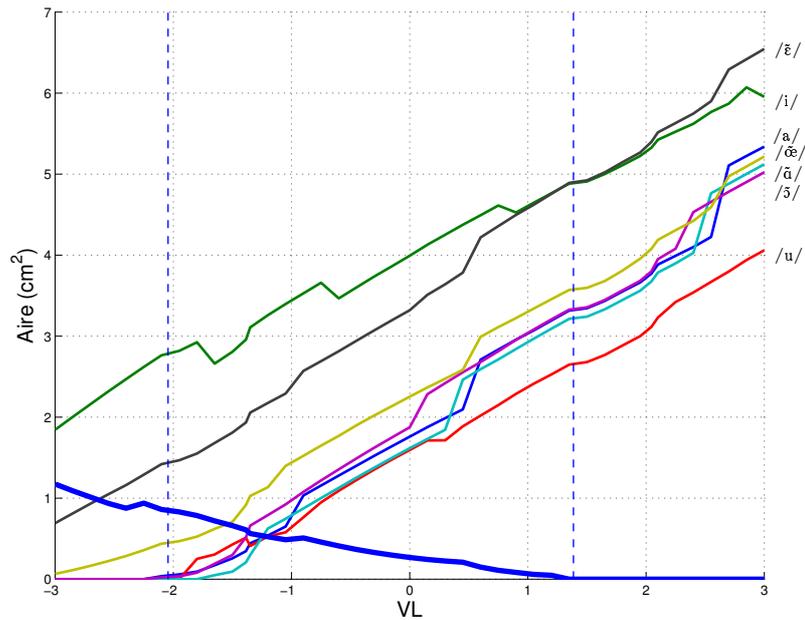


FIG. 4.19 – Aires orales (courbes croissantes) et nasale (courbe décroissante) du premier tube de la fonction d'aire de chacun des deux tuyaux après la séparation effective du conduit nasal en fonction de VL pour les sept articulations indiquées. Les lignes verticales délimitent les valeurs minimale et maximale de VL trouvées dans les données.

posent pour leurs simulations que l'aire totale du conduit est constante et que le voile du palais passe d'une position haute où la proportion d'aire nasale est de 0% à une position basse où la même proportion est de 100% avec une somme des aires orale et nasale toujours constante. La figure 4.20 affiche la somme des aires orale et nasale en fonction de la hauteur du voile du palais pour les sept articulations étudiées. Cette figure confirme que l'aire orale augmente plus vite que l'aire nasale ne diminue au point de séparation des deux tuyaux lorsque le voile du palais remonte. On note même une relation presque linéaire entre l'augmentation de l'aire totale et le paramètre VL. Pour des valeurs de VL inférieure à -2, et à l'exception des voyelles avant /i/ et /ɛ/ et dans une moindre mesure de la voyelle /œ/, le voile du palais provoque une occlusion du conduit oral, dont l'aire est alors nulle (voir la figure 4.19), ce qui conduit à une augmentation de l'aire totale, limitée à l'aire nasale, lorsque le voile du palais descend. Le comportement global observé sur la figure est toujours le même quelle que soit la voyelle considérée : les voyelles sont uniquement différenciées par un décalage des aires orales, avec des aires logiquement plus importantes pour les voyelles avant et moins importantes pour les voyelles arrières.

Le second point étudié, qui est une généralisation du premier, est la covariation des aires orales et nasales sur les tubes qui suivent la séparation du conduit en deux. [Maeda, 1982] et [Feng et Castelli, 1996] soustraient l'aire des tubes du port vélopharyngé à celles des tubes équivalents oraux de manière à conserver une aire totale orale plus nasale constante pour chaque

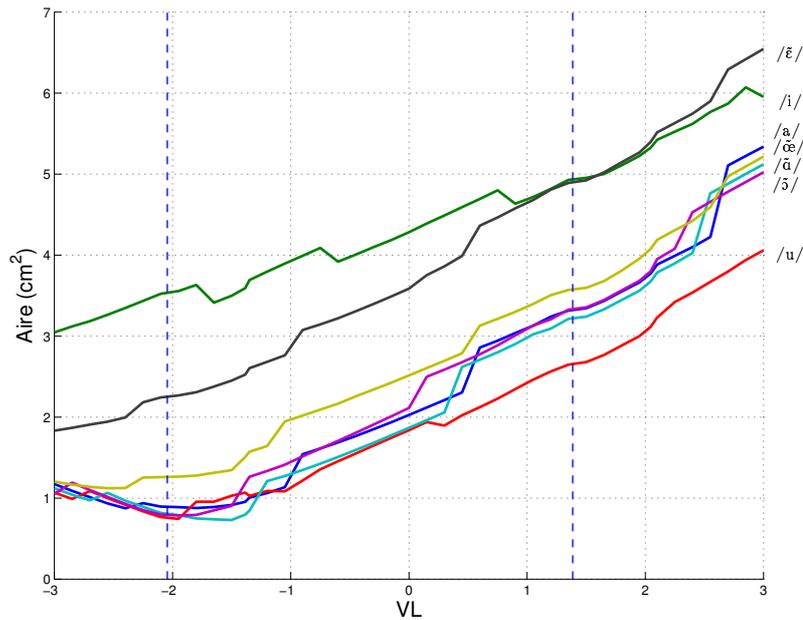


FIG. 4.20 – Somme de l'aire orale et nasale du premier tube de la fonction d'aire de chacun des deux tuyaux après la séparation effective du conduit nasal en fonction de VL pour chacune des sept articulations indiquées. Les lignes verticales délimitent les valeurs minimale et maximale de VL trouvées dans les données.

tube quelle que soit la position du voile du palais. Notons tout de suite qu'il n'y a en réalité pas d'équivalence fixe entre les tubes du port vélopharyngé et ceux de la région orale : le point de connexion du conduit nasal dans le conduit oral dépendant de la hauteur du voile du palais, le premier tube vélopharyngé n'est pas toujours l'*équivalent* nasal du même tube oral. Il en est naturellement de même pour les tubes suivants. Pour pouvoir comparer par paires les variations des aires des tubes des conduit oral et nasal, nous choisissons comme référence le point de séparation du conduit en deux. Ainsi, quelle que soit la hauteur du voile du palais, nous associons les premiers tubes oral et nasal après la séparation, les seconds tubes, les troisièmes tubes, etc. de manière à comparer leurs variations conjointes en fonction de VL. On est donc amené à changer de tube oral pour suivre l'évolution du premier tube nasal en fonction de la hauteur du voile du palais. Cette approche revient en quelque sorte à faire abstraction du déplacement du point de couplage du conduit nasal avec le conduit oral en fonction du mouvement du voile du palais.

La somme des aires orales et nasales sur les cinq tubes qui suivent le point de séparation du conduit vocal en fonction du paramètre VL qui représente la hauteur du voile du palais est affichée sur la figure 4.21 pour les sept articulations étudiées. On constate que quelle que soit la voyelle considérée, plus on est proche du point de séparation du conduit, plus l'augmentation de l'aire orale est forte par rapport à la diminution de l'aire nasale lorsque le voile du palais remonte.

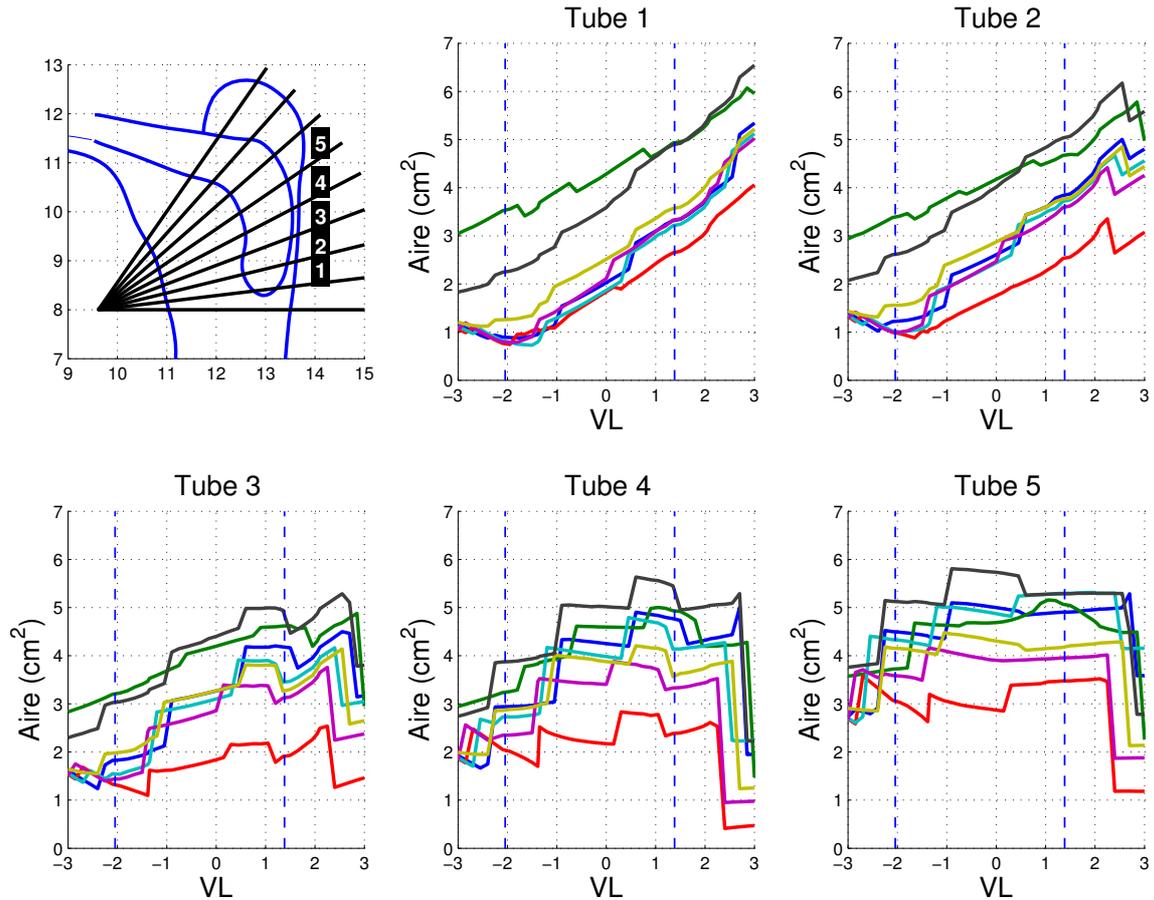


FIG. 4.21 – Somme des aires orale et nasale pour chaque tube numéroté 1 à 5 à partir de la séparation du conduit en deux (figure en haut à gauche) de la fonction d'aire des sept articulations /a, i, u, \tilde{a} , \tilde{i} , \tilde{o} , \tilde{e} / en fonction de VL. Les lignes verticales délimitent les valeurs minimale et maximale de VL trouvées dans les données.

En d'autres termes, l'aire orale augmente plus vite que l'aire nasale dans la région proche du point de séparation lorsque le voile du palais remonte. Le tube numéroté cinq sur la figure 4.21 présente une aire cumulée orale-nasale à peu près constante lorsque le voile du palais remonte ; ceci signifie que la croissance de l'aire orale et la décroissance de l'aire nasale sont du même ordre de grandeur. Ce phénomène peut s'expliquer en partie par la taille des deux tuyaux oral et nasal comme le schématise la figure 4.22 : dans la région inférieure du port vélopharyngé, le tuyau nasal étant généralement de taille inférieure au tuyau oral, une variation de la position du voile du palais induit une variation de l'aire du conduit oral beaucoup plus grande (figure 4.22, gauche). En s'éloignant du point de séparation du conduit en deux, la taille du conduit nasal augmente régulièrement jusqu'à la région du cavum pour atteindre une taille plus comparable à celle du conduit oral et enregistrer des variations d'aires plus proches (figure 4.22, droite), et donc une variation de même ordre de grandeur pour les deux conduits lorsque le voile du palais bouge. Notons que l'aspect irrégulier des courbes présentées à la figure 4.21 est dû simplement au fait qu'à un tube nasal numéroté entre un et cinq correspond une succession de différents

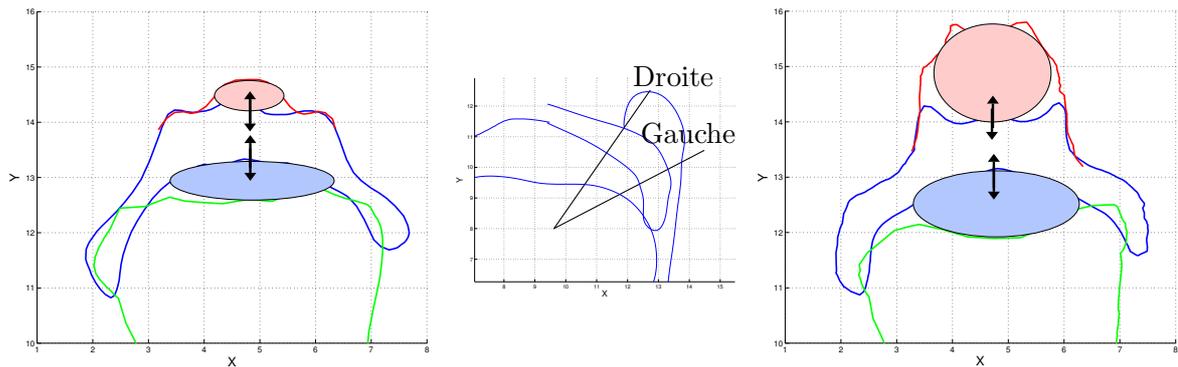


FIG. 4.22 – Sections transverses (à gauche et à droite) pour les deux plans de grille de la figure médiosagittale centrale. Les aires transverses sont schématisées par des ellipses et les mouvements du voile du palais selon VL par des flèches : un même mouvement du voile du palais fait varier d'un ordre de grandeur différent les aires orale et nasale à gauche et d'un même ordre de grandeur à droite.

tubes oraux pour garder constante la distance au point de séparation du conduit.

En résumé, l'étude de la covariation des fonctions d'aire orale et nasale en fonction de la variation de la position du voile du palais a permis de souligner ou rappeler plusieurs points importants indépendants de la voyelle considérée :

1. La fonction d'aire orale est fortement influencée par la position du voile du palais dans la région vélaire, y compris en amont de la séparation effective du conduit vocal en deux conduits oral et nasal.
2. La position du point de séparation du conduit dépend de la position du voile du palais et peut varier sur une longueur d'environ un centimètre.
3. Au niveau du couplage du conduit nasal avec le conduit oral, on observe une variation plus importante de l'aire orale que de l'aire nasale lorsque le voile du palais monte et descend, sauf naturellement en cas d'occlusion de l'un ou l'autre des conduits.
4. Ce phénomène se retrouve globalement dans les tubes en aval de la séparation du conduit en deux et va en s'amenuisant jusqu'à obtenir une variation similaire pour les deux conduits en se rapprochant du palais dur.

4.6 Conclusion

Dans le but de modéliser la propagation plane des ondes acoustiques dans le conduit vocal, nous avons déterminé dans ce chapitre les fonctions d'aire des conduits oral et nasal à partir du modèle articulatoire tridimensionnel. Nous avons proposé pour cela une méthode de calcul des fonctions d'aire à partir des sections transverses des conduits, obtenues soit par recoupes

des maillages du voile du palais, de la paroi du pharynx et de la langue dans la région vélaire, soit directement par le tracé des contours du conduit dans les images IRM transverses dans les autres régions. L'étude des variations de la fonction d'aire du port vélopharyngé en fonction des paramètres de contrôle VL et VS a permis de caractériser les mouvements du voile du palais. Les variations du paramètre VL sont à l'origine de fortes variations de la fonction d'aire : l'aire de couplage du conduit nasal varie de 0 cm^2 à 0.8 cm^2 , aire qui correspond à la valeur maximale de VL trouvée dans les données. Les variations de VL seul suffisent à passer d'une configuration où le conduit nasal est fermé à une configuration où le port vélopharyngé est maximalelement ouvert. L'influence du voile du palais sur l'aire du conduit oral apparaît également très importante. Les variations du paramètre VS correspondent plutôt à de petits mouvements du voile du palais qui entraînent des variations importantes de l'aire de couplage du conduit nasal : celle-ci varie entre 0.1 cm^2 et 0.6 cm^2 . Son influence sur le conduit oral est plus faible que VL mais n'apparaît cependant pas négligeable. Nous avons par ailleurs étudié la covariation des aires de couplage orale et nasale en fonction du paramètre VL. Au point de séparation du conduit en deux, il apparaît que lorsque le voile du palais varie selon le paramètre VL, l'aire de couplage du conduit oral varie plus fortement que l'aire du conduit nasal. Cette tendance va en diminuant au fur et à mesure que l'on s'éloigne du point de séparation du conduit. Enfin, nous avons proposé de nouvelles fonctions d'aire des fosses nasales qui s'avèrent cohérentes avec la littérature récente.

Le calcul des fonctions d'aire permet de déterminer les fonctions de transfert acoustiques du conduit vocal en considérant un mode de propagation des ondes plan. Les résultats présentés dans ce chapitre permettent alors de faire le lien entre les mouvements articulatoires et le signal acoustique. Le chapitre suivant vise à caractériser acoustiquement les mouvements articulatoires obtenus à l'aide du modèle.

Chapitre 5

Modélisation acoustique

La recherche des corrélats acoustiques a longterm constitué un enjeu important pour l'étude de la nasalité (voir par exemple [Lonchamp, 1988] pour une revue détaillée). La simulation acoustique par l'intermédiaire d'une modélisation articulatoire permet d'obtenir des fonctions de transfert acoustiques réalistes. La caractérisation acoustique d'un geste articulatoire est donc au cœur de la compréhension du processus de production de la parole. Le modèle articulatoire présenté dans cette étude offre une nouvelle collection de fonctions d'aire du conduit nasal proches des données. Les relations articulatoire-acoustiques des mouvements du port vélopharyngé et du conduit nasal complet peuvent alors être étudiées à la lumière du modèle. Ceci constitue également une validation du modèle et ouvre la voie à la synthèse acoustique des nasales. Les analyses présentées dans ce chapitre constituent ainsi une première tentative pour étudier les effets du couplage du conduit nasal à l'aide du modèle articulatoire présenté dans les chapitres précédents.

5.1 Modèle acoustique

5.1.1 Principes généraux

Le modèle acoustique utilisé dans cette étude s'appuie sur un analogie électrique-acoustique : la propagation acoustique dans le conduit vocal est similaire à la propagation électrique dans une ligne de transmission. Cette section ne contient que les principes généraux et les choix effectués pour notre étude. Plus de détails seront trouvés dans les nombreux travaux de référence dans ce domaine ([Dunn, 1950], [Stevens *et al.*, 1953], [Fant, 1960], [Flanagan, 1972], [Mrayati, 1976], [Badin et Fant, 1984], [Fant, 1985] ou [Stevens, 1998]).

Nous nous intéressons ici au domaine fréquentiel. Nous faisons donc l'hypothèse classique du modèle source-filtre ([Fant, 1960]) : on peut séparer linéairement la *source* d'excitation, à savoir le débit glottique ou les variations de pression à la constriction, et le *filtre* acoustique que constitue le conduit vocal. Nous ferons abstraction de la source en calculant par la suite des *fonctions de transfert* entre les débits acoustiques à la sortie et à l'entrée du conduit vocal.

Un certain nombre d'hypothèses simplificatrices sont introduites pour simuler la propagation des ondes acoustiques dans le conduit vocal. On suppose ainsi que les modes de propagation transversaux pour des fréquences inférieures à 5000 Hz sont faibles et que la propagation des ondes acoustiques dans le conduit vocal est plane. Les équations de propagation acoustique dans le conduit vocal sont alors similaires à celles de propagation électrique dans une ligne de transmission : les pressions et les débits acoustiques correspondent aux tensions et aux courants électriques. La propagation acoustique dans un tube cylindrique de section uniforme A et de longueur constante L peut alors être représentée par la propagation électrique dans un quadripôle comme celui de la figure 5.1. U_1 et U_2 représentent les débits acoustiques à l'entrée et à la sortie et P_1 et P_2 les pressions. Les valeurs des impédances Z_a et Z_b dépendent de la géométrie du tube et des conditions aux limites qui sont imposées.

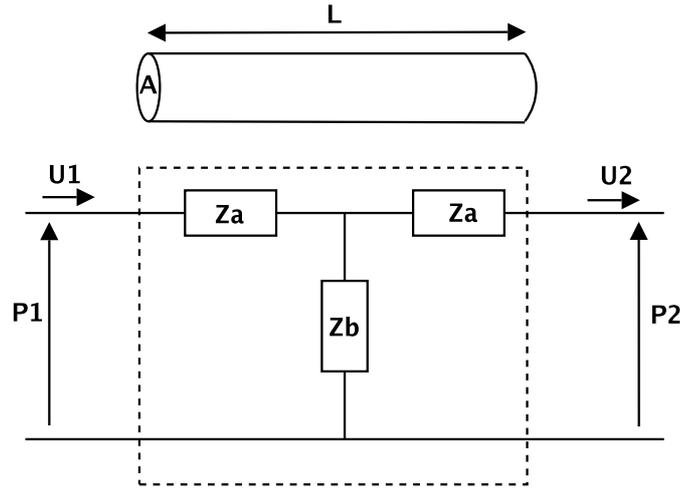


FIG. 5.1 – Représentation d’un tube cylindrique sous forme d’un quadripôle électrique.

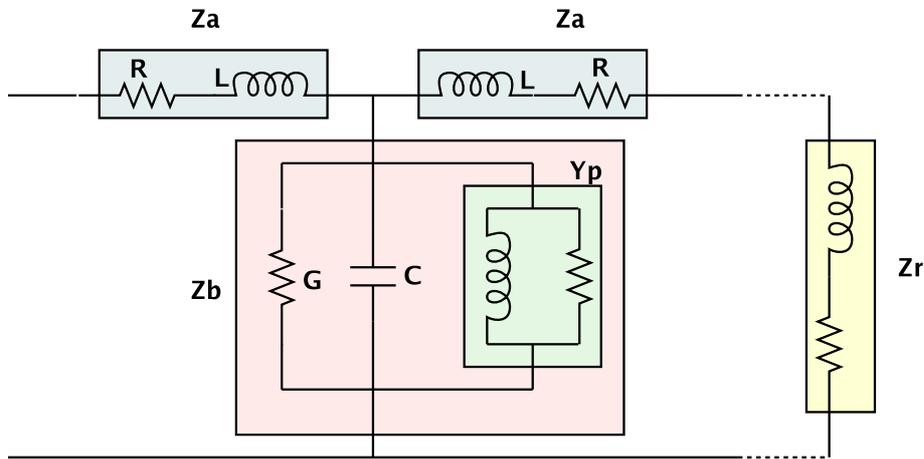


FIG. 5.2 – Représentation du quadripôle électrique équivalent à un tube cylindrique en incluant les pertes. La succession des quadripôles qui représente la succession des tubes cylindrique et donc le conduit vocal est chargée au bout par une impédance de rayonnement.

La forme du quadripôle utilisé pour nos simulations est celui de la figure 5.2. Les bobines L et le condensateur C modélisent les propriétés d’inertie et de compressibilité de l’air dans le tube. La résistance R représente les pertes de surface par viscosité et la conductance G celles par conduction thermique. Enfin, Y_p représente la vibration des parois du tube.

Les valeurs de Z_a et Z_b sont alors les suivantes :

$$Z_a = R + j\omega L = \frac{L}{A} \sqrt{\frac{\pi \rho \mu \omega}{2A}} + j\omega \frac{\rho L}{2A} \tag{5.1}$$

$$Z_b = G + j\omega C + Y_p = 2L\sqrt{\pi A} \frac{\eta - 1}{\rho c^2} \sqrt{\frac{\lambda \omega}{2c_p \rho}} + j\omega \frac{AL}{\rho c^2} + Y_p \tag{5.2}$$

<i>Noms</i>	<i>Symbole</i>	<i>Valeur</i>
Densité	ρ	$1,14 \cdot 10^{-3} \text{ g.cm}^{-3}$
Vitesse du son	c	$3,51 \cdot 10^4 \text{ cm.s}^{-1}$
Viscosité	μ	$1,86 \cdot 10^{-4} \text{ dynes.s.cm}^{-2}$
Constante adiabatique	η	1,4
Coefficient de conductivité de chaleur	λ	$5,5 \cdot 10^{-5} \text{ cal.cm}^{-1} \cdot \text{s}^{-1} \cdot ^\circ\text{C}^{-1}$
Chaleur spécifique pour une pression constante	c_p	$0,24 \text{ cal.g}^{-1} \cdot ^\circ\text{C}^{-1}$

TAB. 5.1 – Constantes de l'air à 0 ° C et 1 atmosphère en unités CGS issues de [Mrayati, 1976].

avec

$$Y_p = 2L\sqrt{\pi A} \frac{b_p}{b_p^2 + m_p^2 + \omega^2} - j\omega 2L\sqrt{\pi A} \frac{m_p}{b_p^2 + m_p^2 + \omega^2}$$

où b_p et m_p sont le coefficient d'amortissement et la masse correspondant aux modèle masse/ressort utilisé pour modéliser la vibration des parois. Ces valeurs sont fixées à 1600 *dynes.s.cm*⁻¹ pour b_p et 1,4 *g.cm*⁻² pour m_p ([Mrayati, 1976]). Les constantes de l'air utilisées pour les équations de Z_a et Z_b sont résumées dans le tableau 5.1.

Le rayonnement dans l'air de l'onde acoustique à la sortie du conduit (c'est-à-dire aux lèvres ou à chacune des deux narines) est modélisé électriquement par une impédance de rayonnement Z_r au bout de la chaîne des quadripôles qui représentent le conduit, comme le montre la figure 5.2. Elle est composée d'une partie résistive, qui représente l'atténuation, en série avec une impédance qui représente l'inertie de l'air à la sortie du tube. Sa valeur est donnée par l'équation :

$$Z_r = \frac{\omega^2 \rho}{2\pi c} + j\omega \frac{8\rho}{3\pi\sqrt{\pi A}} \quad (5.3)$$

Elle représente l'impédance de rayonnement d'un piston vibrant dans un baffle infini.

Le conduit vocal complet de la source glottique aux lèvres est donc modélisé par une succession de quadripôles qui représentent chacun un tube de la fonction d'aire, défini par son aire et sa longueur. La fonction de transfert du conduit vocal, c'est-à-dire le rapport entre le débit acoustique à la sortie et le débit acoustique à la glotte, est donc facilement calculée. Elle dépend uniquement de la fonction d'aire. On obtient ainsi les pôles de la fonction de transfert associée à cette fonction d'aire. Ces pôles, les formants, sont les fréquences de résonance du conduit ainsi défini.

Dans le cas de la nasalité, le conduit nasal est branché en dérivation sur le conduit oral. Ce branchement établit un couplage acoustique entre le conduit nasal et le conduit oral modélisé

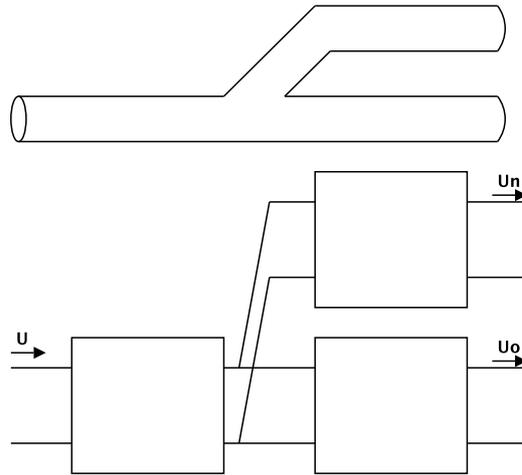


FIG. 5.3 – Représentation du couplage du conduit nasal pour les tuyaux acoustiques et la ligne électrique.

électriquement par le branchement en parallèle d'une chaîne de quadripôles comme le montre la figure 5.3. La ligne électrique du conduit nasal est à nouveau divisée en deux pour tenir compte des deux cavités nasales. En cas de couplage effectif du conduit nasal, c'est-à-dire lorsque le port vélopharyngé est ouvert, l'onde acoustique est donc rayonnée à la fois par l'ouverture aux lèvres et l'ouverture à chacune des deux narines. Nous négligeons dans cette étude les effets du couplage externe entre ces trois tuyaux, et le débit acoustique global à la sortie est calculé de manière simplificatrice comme la somme des trois débits acoustiques distincts. Enfin, les effets d'un éventuel couplage interne à travers le voile du palais ([Dang et Honda, 1996]) sont également négligés.

On définit trois fonctions de transfert acoustiques du conduit vocal :

1. La fonction de transfert *orale* comme le rapport des débits acoustiques aux lèvres et du débit acoustique à la glotte.
2. La fonction de transfert *nasopharyngée* comme le rapport de la somme des débits acoustiques aux narines et du débit acoustique à la glotte.
3. La fonction de transfert *totale* comme le rapport de la somme des débits acoustiques aux narines et aux lèvres et du débit acoustique à la glotte. C'est la somme des deux fonctions de transfert précédentes.

Plusieurs remarques peuvent déjà être formulées à propos des effets du couplage acoustique sur les fonctions de transfert définies ci-dessus. Tout d'abord il convient de rappeler que la fonction de transfert d'un conduit simple sans autre conduit en dérivation ne contient que des pôles et aucun zéro. Ces pôles sont les résonances du conduit. Les zéros apparaissent lors de la connexion d'un autre conduit en dérivation, comme c'est le cas du conduit nasal par rapport au conduit oral lors de l'ouverture du port vélopharyngé. Les zéros qui apparaissent dans la fonction de transfert orale correspondent aux zéros de l'impédance d'entrée du conduit nasal

en dérivation. Réciproquement, les zéros présents dans la fonction de transfert nasopharyngée correspondent aux zéros de l'impédance d'entrée de la cavité buccale, c'est-à-dire du point de séparation du conduit à l'ouverture des lèvres. Les zéros sont donc différents dans les deux fonctions de transfert. Par contre, les pôles des fonctions de transfert orale et nasopharyngée sont identiques. Ils représentent les résonances du conduit complet formé des deux conduits en dérivation. Ils sont donc par nature différents des pôles des fonctions de transfert respectives correspondant aux deux conduits sans que l'un soit en dérivation de l'autre ([Feng, 1986]). Enfin, dans le cas du couplage, les pôles de la fonction de transfert totale, qui est la somme des deux fonctions de transfert orale et nasopharyngée, sont identiques aux pôles de ces deux fonctions de transfert (qui sont déjà identiques entre eux). Les zéros sont par contre encore différents et correspondent à l'opposition de phase des deux sorties ([Feng, 1986]). Nous verrons ultérieurement les évolutions des pôles et des zéros en fonction du couplage, c'est-à-dire de l'ouverture du port vélopharyngé.

Dans la suite du chapitre, les pôles et les zéros sont détectés automatiquement l'un après l'autre par la méthode de Newton-Raphson à partir des fonctions de transfert simulées, puis leur contribution itérativement soustraite pour déterminer le suivant.

5.1.2 Évaluation acoustique du modèle articulaire

Le modèle articulaire développé dans cette étude montre une erreur RMS de reconstruction 0.08 cm pour un seul paramètre et de 0.06 cm pour les deux paramètres. Nous cherchons maintenant à l'aide du modèle acoustique à évaluer ces erreurs de modélisation géométrique par leurs conséquences acoustiques. On dispose pour ceci des sept fonctions d'aire complètes de la glotte aux lèvres et aux narines pour les articulations /a, i, u, \bar{a} , \bar{o} , $\bar{\epsilon}$ et \bar{e} / détaillées dans le chapitre précédent. La fonction d'aire des conduits oral et nasal dans la région vélaire est calculée à partir des maillages du voile du palais et de la paroi du pharynx suivant la méthode décrite dans le chapitre précédent pour trois cas :

- Pour les maillages issus des données servant à construire le modèle.
- Pour les maillages estimés par le modèle limité à son premier paramètre VL.
- Pour les maillages estimés par le modèle complet avec ses deux paramètres VL et VS.

On calcule pour chacun des cas les fonctions de transfert totales. Un exemple de fonctions d'aire et de fonctions de transfert pour les trois cas listés ci-dessus pour l'articulation / \bar{e} / est montré à la figure 5.4. On note la très bonne adéquation entre les trois fonctions d'aire et par conséquent entre les trois fonctions de transfert. Les erreurs sur les six premiers pôles et les deux premiers zéros calculées sur les sept articulations concernées sont résumées pour les deux modèles concernés dans le tableau 5.2. On note la très faible erreur RMS sur l'ensemble des pôles et des zéros qui confirme l'appréciation visuelle de la figure 5.4 : aucune erreur ne dépasse 2 % en fréquence.

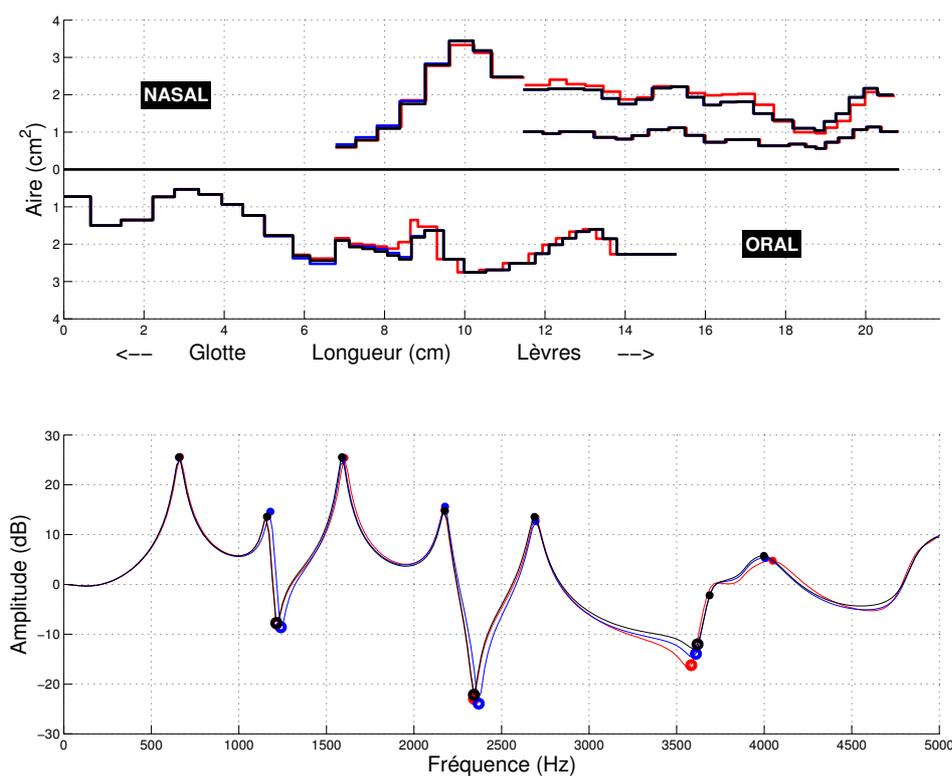


FIG. 5.4 – Superposition des fonctions d’aire du conduit vocal (haut) et des fonctions de transfert totales correspondantes (bas) pour les trois formes du port vélopharyngé considérées (issue des données, modélisée par le modèle à un seul paramètre VL et modélisée par le modèle à deux paramètres VL + VS) pour l’articulation /ê/. Les pôles détectés sont marqués par des points, les zéros par des cercles.

<i>Modèle</i>	<i>P1</i>	<i>P2</i>	<i>P3</i>	<i>P4</i>	<i>P5</i>	<i>P6</i>	<i>Z1</i>	<i>Z2</i>
<i>VL</i>	10 Hz	19 Hz	25 Hz	24 Hz	19 Hz	30 Hz	69 Hz	33 Hz
	1 %	1 %	2 %	1 %	1 %	1 %	2 %	1 %
<i>VL+VS</i>	11 Hz	9 Hz	18 Hz	14 Hz	21 Hz	30 Hz	28 Hz	20 Hz
	2 %	1 %	1 %	1 %	1 %	1 %	2 %	1 %

TAB. 5.2 – Erreurs RMS sur les fréquences des six premiers pôles et des deux premiers zéros calculées sur les sept articulations étudiées grâce au modèle articulaire composé du paramètre VL uniquement ou des deux paramètres VL plus VS.

L’évaluation présentée ci-dessus montre les très faibles erreurs acoustiques correspondant aux erreurs géométriques de reconstruction du modèle montrées dans le chapitre 3. Ces très faibles erreurs nous incitent à utiliser ce modèle dans les sections suivantes pour expliquer l’influence

articulatoire des mouvements du port vélopharyngé et du conduit nasal sur les fonctions de transfert acoustique .

5.2 La nasalité dans le triangle vocalique

Les basses fréquences autour des deux premiers formants renferment une part importante des caractéristiques spectrales des voyelles. Le triangle vocalique est une représentation classique dans le plan des deux premiers formants des voyelles orales. Il permet de caractériser dans les basses fréquences de manière simple et efficace les articulations qui correspondent à des voyelles. Dans cette section, nous nous intéressons d'une part à la représentation dans le plan $F1-F2$ de la cible nasopharyngée en tant que cible articulatoire des voyelles nasales du français et d'autre part à la répartition dans ce même plan des articulations orales des voyelles nasales.

5.2.1 La cible nasopharyngée

[Maeda, 1984] propose comme corrélat acoustique de la nasalité la paire de formants située dans la plan $F1-F2$ autour de 300 Hz – 1000 Hz. Cette région, dite *zone nasale*, correspond dans le plan $F1-F2$ des voyelles du français à un espace vide. [Feng et Castelli, 1996] considèrent cette zone dans le cadre de la nasalité comme la *cible nasopharyngée* d'un mouvement descendant du voile du palais qui partirait d'une configuration très haute conduisant à une fermeture du conduit nasal – c'est-à-dire d'une configuration purement orale – pour atteindre une configuration très basse conduisant à une occlusion du conduit buccal – c'est-à-dire une configuration purement nasopharyngée.

Le conduit nasopharyngé est constitué d'une partie fixe, du cavum aux narines, et d'une

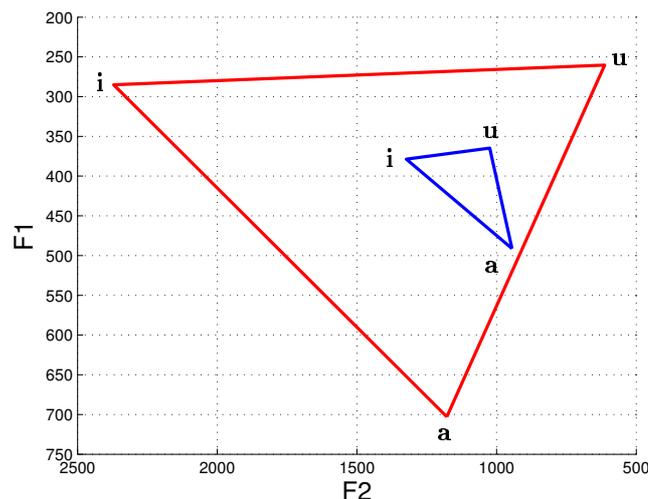


FIG. 5.5 – Positions dans le plan $F1-F2$ des voyelles /a, i, u/ correspondant à une configuration purement orale (grand triangle) et à une configuration purement nasopharyngée (petit triangle).

partie déformable, de la glotte au cavum. Sa forme dans la région pharyngée dépend donc de la voyelle considérée. Une petite variabilité est donc à attendre au niveau des formants $F1$ - $F2$ en fonction de la voyelle considérée. Grâce à des simulations, [Feng et Castelli, 1996] montrent que la répartition dans le plan $F1$ - $F2$ des cibles nasopharyngées correspondant aux onze voyelles orales du français se situe bien la zone nasale, en conservant des positions relatives entre elles à peu près similaires à leurs correspondantes orales. Comme nous l'avons détaillé dans le chapitre précédent, cette étude se concentre sur les trois voyelles orales /a, i, u/ qui bornent l'espace vocalique du français dans le plan $F1$ - $F2$.

La figure 5.5 représente les triangles vocaliques obtenus pour des configurations purement orales et purement nasopharyngées des trois voyelles /a, i, u/. Seul le port vélopharyngé est déformé entre ses deux positions extrêmes suivant le paramètre VL, qui représente le principal mouvement d'ouverture/fermeture. Notons que l'occlusion orale pour la voyelle /i/ est atteinte avec une valeur particulièrement forte de VL qui sort du domaine de variation effectivement observé dans les données. On observe que les trois voyelles se concentrent dans une zone 400 Hz – 1100 Hz, légèrement en-dessous de la zone nasale décrite par [Maeda, 1984] et [Feng et Castelli, 1996]. La structure relative des trois voyelles reste relativement similaire en passant de la configuration orale à la configuration nasopharyngée.

5.2.2 Les voyelles nasales oralisées

Le modèle articulaire nous permet également d'étudier les quatre voyelles nasales sur le même schéma. La figure 5.6 montre les positions dans le plan $F1$ - $F2$ des configurations purement orales et purement nasopharyngées des quatre voyelles / \tilde{a} , \tilde{e} , $\tilde{\text{œ}}$, \tilde{o} /. Sur la figure gauche on peut donc observer la position dans le plan $F1$ - $F2$ des voyelles nasales *oralisées* : elles correspondent aux voyelles obtenues à partir des articulations des voyelles nasales mais en remontant le voile

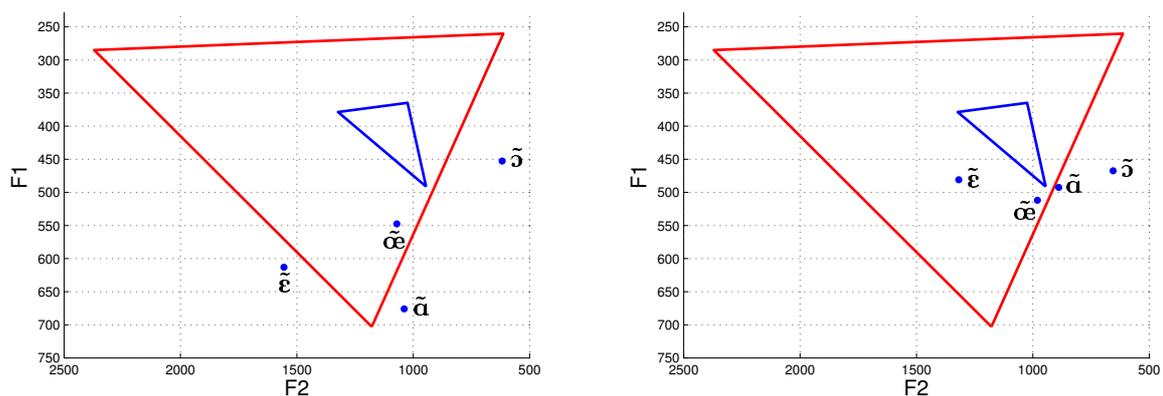


FIG. 5.6 – Positions dans le plan $F1$ - $F2$ des voyelles / \tilde{a} , \tilde{e} , $\tilde{\text{œ}}$, \tilde{o} /. *oralisées* (gauche), c'est-à-dire avec une fermeture du conduit nasal et une ouverture buccale; et représentation des mêmes voyelles avec une occlusion du conduit buccal et une ouverture nasale (droite). Les triangles sont ceux de la figure 5.5

du palais de manière à obtenir la fermeture du conduit nasal. En se basant sur la répartition habituelle des voyelles orales dans le plan $F1$ - $F2$ par rapport à ce triangle, on observe que les articulations artificiellement soutenues / \tilde{a} / et / \tilde{e} / sont proches des articulations des voyelles orales / a / et / ϵ /. Pour le / \tilde{o} /, l'articulation se situe plus ou moins entre / o / et / ɔ /, mais surtout avec un $F2$ plus faible. Enfin de manière intéressante / $\tilde{\text{œ}}$ / semble très proche de / ɔ /. Ces observations acoustiques peuvent être considérées comme relativement cohérentes avec les positions que l'on aurait pu attendre d'après [Zerling, 1984], sauf en ce qui concerne le / $\tilde{\text{œ}}$ /. [Feng et Castelli, 1996] observent dans leurs simulations que les deux premiers formants du / \tilde{a} / oralisé sont plus rapprochés que ceux du / a /, ce qui est vérifié ici. Les valeurs de $F1$ et $F2$ pour les cibles nasopharyngées de ces quatre voyelles nasales sont montrées sur la figure droite. On observe un rapprochement des quatre positions vers le triangle représentant les cibles nasopharyngées de / a , i , u /, sans toutefois pénétrer à l'intérieur. Les quatre positions s'alignent approximativement autour d'un $F1$ de 500 Hz.

D'une manière générale, on peut donc dire que la position dans le plan $F1$ - $F2$ de la cible nasopharyngée des voyelles du français se rapproche de la zone nasale du triangle vocalique avec une certaine variabilité. On note toutefois que la zone atteinte lors de nos simulations est légèrement supérieure en termes de $F1$ à la zone considérée par [Maeda, 1984] et [Feng et Castelli, 1996], avec des valeurs toujours supérieures à 350 Hz et même 450 Hz si l'on fait abstraction du / i / et du / u /. Se pose alors la question de la valeur du fameux $F1$ nasal. La présence dans le spectre d'un pic bas compris entre 250 et 400 Hz lors de la production de voyelles avec ouverture du port vélopharyngé est largement observée dans la littérature (voir par exemple [Lonchamp, 1988] pour une revue détaillée). [Castelli, 1989] confirme même ces valeurs pour notre sujet en mesurant des fonctions de transfert nasopharyngées et calcule les valeurs moyennes des pics pour le cas de la cible nasopharyngée du / a / : les deux premiers sont à 234 Hz et 436 Hz. La présence de ce premier formant a donné lieu à de nombreuses hypothèses sur son origine : la vibration des parois ([Castelli, 1989]), la résonance des cavités sinusales ([Maeda, 1982], [Dang *et al.*, 1994]), la résonance des cavités nasales par une faible ouverture narinaire ([Feng, 1986]), la résonance combinée de plusieurs des cavités citées précédemment ou de la boîte crânienne complète ([Castelli, 1989]), le formant glottique ([Fant, 1979], [Maeda, 1984]), ou encore la résonance des cavités nasales avec une certaine longueur et un certain volume du cavum ([Lonchamp, 1988]). Nos simulations, malgré le très fort lien des fonctions d'aire avec les données enregistrées sur le sujet, et répondant sur ce point à une partie des interrogations, ne nous permettent pas de faire apparaître de formant en-dessous de 350 Hz – 400 Hz. Ces valeurs restent supérieures aux valeurs enregistrées par [Castelli, 1989] sur le même sujet. Notons que si cette section ne présente que les résultats pour la cible nasopharyngée, le reste de l'étude ne fera pas apparaître de formant inférieur au $F1$ présenté ci-dessus. [Lonchamp, 1988] souligne à ce sujet que « la fréquence du premier pic des conduits couplés est toujours comprise entre les fréquences des premiers formants des conduits oral et nasal non couplés », en dehors de toute autre considération tel qu'un couplage sinusal par exemple. La section suivante s'attache justement à décrire les effets du couplage des deux conduits sur les fonctions de transfert.

5.3 Les effets du couplage acoustique

Les effets du couplage du conduit nasal sur les fonctions de transfert acoustiques ont fait l'objet de nombreuses études (voir par exemple [Lonchamp, 1988] pour une revue détaillée). Le modèle articulatoire développé dans notre étude permet un contrôle du mouvement du port vélopharyngé et des fonctions d'aire associées en restant très proche des données enregistrées sur le sujet. Cette section vise à analyser les évolutions des pôles et des zéros des fonctions de transfert en fonction de l'aire de couplage dans la région vélaire.

5.3.1 Les résonances du conduit nasal

Les résonances des cavités nasales sont importantes car elles définissent les fréquences d'apparition des paires pôles/zéros lors des transitions orales vers nasopharyngées détaillées dans le paragraphe suivant. La figure 5.7 affiche les fonctions de transfert du conduit nasal, correspondant aux fonctions d'aire rappelées sur la partie gauche de la figure, en considérant deux cavités distinctes ou une seule qui serait la somme des deux en termes d'aires. Lorsqu'on ne considère

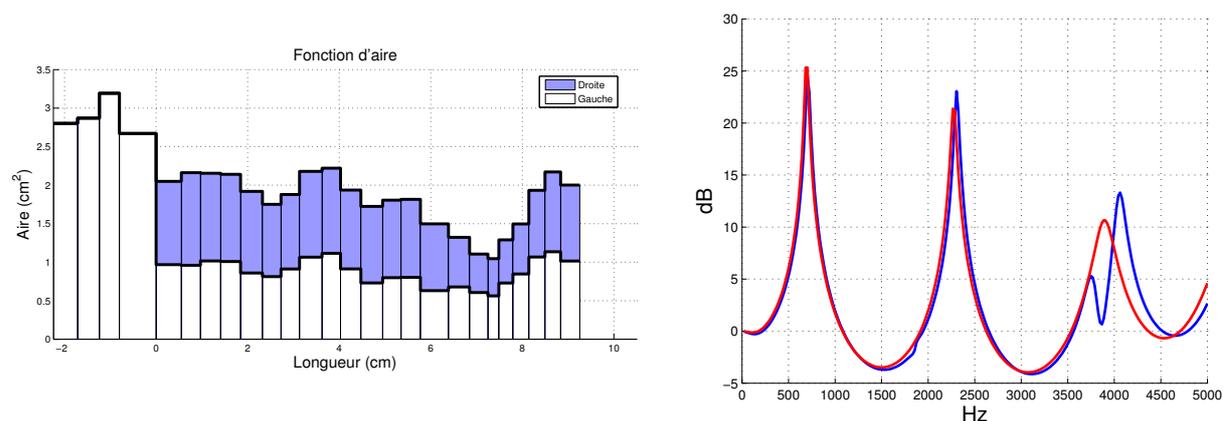


FIG. 5.7 – Fonction d'aire de la partie rigide du conduit nasal (gauche) et fonctions de transfert associées modélisé par un seul ou deux tuyaux (droite).

qu'une seule cavité globale, la fonction de transfert ne contient que des pôles, exactement trois entre 0 Hz et 5000 Hz (700 Hz, 2300 Hz et 3900 Hz). Ils correspondent approximativement aux résonances en quart d'onde d'un tuyau de 12.5 cm de long ouvert en sortie. La longueur ainsi prédite est légèrement supérieure à celle visible sur la figure 5.7 mais les ordres de grandeur sont tout à fait cohérents. Notons donc qu'en dépit d'une constriction au niveau des narines, certes faible, la cavité se comporte comme un tuyau ouvert en sortie. Ces valeurs sont proches de celles qu'on peut trouver dans la littérature. Sur la base des données de [Bjuggren et Fant, 1964], [Lonchamp, 1988] trouve des résonances à 670 Hz, 2200 Hz et 4050 Hz ; sur la base des données de [House et Stevens, 1956] et [Fant, 1960], [Maeda, 1982] trouve les résonances à 670 Hz, 2800 Hz et 4200 Hz. [Dang *et al.*, 1994] trouvent une première résonance entre 540 Hz et 710 Hz pour leurs quatre sujets et une seconde résonance autour de 2500 Hz. Deux études

livrent des valeurs sensiblement différentes, avec notamment des $F1$ plus faibles, principalement à cause d'une ouverture nasale artificiellement basse ([Lonchamp, 1988]) : [House et Stevens, 1956] livrent des fonctions d'aire qui conduisent à des résonances à 356 Hz, 2098 Hz et 3344 Hz (selon [Lonchamp, 1988]) avec une ouverture nasale de seulement 0.23 cm^2 et [Feng, 1986], en se basant sur les données de [Bjuggren et Fant, 1964], trouve des résonances à 450 Hz, 2700 Hz et 4000 Hz avec une ouverture nasale à 0.6 cm^2 .

[Lindqvist-Gauffin et Sundberg, 1972] ont suggéré que la dissymétrie des cavités nasales droite et gauche pouvait jouer un rôle important sur les fonctions de transfert. La figure 5.7 montre les fonctions de transfert obtenues en ne considérant d'une part qu'un large tuyau commun et d'autre part les deux tuyaux distincts. La présence d'un tuyau en dérivation de l'autre joue le même rôle que le tuyau nasal en dérivation du tuyau oral : l'apparition de paires pôles/zéros dans la fonction de transfert globale. [Lonchamp, 1988] souligne d'ailleurs que « deux conduits identiques (sans pertes) au même point de couplage sur un troisième se comportent exactement comme un conduit unique ayant une aire double ». On voit que la relative faible dissymétrie observée dans les données, inférieure à 20% pour les aires (voir la section 3.1), conduit à l'apparition d'une paire pôle/zéro autour de 3800 Hz qui déplace la troisième résonance de 3900 Hz à 4050 Hz. Les deux premières résonances restent inchangées, avec des variations inférieures à 2%. La valeur de cette paire pôle/zéro apparaît plus de deux fois supérieure à celle proposée par [Lonchamp, 1988] (autour de 1700 Hz) et presque le double de celle observée par [Dang *et al.*, 1994] sur quatre sujets (entre 2000 Hz et 2500 Hz). Cette forte valeur dans nos simulations est probablement à attribuer à la faible dissymétrie de nos données : nous avons vérifié que l'augmentation artificielle de la dissymétrie entre les cavités droite et gauche dans nos données faisait effectivement apparaître une paire pôle/zéro dans la zone 1600 Hz – 1700 Hz. Dans la suite, nous utiliserons toujours deux conduits distincts pour les cavités nasales.

L'étude des résonances des cavités nasales constituait un préalable à l'étude du couplage proprement dit. Le paragraphe suivant détaille les effets du couplage sur les fonctions de transfert.

5.3.2 La transition du conduit oral vers le conduit nasopharyngé

[Feng et Castelli, 1996] considèrent la nasalité comme la transition d'une articulation purement orale vers l'articulation purement nasopharyngée correspondante. La figure 5.8 montre ainsi les trois étapes de la nasalisation pour l'articulation $/\bar{\epsilon}/$ artificiellement soutenue dans notre corpus : fonction de transfert orale, fonction de transfert nasale et fonction de transfert nasopharyngée. On observe que la fonction de transfert de l'articulation nasale observée dans notre corpus est bien plus complexe, avec plus de pôles et des zéros, que les deux cibles affichées sur la figure. La vision dynamique de la nasalité proposée par [Feng et Castelli, 1996] suppose une évolution continue et un suivi des différents pôles et zéros de la fonction de transfert totale entre les deux configurations : l'origine des résonances observées peut donc être déterminée. Nous avons montré que le paramètre VL permet de contrôler l'ouverture et la fermeture du port vélopharyngé et représente le mouvement principal du voile du palais. C'est donc ce paramètre

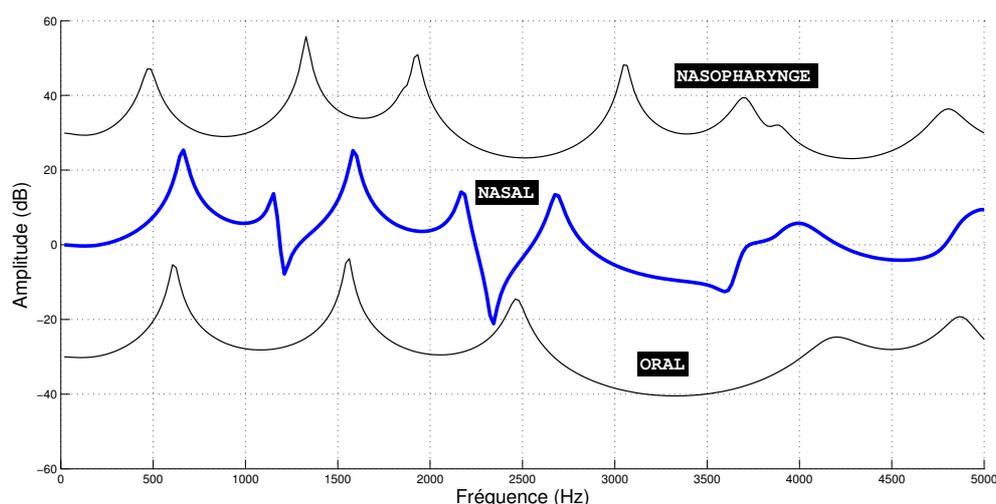


FIG. 5.8 – Fonctions de transfert totales pour l’articulation purement orale ($VL = +1.4$, courbe du bas), pour l’articulation nasale ($VL = -1.5$, courbe du milieu), possédant un couplage entre les deux conduit, et pour l’articulation purement nasopharyngée ($VL = -4.05$, courbe du haut), ceci pour l’articulation $/\bar{e}/$. Les courbes sont décalées entre elles de 30 dB pour une meilleure lisibilité.

que nous allons faire varier pour obtenir les transitions de passage d’un conduit purement oral à un conduit purement nasopharyngé. La figure 5.9 affiche les nomogrammes des fonctions de transfert totales lorsque le paramètre VL varie linéairement entre sa valeur liée à une fermeture du conduit nasal (les courbes du bas avec $VL = +1.4$) et sa valeur liée à une occlusion du conduit buccal (les courbes du haut), et ce pour les trois articulations $/a, i$ et $u/$. Les mêmes nomogrammes ont été réalisés pour les voyelles $/\bar{a}, \bar{i}, \bar{e}$ et $\bar{u}/$ mais ne sont pas affichés ici pour des raisons de place. Notons que l’obtention d’une occlusion buccale dans le cas du $/i/$ n’est possible que grâce à une valeur de $VL = -4.5$ largement inférieure aux valeurs minimales trouvées dans les données ($VL = -2$) pour abaisser suffisamment le voile du palais et assurer le contact avec la langue qui est dans une position très avancée. Ceci explique le plus grand nombre de fonctions présentés pour cette articulation que pour le $/a/$ ($VL = -2$ à son minimum), lui-même représenté avec plus de fonctions que pour le $/u/$ ($VL = -1.6$ à son minimum).

On observe sur la figure 5.9 des évolutions complexes des pôles et des zéros entre les deux cibles. Depuis [Fant, 1960], ces évolutions ont été très bien détaillées dans la littérature, en particulier par [Fujimura et Lindqvist, 1971] et [Maeda, 1993]. Nous rappelons les principales propriétés des pôles et des zéros lors des transitions orales - nasopharyngées :

1. Les fonctions de transfert correspondant aux deux configurations extrêmes, en tant que fonctions de transfert de conduits simples sans dérivation, ne possèdent que des pôles et aucun zéro, à l’exception naturellement des zéros pouvant résulter de la dissymétrie des cavités nasales.

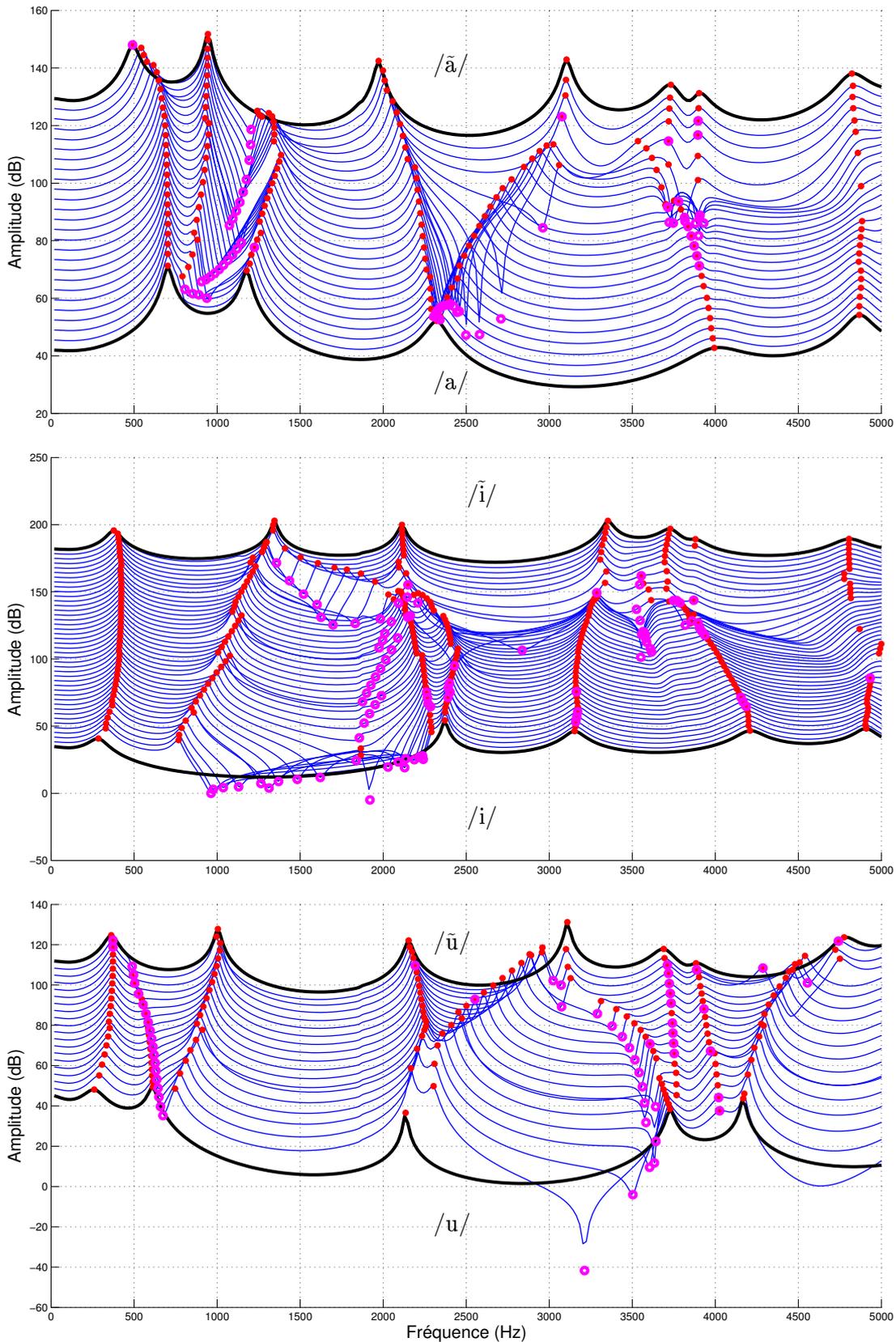


FIG. 5.9 – Fonctions de transfert totales pour le passage d'un conduit oral (courbes du bas) à un conduit nasopharyngé (courbes du haut) pour les articulations /a, i, u/. Les courbes sont décalées entre elles de 3.5 dB pour une meilleure lisibilité. Les pôles détectés sont représentés par des points et les zéros par des cercles. VL varie linéairement de +1.4 à -2 pour /a/, de +1.4 à -4.5 pour /i/ et de +1.4 à -1.65 pour /u/ par pas de 0.15.

2. Les pôles et les zéros apparaissent et disparaissent par paires en s'annulant mutuellement. Combinée avec la propriété précédente, il s'ensuit que les deux fonctions de transfert extrêmes orale et nasopharyngée doivent posséder le même nombre de pôles, toujours à l'exception de pôles pouvant résulter de la dissymétrie des cavités nasales.
3. Lors d'une transition orale vers nasopharyngée, les paires pôles/zéros apparaissent autour des fréquences de résonance du conduit nasal seul, c'est-à-dire du point de fermeture du port vélopharyngé jusqu'aux sorties narinaires, et disparaissent autour des fréquences de résonance du conduit buccal seul, c'est-à-dire du point d'occlusion dans la région vélaire jusqu'à la sortie aux lèvres. On peut donc considérer que les deux fonctions de transfert extrêmes orale et nasopharyngée possèdent des paires pôles/zéros à ces fréquences de résonance qui sont invisibles car les zéros annulent exactement les pôles ; ces paires vont effectivement apparaître dès que l'occlusion sera relâchée, c'est-à-dire dès que les deux tuyaux seront couplés.
4. Lors des transitions, il est admis que l'ordre des pôles reste invariant, de même que celui des zéros ([Fant, 1960]) : deux formants (deux pôles) ne peuvent jamais se « croiser ». Cette propriété n'est pas vraie pour les pôles et les zéros les uns envers les autres. Les zéros, par définition moins nombreux que les pôles, et avec généralement une amplitude de déplacement fréquentiel plus importante, peuvent donc croiser des pôles ; ils peuvent avoir un déplacement assez important pour croiser plusieurs pôles lors de la transition.
5. La conséquence de la propriété précédente est qu'un zéro qui apparaît associé avec un pôle peut disparaître avec autre pôle. On assiste donc lors d'une transition orale - nasopharyngée à l'apparition de pôles liés au conduit nasal et à la disparition de pôles liés au conduit oral. Notons qu'un zéro peut également disparaître avec le pôle avec lequel il est apparu. [Feng et Castelli, 1996] listent les différents types d'apparition-disparition possibles des paires pôles/zéros.

Ces propriétés d'évolution des paires pôles/zéros permettent de distinguer dans la fonction de transfert de la cible nasopharyngée les formants issus d'une résonance orale et déplacés sous l'effet du couplage (*shifted formant* dans la littérature) et les formants issus d'une résonance de la cavité nasale et également déplacés à cause de la variation du couplage.

Les propriétés d'évolution des pôles et des zéros décrites ci-dessus offrent une base pour comprendre les évolutions formantiques observées lors de nos simulations de transitions orales - nasopharyngées. Un schéma type des évolutions rencontrées dans nos simulations est montré à la figure 5.10 pour la voyelle /œ/. En référence aux termes adoptés par [Maeda, 1993], les pôles de la fonction de transfert d'une configuration purement orale seront nommés les *formants*, ceux d'une configuration purement nasopharyngée les *résonances nasales*, les fréquences d'apparition des paires pôles/zéros en partant du conduit purement oral les *fréquences critiques de la cavité nasale*, et enfin celles de disparition des paires en arrivant vers le conduit purement nasopharyngé les *fréquences critiques de la cavité buccale*.

On observe sur la figure 5.10 cinq formants oraux pour sept résonances nasales, en contradic-

tion avec la propriété 2 ci-dessus. Nous avons vérifié que les deux résonances nasales légèrement inférieures à 4000 Hz sont à attribuer l'une à une résonance nasale et l'autre à la dissymétrie des cavités nasales qui présentent une paire pôle/zéro à cette fréquence (voir la figure 5.7). Ces deux résonances très proches sont visibles de manière très stable sur toutes nos simulations (voir la figure 5.9 pour /a, i et u/). Notons paradoxalement que le zéro qui apparaît dans cette région sous l'effet du couplage semble disparaître des observations sans annuler de pôle. En dehors de ces deux résonances, on constate une série d'apparitions et de disparitions de paires pôles/zéros qui maintiennent le nombre de résonances nasales égal au nombre de formants oraux, c'est-à-dire cinq dans notre cas entre 0 Hz et 5000 Hz. Lorsque le port vélopharyngé s'ouvre très légèrement et que les deux conduits sont couplés, deux paires pôles/zéros apparaissent systématiquement à 700 Hz et 2300 Hz, c'est-à-dire aux résonances critiques du conduit nasal (voir la figure 5.7). Les zéros apparus disparaissent soit avec le pôle apparu conjointement, soit avec d'autres pôles, aux fréquences de résonance de la cavité buccale. Sur la figure 5.10 le premier zéro disparaît avec le second formant et le second zéro disparaît avec le pôle avec lequel il est apparu. Sur la figure 5.9 le premier zéro disparaît dans le cas du /a/ avec le second formant, qui correspond justement à une résonance de la cavité buccale et qui disparaît logiquement dans le cas du conduit nasopharyngé pur à la fréquence critique de la cavité buccale. Le même phénomène s'observe pour le /u/ où c'est le second formant, correspondant à une résonance de la cavité avant, qui disparaît annulé par le premier zéro à la fréquence critique de la cavité buccale. Sur le cas plus difficile du /i/, le premier zéro vient annuler l'un des deux pôles apparus lors du couplage. Le second zéro qui apparaît vers 2300 Hz se déplace plus que le premier et vient annuler des pôles jusque vers 3400 Hz, comme il semble que cela soit également le cas pour le /a/.

Suite à l'analyse systématique des transitions orales - nasopharyngées pour les sept articulations /a, i, u, \tilde{a} , \tilde{i} , \tilde{e} et $\tilde{\text{œ}}$ /, nous pouvons formuler quelques remarques générales :

- Le premier pôle est toujours le premier formant oral déplacé. Ce formant ne disparaît jamais annulé par un zéro. Il apparaît d'ailleurs logiquement affilié à une résonance d'une cavité arrière pour le /a/, le /i/ et le /u/. Comme le montre la figure 5.6, sa valeur est toujours comprise entre 350 Hz et 550 Hz pour le conduit nasopharyngé, et il est toujours inférieur à la première résonance critique du conduit nasal dans le cas du conduit oral ; il reste donc toujours inférieur à la première paire pôle/zéro qui apparaît.
- La seconde résonance nasale est toujours issue d'une résonance de la cavité nasale, c'est-à-dire d'un pôle apparu conjointement avec un zéro aux résonances critiques du conduit nasal, sauf pour le / \tilde{e} /.
- Les troisième et quatrième résonances sont assez variables et plus difficiles à interpréter en raison de la grande variabilité des pôles et des zéros dans cette région. Il semble que la troisième résonance partage une origine entre formant oral déplacé (/i, u, \tilde{a} , \tilde{i} et $\tilde{\text{œ}}$ /) et résonance nasale déplacée (/a et \tilde{e} /) ; il semble que / \tilde{a} / soit la seule articulation pour laquelle la quatrième résonance nasale est issue d'une résonance nasale déplacée.
- Deux faibles résonances très proches, généralement légèrement inférieures à 4000 Hz, apparaissent entre les résonances 4 et 5 (voir la figure 5.10) du conduit nasopharyngé. Elles ont pour origine une résonance nasale fixe et la dissymétrie des fosses nasales.

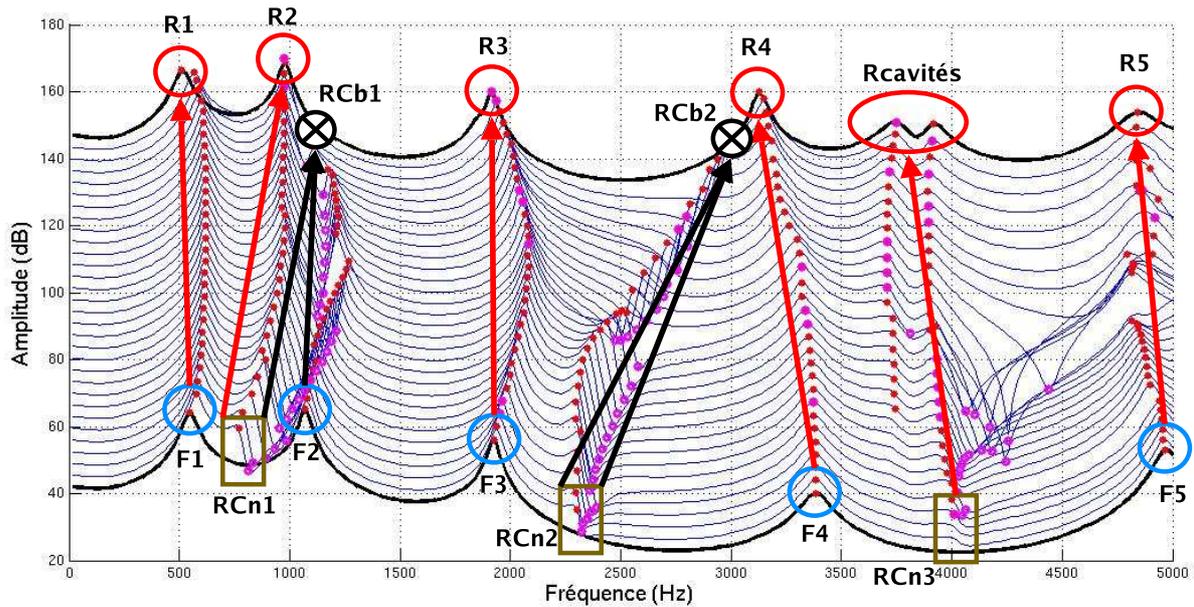


FIG. 5.10 – Schéma d'évolution des pôles et des zéros entre une configuration orale (bas) et nasopharyngée (haut) pour la voyelle /œ/. Les symboles \circ désignent les formants et les résonances nasales deux fonctions extrêmes, les symboles \square désignent les zones d'apparition des résonances critiques du conduit nasal dans la fonction de transfert orale et les symboles \otimes désignent les appariements d'un pôle et d'un zéro qui s'annulent mutuellement ; enfin, les flèches représentent les évolutions des pôles et des zéros entre les deux fonctions extrêmes.

- La résonance numérotée 5 du conduit nasopharyngé paraît toujours provenir du cinquième formant oral, par ailleurs peu déplacé en général.

La principale information apportée par ces simulations est que la première résonance nasale correspond toujours au premier formant oral déplacé, puisque celui-ci se situe toujours en-deçà de la première résonance critique de la cavité nasale. Avec une constriction narinaire de 0.6 cm^2 , [Feng et Castelli, 1996] abaissent la première résonance à 450 Hz et peuvent faire apparaître une paire en-deçà du premier formant oral pour les cas où celui-ci est élevé (pour les configurations /a, ɑ, ɔ et ā/). Ce n'est pas possible avec les fonctions d'aire proposées dans cette étude. On note également que ce formant se déplace sous l'effet du couplage d'un ordre de grandeur compris entre 6% et 35%. Deux sources sont possibles pour le fameux premier formant nasal observé dans la littérature entre 250 Hz et 400 Hz : une résonance issue du couplage des deux conduits oral et nasal ou un phénomène différent qui vient se greffer aux résonances de couplage. Nos résultats plaident pour la seconde solution. [Dang *et al.*, 1994] suggèrent également une résonance des sinus à l'issue de leurs simulations.

5.3.3 Les effets acoustiques des paramètres VL et VS

L'étude précédente a permis d'expliquer les évolutions des pôles et des zéros entre une configuration de fermeture du conduit nasal jusqu'à une configuration d'occlusion du conduit buccal par une évolution continue. Cette étude s'est appuyé sur les mouvements du voile du palais obtenus en faisant varier le paramètre de contrôle VL. Néanmoins les valeurs de ce dernier ont été exagérées de manière à obtenir la fermeture des conduits aux deux extrémités de la transition. On s'intéresse ici aux mêmes évolutions mais pour un ordre de grandeur de VL limité aux valeurs trouvées dans les données de manière à obtenir des évolutions réalistes. L'analyse topologique des pôles et des zéros présentée dans le paragraphe précédent permet d'étiqueter les pôles en connaissant leur provenance, à savoir un formant oral ou une résonance nasale. Il ressort des mesures de variations de chaque pôle une très grande variabilité entre les articulations. Ces variations sont résumées dans le tableau 5.3. On note que seul le pôle *Rn1* issu de la première résonance du conduit nasal varie de manière systématique sur toutes les articulations.

<i>Articulation</i>	<i>F1</i>	<i>F2</i>	<i>F3</i>	<i>Rn1</i>	<i>Rn2</i>
<i>/a /</i>	10 %	10 %	25 %	17 %	9 %
<i>/i /</i>	18 %	1 %	1 %	33 %	5 %
<i>/u /</i>	25 %	23 %	1 %	31 %	22 %
<i>/ā /</i>	17 %	1 %	29 %	15 %	3 %
<i>/õ /</i>	8 %	1 %	38 %	14 %	3 %
<i>/œ /</i>	6 %	11 %	5 %	21 %	15 %
<i>/ê /</i>	3 %	2 %	8 %	38 %	6 %

TAB. 5.3 – Pourcentages de variation des pôles issus des formants oraux (*F1* à *F3*) et issus des résonances nasales (*Rn1* et *Rn2*) pour des variations de VL comprises entre les deux valeurs extrêmes trouvées dans les données.

Les variations de la forme du voile du palais sous l'action du paramètre VS sont moindres que celles correspondant à VL. Le mouvement correspondant à VS est plutôt caractérisé par un petit mouvement en amplitude qui agit significativement sur l'aire de couplage du port vélopharyngé, c'est-à-dire l'aire de couplage nasal. Il ne correspond pas, contrairement à VL, à un large mouvement qui irait de la fermeture nasale jusqu'à l'occlusion buccale comme on l'a vu dans le paragraphe précédent. On étudie donc dans ce paragraphe les effets formantiques du mouvement du voile du palais sous l'action de VS dans une gamme d'amplitude correspondant à celle effectivement constatée dans les données. La figure 5.11 montre un exemple de variation de la fonction de transfert totale pour l'articulation */ê/*, pour VS variant entre les deux valeurs

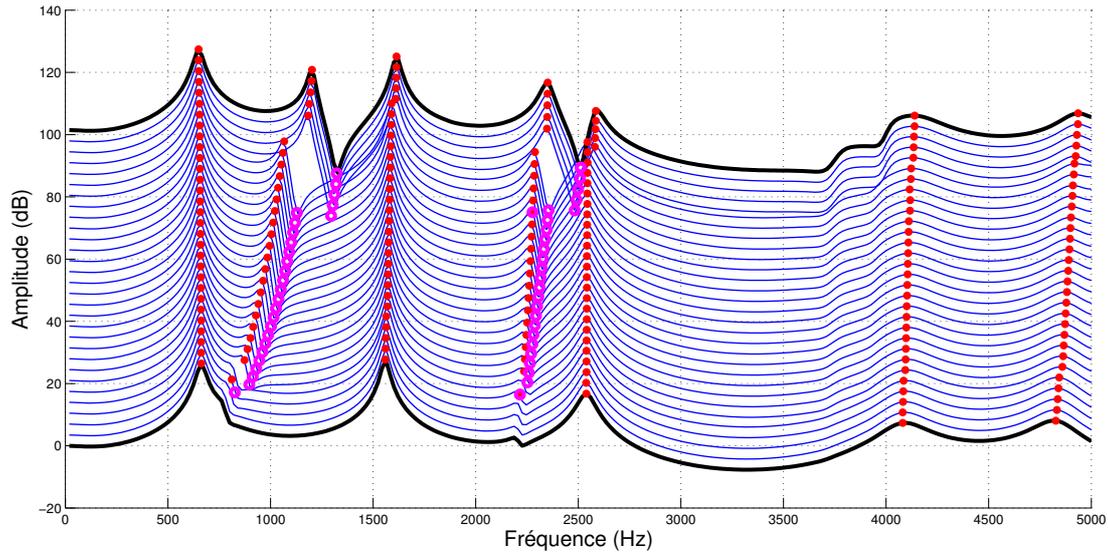


FIG. 5.11 – Fonctions de transfert totales pour des variations linéaires de VS entre les deux valeurs extrêmes trouvées dans les données pour l’articulation $/\bar{\epsilon}/$. La courbe du bas correspond à la valeur minimale de VS, celle du haut à la valeur maximale. Les courbes sont décalées entre elles de 3.5 dB pour une meilleure lisibilité. Les pôles détectés sont représentés par des points, les zéros par des cercles. Le saut brusque entre deux fonctions de transfert successives observé dans la partie supérieure est dû à l’échantillonnage du conduit vocal pour le calcul de la fonction d’aire et non au phénomène de couplage.

extrêmes trouvées dans les données. Lorsque VS est minimum (courbe du bas) l’aire de couplage est aussi minimum (on voit sur la figure 4.16 page 154 qu’elle est inférieure à 0.1 cm^2) mais néanmoins non nulle : la fonction de transfert n’est donc pas constituée uniquement de pôles mais aussi de zéros, aussi faibles soient-ils. De manière plus générale, les courbes de fonctions de transfert obtenues en faisant varier VS ne conduisent jamais à l’occlusion totale de l’un ou l’autre des deux conduits et sont donc constituées de pôles et de zéros. L’analyse topologique des pôles et des zéros présentée dans le paragraphe précédent permet de savoir quels sont les pôles issus de formants oraux et quels sont les pôles issus des résonances nasales. Dans le cas du $/\bar{\epsilon}/$ présenté sur la figure 5.11 on sait pas exemple que l’ordre des pôles observés est le suivant : $F1 < Rn1 < F2 < Rn2 < F3 < F4 < F5$, où F_i désigne le $i^{\text{ème}}$ formant oral et $Rn1$ et $Rn2$ les pôles issus des deux résonances nasales qui apparaissent à 700 Hz et 2300 Hz lors de l’ouverture du port vélopharyngé. Cet étiquetage des différents pôles observés nous permet de calculer les variations de chacun d’eux pour des variations de VS comprises entre les deux valeurs extrêmes trouvées dans les données. Le tableau 5.4 affiche les pourcentages de variation des pôles correspondant à ces variations de VS.

La seule variation des pôles supérieure en moyenne à 10 % est celle de la première résonance nasale avec une variation moyenne de 28 % sur l’ensemble. Considérant que l’impact de la position du voile du palais variant selon VS reste relativement faible sur la fonction d’aire orale mais

<i>Articulation</i>	<i>F1</i>	<i>F2</i>	<i>F3</i>	<i>Rn1</i>	<i>Rn2</i>
/a /	0 %	12 %	4 %	22 %	2 %
/i /	9 %	2 %	1 %	30 %	3 %
/u /	15 %	1 %	1 %	32 %	6 %
/ā /	0 %	13 %	2 %	20 %	8 %
/̄ /	2 %	1 %	3 %	XX	9 %
/œ /	1 %	16 %	3 %	25 %	12 %
/ē /	2 %	3 %	2 %	39 %	6 %

TAB. 5.4 – Pourcentages de variation des pôles issus des formants oraux ($F1$ à $F3$) et issus des résonances nasales ($Rn1$ et $Rn2$) pour des variations de VS comprises entre les deux valeurs extrêmes trouvées dans les données. Les valeurs XX correspondent à des valeurs difficiles à obtenir en pratique sur les courbes.

change significativement l'aire de couplage du port vélopharyngé, il paraît logique d'observer une forte variation des pôles issus directement du couplage et moindre des pôles issus des formants oraux. Les variations du second pôle issu d'une résonance nasale $Rn2$ restent cependant plus faibles que nos attentes. Les seules variations importantes du premier formant sont observées pour le /i/ et le /u/, avec des valeurs respectives de 9 % et 15 %, lorsque le voile du palais modifie la cavité arrière du conduit oral à l'origine de ce premier formant. Pour appuyer ces observations, nous avons reproduit les mêmes calculs en faisant cette fois varier VL entre ses valeurs maximales trouvées dans les données mais sans faire varier la fonction d'aire du conduit oral : seule varie la fonction d'aire du conduit nasal. Cette simulation permet de faire varier sensiblement le couplage des conduits sans faire varier la fonction d'aire du conduit oral. Les résultats sont exactement similaires à ceux présentés ci-dessus : les pôles issus de $F1$ et $F2$ varient en moyenne respectivement à 8 % et 5 % tandis que ceux issus de $Rn1$ et $Rn2$ de 27 % et 7 %. Les seules variations importantes du pôle issu de $F1$ sont à nouveau observées pour /i/ et /u/ lorsque l'ouverture du port vélopharyngé vient modifier les constriction des cavités de Helmholtz responsables des résonances. À nouveau une variation importante de l'aire couplage sans modification importante de la fonction d'aire orale conduit à la variation importante du seul pôle issu de la première résonance nasale. À ce titre la variation du paramètre VS, en amélioration de VL, peut être vue comme un contrôle plus fin du couplage du conduit nasal, et par conséquent des pôles qui sont liés à ce phénomène (au moins le premier pôle), sans modifier exagérément les autres pôles contrôlés par VL.

À l'image de l'analyse précédente qui souligne qu'un petit mouvement du voile du palais modifiant sensiblement l'aire de couplage du conduit nasal conduit à de faibles variations des pôles

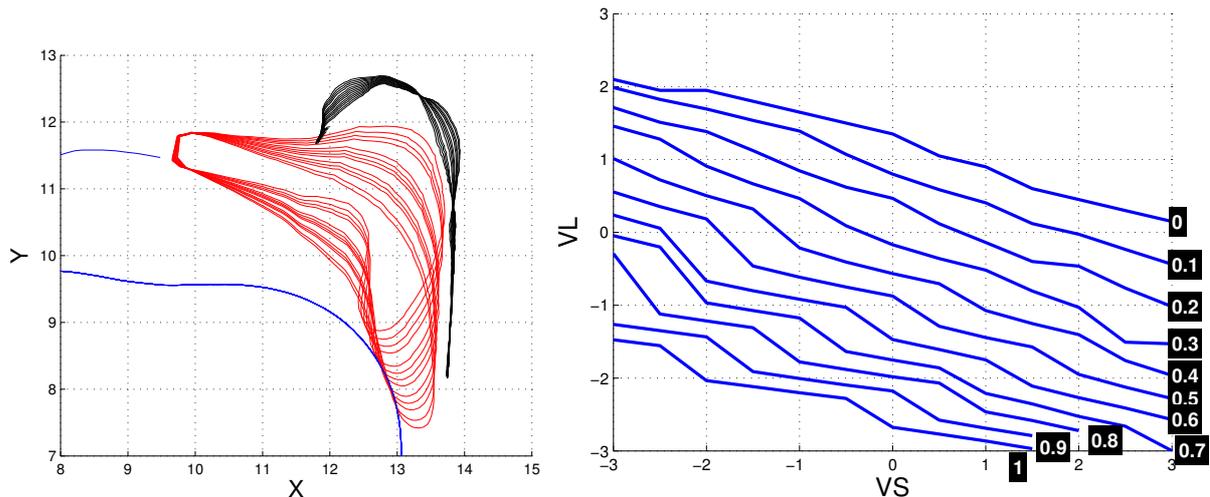


FIG. 5.12 – À gauche, contours médiosagittaux de la langue, du voile du palais et de la paroi du pharynx obtenus en faisant varier conjointement VL et VS de manière à conserver constante l'aire de couplage du port vélopharyngé à 0.4 cm^2 . À droite, courbes le long desquelles l'aire de couplage du port vélopharyngé est constante en fonction des valeurs de VL et VS; les onze courbes représentent une aire de couplage variant de 0 cm^2 à 1 cm^2 par pas de 0.1 cm^2 .

issus des formants oraux et à de fortes variations du premier pôle issu d'une résonance nasale, on peut se demander si la réciproque est vraie : une aire de couplage fixe accompagnée de mouvements du voile du palais qui modifieraient les fonctions d'aire conduit-elle à de fortes variations des pôles issus des formants oraux et à de faibles variations des pôles issus des résonances nasales ? La variation conjointe des valeurs des paramètres de contrôle VL et VS permet d'obtenir des variations de la forme du voile du palais selon les deux mouvements auxquels ils correspondent tout en conservant une aire de couplage constante : voir la figure 5.12 (gauche) pour de tels nomogrammes médiosagittaux correspondant à l'aire de couplage du conduit nasal fixe à 0.4 cm^2 . La figure 5.12 (droite) montre les variations conjointes de VL et VS qui conduisent à une aire de couplage du port vélopharyngé constante. On observe des courbes presque linéaires avec logiquement des variations de VL plus faibles par rapport à celles de VS qui représentent une amplitude de mouvement plus petite.

On étudie les variations des pôles de la fonction de transfert totale le long de ces courbes qui définissent des aires de couplage du port vélopharyngé constantes. Un exemple de la variation de la fonction de transfert totale pour une aire de couplage fixée à 0.4 cm^2 et pour l'articulation /a/ est montré à la figure 5.13. L'analyse topologique des pôles et des zéros détaillée dans le paragraphe précédent nous permet d'identifier les pôles issus des formants oraux et ceux issus des résonances nasales. Pour le /a/, l'ordre des quatre premiers pôles est le suivant : $F1 < Rn1 < F2 < Rn2$, où $F1$ et $F2$ désignent les pôles issus des formants oraux et $Rn1$ et $Rn2$ ceux issus des résonances nasales.

La figure 5.14 montre pour l'articulation /a/ les pourcentages de variation des quatre pre-

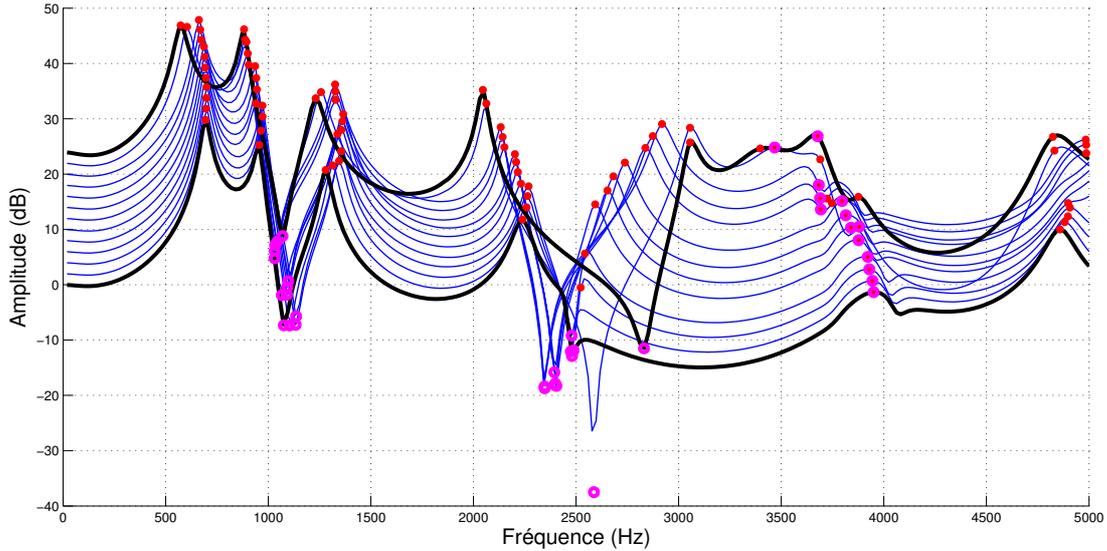


FIG. 5.13 – Fonctions de transfert totales pour des variations conjointes de VL et VS qui conservent l’aire de couplage du port vélopharyngé à 0.4 cm^2 pour l’articulation /a/. Les courbes sont décalées entre elles de 2 dB pour une meilleure lisibilité. Les pôles détectés sont représentés par des points, les zéros par des cercles.

miers pôles de la fonction de transfert totale en fonction de l’aire de couplage du port vélopharyngé pour des variations de VL et VS qui conservent l’aire de couplage constante. On observe que pour une aire de couplage inférieure à $0.3 \text{ cm}^2 - 0.4 \text{ cm}^2$, aucun pôle ne varie de plus de 15 % pour des variations linéaires de VS comprises entre -3 et +3 comme le montre la figure 5.12. Au-delà de cette aire de couplage, on constate une forte augmentation de la variation du premier pôle, qui correspond au premier formant oral, qui vient se stabiliser entre 30 % et 35 % tandis que les variations des autres pôles restent inférieures 15 %. Pour des valeurs d’aires de couplage faibles, le voile du palais est en position haute et/ou arrière (figure 5.12, gauche) : son impact sur les pôles issus des formants oraux reste faible dans la mesure où il modifie peu les fonctions de transfert orales. L’augmentation de l’aire de couplage du port vélopharyngé s’accompagne naturellement d’un abaissement et/ou d’une avancée de la position du voile du palais dans le conduit oral (figure 5.12, gauche) : ce dernier modifie alors sensiblement la fonction d’aire orale dans la région vélaire, conduisant à une variation plutôt forte des pôles associés, en l’occurrence le pôle issu du premier formant oral dans le cas du /a/, qui correspond à une résonance de la cavité arrière. Le pôle issu du second formant oral, correspondant à une résonance de la cavité avant, ne semble pas modifié de manière plus importante que pour une faible aire de couplage. Dans tous les cas les variations des deux pôles issus des résonances nasales restent inférieures à 15 % lorsque l’aire de couplage du port vélopharyngé est stable, quelles que soient les variations du voile du palais. Pour appuyer ces observations, nous avons effectué les mêmes calculs en faisant varier linéairement VL entre ses deux valeurs extrêmes trouvées dans les données mais en gardant fixe la fonction d’aire du port vélopharyngé pour contrôler l’aire de couplage. Cette simulation a été réalisée pour plusieurs valeurs de couplage du port vélopharyngé entre 0 cm^2 et

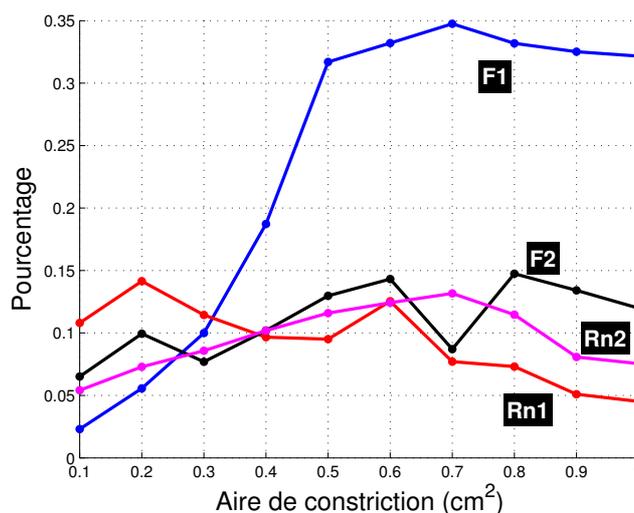


FIG. 5.14 – Pourcentages de variation des quatre premiers pôles de la fonction de transfert totale de l’articulation /a/ en fonction de l’aire de couplage du port vélopharyngé pour des variations de VL et VS conformes à celles de la figure 5.12. *F1* et *F2* désignent les pôles issus des formants oraux et *Rn1* et *Rn2* ceux issus des résonances nasales.

0.8 cm² pour les articulations /a/, /i/ et /u/. Les mêmes résultats sont observés : pour le /a/, seul le pôle issu du premier formant oral enregistre des variations supérieures à 25 % lorsque les autres pôles varient de moins de 10 % en général. Une légère tendance à la décroissance est néanmoins observée lorsque l’aire de couplage du port vélopharyngé augmente : la plus grande ouverture du port vélopharyngé a tendance à amoindrir les effets de la constriction orale obtenue par l’abaissement du voile du palais vers la langue. C’est une situation différente du cas précédent où VL et VS covariaient ; ici VL varie maximale pendant que VS est toujours nul. Ces résultats sont corroborés par des observations similaires pour le /i/. Le /u/ présente cependant des résultats plus mitigés avec des variations des pôles issus des résonances nasales comparables à celles des pôles issus des formants oraux.

Pour résumer, quelques tendances ressortent de ces analyses :

- Une variation de l’aire de couplage du conduit nasal modifie au moins le pôle de la fonction de transfert totale qui est issu de la première résonance nasale, mais pas forcément les autres.
- Une variation des fonctions d’aire qui conserve l’aire de couplage constante modifie peu les pôles issus d’une résonance nasale et fortement ceux issus d’un formant oral affilié à une cavité vélaire (dans le cas du /a/ et du /i/).
- Le paramètre de contrôle VL, ayant un impact important à la fois sur les fonctions d’aire et l’aire de couplage, modifie à la fois les pôles issus des formants oraux et ceux issus des résonances nasales.

- Le paramètre de contrôle VS, ayant un impact important sur l’aire de couplage et moins important sur les fonctions d’aire, modifie plus particulièrement le pôle issu de la première résonance nasale.

Dans la section suivante nous nous intéressons plus particulièrement aux effets acoustiques du mouvement du voile du palais sans prendre en compte le couplage.

5.4 L’influence de la modification de la fonction d’aire orale

[Feng et Kotenkoff, 2006] considèrent la nasalité avant tout comme une modification du conduit oral à laquelle vient s’ajouter une *signature* nasale relativement constante pour l’ensemble des phonèmes. La section présente s’attache à montrer les effets de la constriction vélaire dans le conduit oral sans prise en compte du couplage du conduit nasal. On sait que VS représente de petits mouvements du voile du palais qui vont plutôt modifier l’aire de couplage du port vélopharyngé. On s’intéresse donc à des variations linéaires de VL comprises entre ses deux valeurs maximales trouvées dans les données. La figure 5.15 montre deux exemples d’évolution des fonctions de transfert en fonction de VL pour les articulations /i/ et /u/. On observe de forts mouvements formantiques pour l’articulation /u/ tandis que l’articulation /i/ ne présente que de faibles changements. La variations des formants est résumée pour les trois articulations /a, i, u/ dans le tableau 5.5. La première remarque est que lorsque la cavité dans laquelle se situe le voile

<i>Articulation</i>	<i>F1</i>	<i>F2</i>	<i>F3</i>	<i>F4</i>
/a /	31 % -188 Hz	5 % +66 Hz	3 % -69 Hz	12 % -465 Hz
/i /	8 % +25 Hz	5 % -105 Hz	2 % +57 Hz	9 % -346 Hz
/u /	13 % -33 Hz	28 % -149 Hz	29 % +713 Hz	21 % -711 Hz

TAB. 5.5 – Pourcentages de variation des formants ($F1$ à $F4$) sans couplage nasal pour des variations de VL comprises entre les deux valeurs extrêmes trouvées dans les données pour les articulations /a, i, u/.

du palais est importante, comme c’est le cas pour le /i/, l’influence de celui-ci est relativement faible sur les formants (aucun formant ne varie de plus de 10 %). A l’inverse, lorsque celle-ci est petite comme pour le /u/, les formants varient fortement (ils varient tous de plus de 10 %). Pour le /a/ et le /u/, cette variation de la fonction d’aire environ à mi-distance entre la glotte et les lèvres correspond à une variation d’impédance à la jonction de deux cavités importantes,

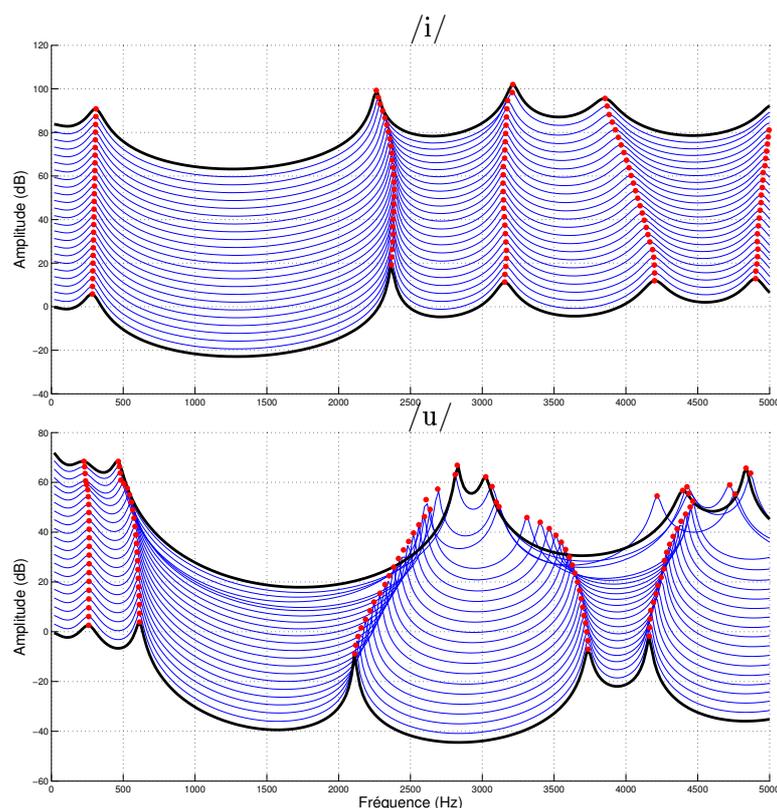


FIG. 5.15 – Fonctions de transfert orales sans couplage nasal pour des variations linéaires de VL entre ses deux valeurs extrêmes trouvées dans les données pour les articulations /i/ et /u/. La courbe du bas correspond à un voile du palais en position haute et inversement. Les courbes sont décalées entre elles de 3.5 dB pour une meilleure lisibilité. Les pôles détectés sont représentés par des points.

conduisant ainsi à une modification des résonances liées à l'une ou l'autre des cavités ou aux deux. Dans le cas du /a/, c'est particulièrement le premier formant, lié à la cavité arrière, qui est sensible au mouvement du voile du palais ; dans le cas du /u/ la modification du tuyau fin entre les deux grandes cavités avant et arrière modifie les résonances de ces deux cavités ainsi que celle du tuyau de liaison. En se basant sur un modèle articulatoire géométrique ([Mermelstein, 1973]), [Bell-Berti et Baer, 1983] ont effectué les mêmes simulations pour les voyelles /u, i, æ, a, ʌ/. Ils n'observent que des variations faibles des deux premiers formants, sauf pour le /u/. Nos résultats confirment la plus forte variation enregistrée pour le /u/ mais montrent des variations dans l'ensemble beaucoup plus importantes. De même [Lonchamp, 1988] enregistre des variations inférieures à 10 % pour les trois premiers formants du /a/, ce que nos mesures ne confirment pas pour le premier formant.

Les analyses présentées dans cette section confirment que la constriction du conduit oral provoquée par un abaissement du voile du palais peut avoir une influence importante sur les formants lorsque l'impédance du conduit varie significativement. Elle montre ainsi une forte influence dans le cas du /u/ et beaucoup plus faible dans le cas du /i/.

5.5 Comparaison avec des enregistrements acoustiques

L'ultime étape de ce travail consiste à comparer les transitions acoustiques décrites dans les sections précédentes avec des enregistrements réels. Les sections précédentes ont permis d'acquérir une bonne connaissance de la topologie des pôles et des zéros lors du couplage du conduit nasal et du mouvement du voile du palais qui constitue une aide précieuse pour l'interprétation des comparaisons des simulations avec les enregistrements réels. Nous avons effectué des enregistrements combinés acoustiques plus bobines EMA pendant lesquels le sujet devait réaliser des transitions du voile du palais d'une position maximale basse à une position maximale haute sans bouger les autres articulateurs. En réalisant cette tâche parfaitement, c'est-à-dire de l'occlusion buccale jusqu'à la fermeture nasale, on obtiendrait des enregistrements de transitions passant d'un conduit purement oral à un conduit purement nasopharyngé. Devant la difficulté à contrôler le voile du palais sans bouger les autres articulateurs, le sujet s'est entraîné longtemps avant l'enregistrement. Pour lui faciliter la tâche, quatre types de transitions sont enregistrés, correspondant aux quatre configurations du conduit vocal pour les quatre voyelles nasales /ã, õ, œ, ê/. Le sujet part de la voyelle nasale qu'il maîtrise bien et tente de remonter son voile du palais vers la fermeture nasale sans bouger les autres articulateurs. Dans la suite du paragraphe, les voyelles oralisées ainsi obtenues des quatre voyelles nasales sont surmontées d'une barre pour les différencier des vraies voyelles orales. Le son est enregistré par un micro positionné devant les lèvres et le nez. Les enregistrements EMA synchrones sont ceux détaillés au paragraphe 2.2.6, avec les cinq bobines collées comme sur la figure 2.10 (page 52).

La connaissance des positions des bobines lors des transitions permet d'obtenir des informations sur deux points essentiels de l'expérience :

- Les enregistrements des bobines collées sur la mâchoire et sur la pointe et au milieu de la langue permettent de vérifier que ces articulateurs restent à peu près fixes lors des transitions.
- L'enregistrement de la bobine collée sur le voile du palais permet d'obtenir les positions du voile du palais synchronisées avec le son.

On a montré au paragraphe 3.7.1 qu'il était possible de retrouver la forme complète du port vélopharyngé avec une erreur très faible à partir des coordonnées de la bobine EMA collée sur le voile du palais par inversion du modèle articulatoire. Les enregistrements des coordonnées de la bobine EMA du voile du palais permettent donc d'obtenir la forme tridimensionnelle du port vélopharyngé et par conséquent des fonctions d'aire qui correspondent à l'enregistrement acoustique synchrone. On mesure par ailleurs les écarts maximaux enregistrés entre les coordonnées des autres bobines et leur valeur moyenne sur l'ensemble des enregistrements : 0.03 cm pour celle de la mâchoire, 0.1 cm pour celle de la pointe de la langue et 0.13 cm pour celle du milieu de la langue. On fait l'hypothèse que l'arrière de la langue ne bouge pas plus que le milieu et on considère ces variations négligeables. On utilise donc la fonction d'aire du conduit oral mesurée pour chacune des quatre voyelles concernées et on calcule la fonction d'aire totale orale plus nasale en fonction de la position du port vélopharyngé de la manière indiquée dans le chapitre 4. Enfin on calcule la fonction de transfert acoustique totale en fonction de la position du port

vélopharyngé. Notons que les fonctions d'aire obtenues par cette méthode montrent que le sujet n'atteint presque jamais l'occlusion parfaite, sauf l'occlusion buccale dans le cas du $/\tilde{a}/$ et du $/\tilde{y}/$. En résumé, la méthode décrite ci-dessus permet d'obtenir à l'aide du modèle les fonctions de transfert totales qui correspondent exactement aux enregistrements effectués.

Les sonagrammes des transitions correspondant aux trois articulations $/\tilde{a}/$, $\tilde{œ}$ et $\tilde{ɛ}/$ sont affichés sur les figures 5.16, 5.17 et 5.18. La figure gauche représente le sonagramme obtenu par le modèle et la figure droite celui des enregistrements correspondants. La transition du $/\tilde{y}/$ vers sa position oralisée n'est pas représentée en raison du bruit plus fort observé. La gauche des figures correspond toujours à la position la plus basse du voile du palais et la droite toujours à la position la plus haute. Notons que les enregistrements ne sont pas d'une qualité excellente et les différentes tentatives pour réhausser le rapport signal sur bruit pour obtenir de meilleurs sonagrammes n'ont pas donné de résultats convaincants ; les pôles détectés grâce au logiciel *Praat* ont néanmoins été superposés sur les figures pour une meilleure interprétation des sonagrammes enregistrés. Ces comparaisons entre simulations et enregistrements constituent donc une première ébauche et nécessiteront par la suite de nouveaux enregistrements. Avant de

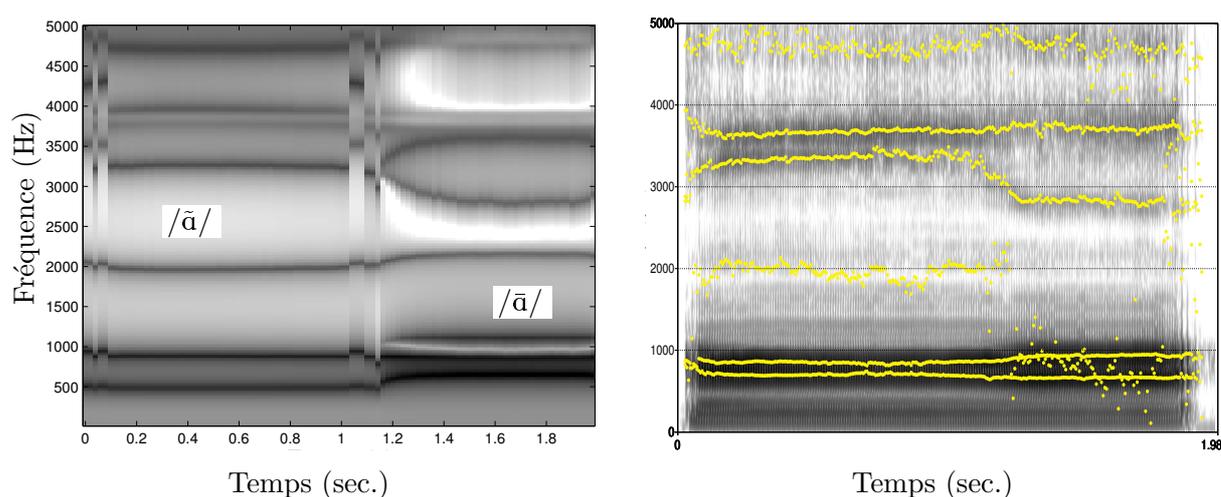


FIG. 5.16 – Sonagrammes de la transition $/\tilde{a}/ \rightarrow /ā/$ obtenus par modélisation (gauche) et par enregistrement sur le sujet avec superposition des pôles détectés (droite). Le sujet remonte le voile du palais au cours du temps.

comparer les simulations obtenues aux enregistrements pour les trois cas, quelques remarques générales peuvent être formulées :

- Comme nous l'avons souligné précédemment, notre modèle ne nous permet pas d'obtenir de formant en dessous de 350 Hz (voir figure 5.6). On observe cependant systématiquement sur les enregistrements un pic plus faible autour de 200 Hz – 300 Hz, qui n'est pas reproduit dans nos simulations.
- Le même schéma d'évolution se reproduit pour toutes les transitions : deux paliers relativement stables avant 1 sec. et après 1.5 sec. et une transition pour passer d'un palier à l'autre entre 1 sec. et 1.5 sec. Cette transition paraît assez cohérente avec la durée moyenne

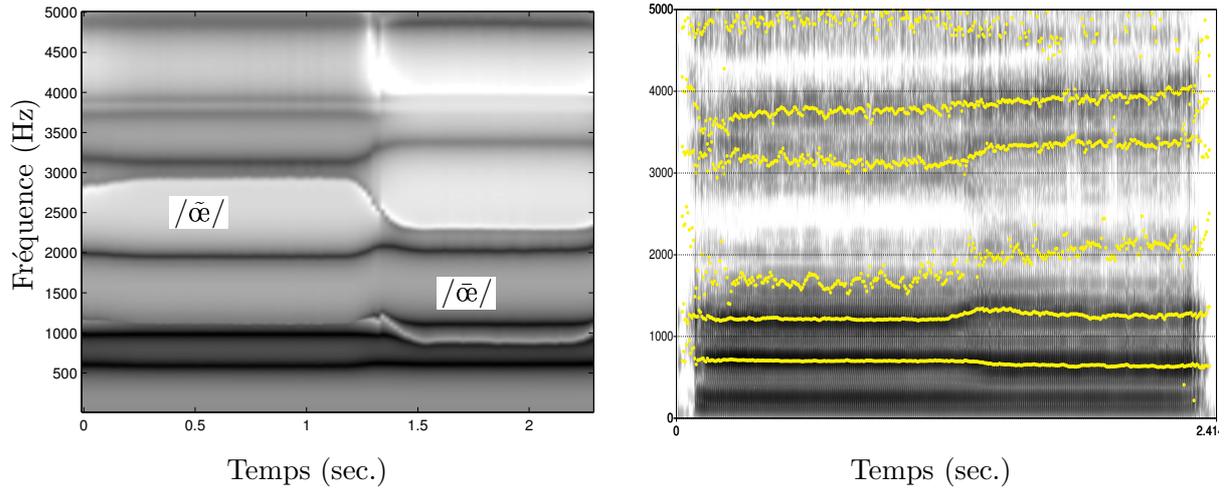


FIG. 5.17 – Sonagrammes de la transition $/\text{œ̃}/ \rightarrow / \text{œ̄}/$ obtenus par modélisation (gauche) et par enregistrement sur le sujet avec superposition des pôles détectés (droite). Le sujet remonte le voile du palais au cours du temps.

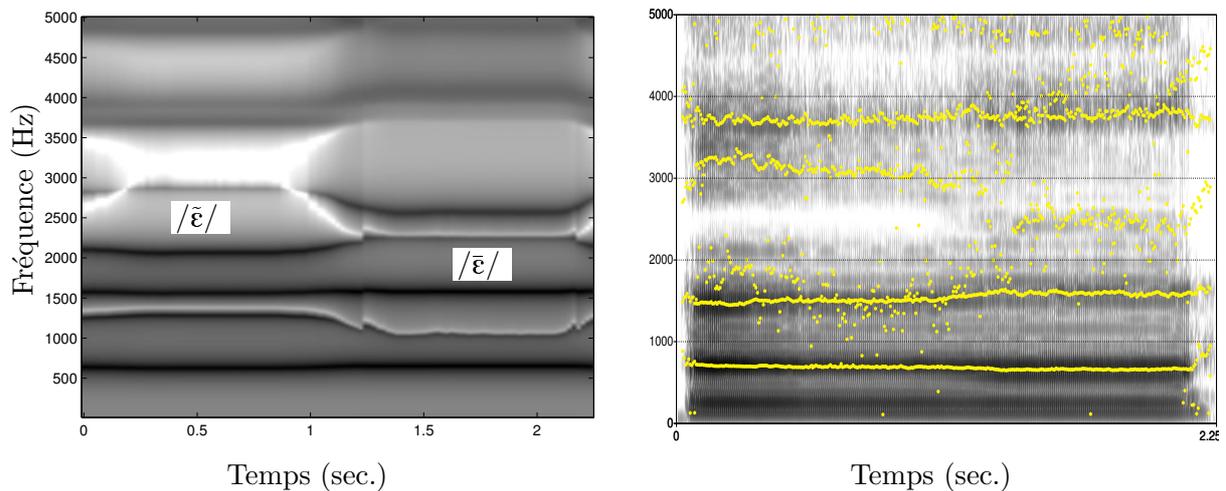


FIG. 5.18 – Sonagrammes de la transition $/\text{ẽ̃}/ \rightarrow / \text{ẽ̄}/$ obtenus par modélisation (gauche) et par enregistrement sur le sujet avec superposition des pôles détectés (droite). Le sujet remonte le voile du palais au cours du temps.

d'une syllabe.

- Sur tous les enregistrements on constate à l'origine du temps, c'est-à-dire lorsque le voile du palais est en position abaissée, une concentration d'énergie entre 3500 Hz et 4000 Hz, quelle que soit la transition considérée ; cette concentration n'est pas sans rappeler les résonances systématiques que nous avons observées dans nos simulations entre 3500 Hz et 4000 Hz.
- On retrouve sur les figures gauches l'enchevêtrement des pôles et des zéros comme ils ont été décrit dans les sections précédentes. Il est possible de connaître pour chacun des cas

la provenance des pôles observés, à savoir un des formants oraux ou une résonance nasale.

Regardons maintenant chaque cas un peu plus en détail. Les descriptions sont faites dans l'ordre temporel inverse, c'est-à-dire de la droite vers la gauche sur les figures, pour passer d'un conduit plutôt oral à un conduit plutôt nasopharyngé.

Pour la transition $/\tilde{a}/ \rightarrow /ā/$ visible sur la figure 5.16, on constate sur les données qu'une grande partie de l'énergie est concentrée dans la région 500 Hz – 1000 Hz. En regardant les données plus précisément on constate deux pôles distincts : l'un légèrement au-dessus de 500 Hz, assez stable en fréquence, et l'autre, plutôt vers 1000 Hz qui baisse légèrement en allant vers la gauche. Ces deux pôles peuvent correspondre aux deux pôles autour de 500 Hz et 1000 Hz observés dans nos simulations, le pôle autour de 1000 Hz étant lui-même composé de deux pôles avant la transition, dont l'un disparaît annulé par un pôle, provoquant une baisse de la concentration d'énergie comme dans les données. On note sur ces dernières une paire pôle/zéro qui disparaît à 1.2 sec et un léger abaissement du premier pôle. Il apparaît presque impossible d'observer dans nos enregistrements le pôle visible à 2000 Hz dans nos simulations. De manière plus intéressante, on constate sur les données deux pôles présents au départ autour de 2900 Hz et 3700 Hz qui se rapprochent fortement au moment du saut de niveau, pour former deux pôles beaucoup plus rapprochés de part et d'autre de 3500 Hz. On constate parallèlement dans les simulations que deux pôles à 2900 Hz et 3700 Hz se rapprochent fortement lors de la transition vers 1.2 secondes. L'un des deux disparaît alors annulé par un zéro. On retrouve alors après cette transition un pôle vers 3200 Hz – 3300 Hz et deux autres de l'autre côté des 3500 Hz. Mis à part le premier formant vers 200 Hz – 300 Hz que l'on ne trouve pas dans les simulations et le pôle vers 2000 Hz que l'on observe presque pas dans les données, l'évolution des pôles obtenue grâce au modèle semble correspondre à peu près aux données, même si de nouveaux enregistrements plus précis s'avèrèrent nécessaires.

Pour la transition $/\tilde{e}/ \rightarrow /ē/$ visible sur la figure 5.17, si l'on ne s'intéresse pas au premier formant en dessous de 500 Hz, on peut considérer que les pôles simulés autour de 500 Hz et 1000 Hz correspondent aux deux pôles un peu plus écartés visibles sur les données. Le second pôle simulé vers 1000 Hz correspond à un regroupement lors de la transition de deux pôles situés de part et d'autre de 1000 Hz ; cette observation peut correspondre dans les données à l'abaissement à la transition du pôle observé au-delà de 1000 Hz. Comme lors de la transition $/\tilde{a}/ \rightarrow /ā/$, le pôle simulé vers 2000 Hz est à peine visible sur les enregistrements, même si une petite concentration d'énergie se retrouve à la fois dans les données et les simulations autour de cette fréquence avant la transition. Le pôle légèrement supérieur à 3000 Hz suit une évolution assez stable et semblable dans les données et les simulations. Un pic d'énergie est toujours visible entre 3500 Hz et 4000 Hz. On constate encore une fois, malgré l'inexistence de formant à 2000 Hz dans les enregistrements, une évolution assez cohérente entre les données et les simulations.

Enfin les changements de formants lors de la transition $/\tilde{e}/ \rightarrow /ē/$ visible sur la figure 5.18 apparaissent moins importants que les précédents. Un premier pôle vers 700 Hz et un second

vers 1500 Hz sont observés à la fois dans les données et les simulations. La paire pôle/zéro visible dans les simulations entre 1000 Hz et 1500 Hz n'est cependant pas visible sur les données. À des fréquences plus élevées on observe dans les simulations deux pôles proches autour de 2500 Hz, à peine visible sur les données, qui se séparent à la transition, le premier descendant vers 2000 Hz, le second montant vers 2800 Hz, laissant un vide vers 2500 Hz, également visible sur les données. Entre 3500 Hz et 4000 Hz, toujours une concentration d'énergie. À nouveaux à l'exception des fréquences centrales autour de 2000 Hz, les simulations montrent des évolutions assez proches de celles des données pour cette articulation également.

Dans cette dernière section on compare les simulations acoustiques obtenues par le modèle avec des enregistrements réels. On note des évolutions qualitatives assez cohérentes entre les résultats du modèle et les enregistrements. Il s'avère néanmoins que certains pôles qui apparaissent dans les simulations ne se retrouvent pas dans les données, comme une concentration d'énergie autour de 2000 Hz systématiquement obtenue en simulation et pas dans les enregistrements. Plusieurs hypothèses sont possibles : trop grande proximité avec un zéro qui rend le pôle invisible sur les données, qualité trop médiocre des enregistrements, fonction d'aire trop éloignée de la réalité, approximations du modèle acoustique trop fortes, etc. Les résultats sont donc très encourageants mais nécessitent un approfondissement sur ce point, en particulier avec des données de meilleure qualité.

5.6 Conclusion

Nous avons présenté dans ce chapitre l'étude des caractéristiques acoustiques liées au modèle articulatoire. Nous avons vérifié dans un premier temps que les effets acoustiques des erreurs géométriques de reconstruction du conduit vocal à l'aide du modèle tridimensionnel étaient négligeables : une différence inférieure à 2 % est constatée entre les pôles et les zéros estimés avec les données ayant servi à construire le modèle et ceux estimés avec le modèle. Le modèle a donc été utilisé pour simuler des transitions d'un conduit oral vers un conduit nasopharyngé. Nous avons en particulier confirmé le rétrécissement du triangle vocalique correspondant au conduit oral vers un triangle plus petit situé dans la *zone nasale* proposée par [Maeda, 1984] avec une position relative à peu près conservée pour les trois voyelles /a, i, u/. La topologie des pôles et des zéros pour les transitions d'une position purement orale à une position purement nasopharyngée a par ailleurs été largement détaillée. Cette analyse a souligné quelques points importants des effets du couplage nasal : l'apparition systématique de deux paires pôles/zéros vers 700 Hz et 2300 Hz, l'apparition d'une concentration d'énergie entre 3500 Hz et 4000 Hz, l'apparition de pôles liés aux résonances nasales et la disparition de pôles liés aux résonances orales. Il apparaît assez clairement que la variation de l'aire de couplage nasal modifie plus particulièrement les pôles issus de résonances nasales tandis que la variation de la forme des fonctions d'aire modifie plus particulièrement les pôles issus de résonances orales. Il s'avère cependant impossible d'obtenir avec ce modèle le fameux premier formant nasal largement discuté dans la littérature (voir par exemple [Maeda, 1982], [Castelli, 1989], [Feng et Castelli, 1996] ou [Dang *et al.*, 1994]). Le rôle de la position du voile du palais dans le conduit oral et la constriction qu'il

produit en conjonction avec la langue lorsqu'il baisse a également été étudié : la constriction orale dans la région vélaire modifie sensiblement les formants, en particulier lorsque la langue est en position arrière et haute et que la région vélaire constitue une constriction du conduit oral comme dans le /u/. Enfin nous avons piloté le modèle à partir des enregistrements EMA pour comparer les enregistrements acoustiques avec nos simulations. Le modèle présente un comportement cohérent avec les données en ce qui concerne l'évolution des pôles lors du mouvement du voile du palais même si l'adéquation avec les données mérite encore d'être approfondie notamment en ce qui concerne les résonances inférieures à 400 Hz absentes dans nos simulations et les résonances autour de 2000 Hz présentes dans les simulations mais absentes dans les enregistrements.

Un certain nombre de résultats présentés dans ce chapitre apportent un nouveau point de vue sur les diverses études articulatoire-acoustiques menées sur la nasalité. Ils s'appuient notamment sur des fonctions d'aire et des mouvements du port vélopharyngé très proches de la réalité, ce qui n'avait jamais été réalisé avant cette étude à notre connaissance. Ils permettent donc de mieux comprendre la complexité des spectres obtenus dans le cas de la nasalité. Ce chapitre se conclut sur une comparaison des résultats obtenus en acoustique avec des enregistrements de données, ce qui résume assez fidèlement la démarche générale adoptée lors de ce travail de thèse !

Conclusion et perspectives

Dans le cadre de cette thèse, nous nous sommes fixé l'objectif de contribuer à la compréhension du phénomène de nasalité, phénomène qui trouve son origine dans le geste d'ouverture du port vélopharyngé. Nous avons choisi une approche modélisatrice au cours de laquelle le modèle émerge des données, qui ont alors logiquement occupé une place centrale dans ce travail. Ces données nous ont ainsi permis de développer un modèle articulatoire tridimensionnel du port vélopharyngé et donc d'étudier et caractériser les conséquences acoustiques de l'abaissement du voile du palais.

Le modèle tridimensionnel linéaire complet du voile du palais et de la paroi du pharynx construits et la caractérisation fine de la géométrie des cavités nasales complètent les travaux de modélisation articulatoire menés au sein du laboratoire dans le cadre du développement de têtes parlantes audiovisuelles virtuelles. Une partie des données traitées au cours de cette thèse, ainsi que la démarche largement détaillée aux chapitres 2 et 3, ont en outre contribué au développement d'un modèle articulatoire linéaire tridimensionnel de la langue basé sur le même locuteur ([Badin et Serrurier, 2006]). Les apports de cette thèse sont discutés plus en détail ci-dessous.

Les données De nombreuses données statiques tridimensionnelles ont été utilisées dans ce travail pour l'élaboration du modèle : images CT, images IRM, moulages dentaires. Ces données de natures diverses ont été calées dans un repère commun pour permettre l'exploitation conjointe et complémentaire des avantages et des caractéristiques de chacune d'entre elles. Nous avons ainsi élaboré une base de formes tridimensionnelles d'articulateurs du conduit vocal, les formes des structures déformables comme le voile du palais et la paroi du pharynx étant obtenues pour chacune des articulations d'un corpus représentatif de la tâche de parole. Cette base de données inédite en parole constitue un atout précieux pour l'étude et la modélisation articulatoire du conduit vocal. Des données de nature dynamique, enregistrées sur le même sujet, et calées dans le même repère, constituent un complément à l'ensemble des données statiques utilisées pour l'extraction de formes tridimensionnelles.

Un modèle articulatoire tridimensionnel L'élaboration d'un modèle articulatoire tridimensionnel conjoint du voile du palais et de la partie supérieure de la paroi du pharynx basé sur les données constitue un autre apport essentiel de ce travail. Nous avons en particulier montré que le voile du palais, principal responsable du mécanisme d'ouverture ou de fermeture du port vélopharyngé, est contrôlé par deux paramètres. Le premier, largement dominant – il

explique à lui seul jusqu'à 83% de la variance du voile du palais –, correspond au mouvement de montée / descente du voile du palais. De manière intéressante, ce paramètre contrôle aussi une légère avancée de la région supérieure de la paroi du pharynx qui vient compléter la fermeture du port vélopharyngé lorsque le voile du palais remonte, ce qui va dans le sens de l'*effet sphincter* mentionné dans la littérature. Ce premier paramètre suffit à contrôler complètement l'ouverture et la fermeture du port vélopharyngé. Le second paramètre, qui explique 6% de la variance du voile du palais, correspond à un léger mouvement avant-arrière du voile du palais qui vient modifier sensiblement l'aire de couplage du port vélopharyngé. Les deux paramètres ont été également observés sur les enregistrements dynamiques effectués sur le même sujet par articulographie électromagnétique ; ceci confirme l'existence de deux degrés de libertés indépendants des mouvements du voile du palais, même si les questions de l'origine biomécanique et du contrôle de la deuxième composante, de plus faible amplitude, restent ouvertes. Le modèle a ainsi permis de faire le lien entre les données statiques à l'origine de la modélisation et les données dynamiques, ce qui constitue une avancée intéressante pour la modélisation articulaire de la production de la parole. Enfin, nous avons confirmé le fort lien entre les contours médiosagittaux et les formes tridimensionnelles du voile du palais déjà observé pour la langue ([Badin *et al.*, 2002]), et montré qu'il est possible de prédire la forme complète tridimensionnelle du voile du palais à partir de mesures articulatoires dynamiques dans le plan médiosagittal. Cette modélisation tridimensionnelle réaliste de la géométrie du conduit vocal a permis la détermination de fonctions d'aire précises du conduit nasal complet et du conduit oral dans la région vélaire nécessaires à une modélisation acoustique par analogie électrique, et ouvre des perspectives pour une modélisation acoustique plus élaborée tenant compte de l'aspect tridimensionnel. Pour terminer, nous avons apporté de nouvelles mesures et une fonction d'aire réaliste des cavités nasales, qui viennent compléter les rares données disponibles dans la littérature sur ce sujet. Nous avons en particulier observé que la fonction d'aire des cavités nasales pour ce sujet ne présente pas de constriction marquée au seuil narinaire, ce qui est en contradiction avec les hypothèses d'une partie de la littérature sur le sujet.

La caractérisation articulatoire-acoustique des mouvements du voile du palais Dans le cadre de la recherche sur la nasalité, les évolutions spectrales liées au couplage du conduit nasal ont été étudiées grâce au nouveau modèle articulaire. Suivant la longue tradition d'analyse des relations articulatoire-acoustiques du laboratoire, nous nous sommes attachés à caractériser les évolutions formantiques du spectre acoustique en lien avec les gestes articulatoires des organes pour un ensemble de configurations vocaliques du conduit oral. Nous avons montré que, en l'absence de constriction narinaire marquée, les cavités nasales résonnent approximativement comme un tuyau ouvert à une extrémité, aux environs de 700 Hz et de 2300 Hz. Nous ne vérifions donc pas l'hypothèse de constriction narinaire avancée par [Feng, 1986]. L'étude de la topologie des pôles et des zéros de la fonction de transfert acoustique observée lorsque les deux conduits sont couplés a permis de leur attribuer une origine orale ou nasale. Les pôles d'origine orale se sont avérés plus sensibles aux changements de la fonction d'aire orale induite par l'abaissement du voile du palais, tandis que les pôles d'origine nasale se sont avérés plus sensibles aux changements de l'aire de couplage du port vélopharyngé. Une étude plus approfondie sur l'ensemble des voyelles du français mérite néanmoins d'être menée pour étayer ce point. La comparaison de

simulations acoustiques obtenues à l'aide du modèle avec des enregistrements du même sujet donne des résultats encourageants, même si les différences sont encore très sensibles ; nous avons en particulier observé que notre modèle ne permettait pas d'obtenir des résonances inférieures à 400 Hz comme le mesure pourtant [Castelli, 1989] sur le même sujet et qu'une résonance systématique autour de 2000 Hz ne se retrouvait pas dans les enregistrements. Ces simulations ouvrent cependant la voie à une future utilisation du modèle pour la synthèse acoustique des nasales, même si la question de l'origine du premier formant nasal, souvent attribuée à une constriction au seuil narinaire dans la littérature, ne semble néanmoins pas pouvoir être résolue avec notre modèle et reste donc ouverte.

Perspectives

Les avancées apportées par ce travail, dont nous avons résumé les principaux résultats ci-dessus, méritent d'être étoffées sur un certain nombre de points. La partie de la géométrie complexe des cavités nasales restée inexploitée dans cette étude pourrait contribuer à l'abaissement de la fréquence du formant bas qui constitue une partie importante des corrélats acoustiques de la nasalité : les divers sinus paranasaux – et les sinus maxillaires en particulier qui sont les plus importants en terme de volume – seront intégrés à court terme dans les fonctions d'aire des cavités nasales et leur influence acoustique évaluée. Cela permettra de mesurer l'influence de ces cavités sur les formants bas. Cet ajout permettrait peut-être de diminuer l'erreur entre formants obtenus à l'aide du modèle acoustique et formants mesurés sur le sujet pour les mêmes articulations. Sur ce dernier point des enregistrements acoustiques bouche-nez séparés effectués sur le même sujet ([Feng et Kotenkoff, 2006]) pourraient se révéler très utiles. La forme complexe des cavités nasales pourrait en outre être exploitée à moyen terme pour élaborer un modèle acoustique plus réaliste, notamment au niveau des pertes et des modes de propagation non plans.

Pour être complètement opérationnel pour l'étude de la nasalité, notre nouveau modèle sera intégré à dans un modèle articulatoire complet, incluant notamment la langue ([Badin et Serurier, 2006]), mais aussi l'ensemble des autres articulateurs du conduit vocal. Ceci permettrait de contrôler précisément la géométrie tridimensionnelle complète des conduits oral et nasal et les fonctions d'aires associées. Ce modèle complet ouvrira en particulier la possibilité de caractériser l'importance du recul de la langue pour les voyelles nasales du français comme suggéré par [Maeda, 1993] ou [Feng et Kotenkoff, 2006].

Enfin, ce modèle pourrait être utilisé comme synthétiseur articulatoire. La synthèse par copie articulatoire basée sur des mesures articulatoires dynamiques des articulateurs pour estimer la forme tridimensionnelle du conduit, comme nous avons déjà commencé à l'expérimenter dans ce travail, pourrait constituer une première évaluation des capacités du modèle pour la synthèse. La synthèse articulatoire complète pourrait être utile à terme pour étudier la perception de la nasalité, comme l'ont fait notamment [Maeda, 1993] ou [Teixeira *et al.*, 2003] à partir de leur modèle, et poursuivre la recherche des corrélats acoustiques de la nasalité.

Cléopâtre n'en a donc pas fini de nous mener par le bout du nez. . .

Annexe A

Modélisation linéaire

Considérons n observations d'un ensemble de m variables. Les observations sont définies par $\{Y_j, j = 1..n\}$ vecteurs de taille $m \times 1$. D'après [Lebart *et al.*, 1995], les résultats mathématiques obtenus en considérant m observations de n données sont identiques. On considère donc pour la suite $n \leq m$. La première opération consiste à centrer les données :

$$\forall j \in \{1..n\}, Y_j^c = Y_j - \frac{1}{n} \sum_{j=1}^n Y_j \quad (\text{A.1})$$

On considère pour la suite de l'explication que les données sont dorénavant centrées. La matrice de covariance qui définit l'ensemble des variances croisées des m variables est alors égale à la matrice de corrélation des données centrées :

$$\text{cov}(Y) = \text{cor}(Y^c) = E\{Y^c Y^{c^t}\} = \frac{1}{n} \sum_{j=1}^n Y_j^c Y_j^{c^t} \quad (\text{A.2})$$

où t définit l'opération de transposition.

Le principe de la modélisation linéaire est de représenter une réalisation quelconque Y_j des données de manière linéaire comme une somme pondérée de p facteurs :

$$\forall j \in \{1..n\}, Y_j = C X_j + Z_j \quad (\text{A.3})$$

où C est une matrice déterministe rectangulaire de taille $m \times p$ dont les colonnes sont les vecteurs définissant chaque composante du modèle ; c'est donc elle qui définit le modèle. X_j , de taille $p \times 1$ est un vecteur de pondération des p composantes du modèle ; il change pour chaque réalisation et constitue le vecteur d'entrée du modèle. Enfin Z_j , de taille $m \times 1$ constitue l'erreur de modélisation pour la réalisation j . Un des critères de modélisation étant de réduire au maximum le nombre de facteurs tout en minimisant l'erreur de modélisation, on choisira $p < m$. L'objectif de la modélisation est de trouver C en minimisant un critère d'erreur, forcément relié à Z , matrice de taille $m \times n$ contenant les erreurs de l'ensemble des observations. On choisit un critère des moindres carrés qui revient à calculer la somme des variances des m composantes individuelles de l'erreur Z . Mathématiquement, elle s'exprime comme suit :

$$\varepsilon = \sum_{i=1}^m \frac{1}{n} \sum_{j=1}^n z_{ij}^2 \quad (\text{A.4})$$

où z_{ij} est l'élément (i, j) de la matrice Z de taille $m \times n$. Cette erreur s'exprime également comme la somme des valeurs diagonales, généralement appelée la *trace*, de la matrice de covariance de Z :

$$\varepsilon = \text{tr}(E\{ZZ^t\}) \quad (\text{A.5})$$

Cette erreur définit donc une erreur globale de l'approximation de l'ensemble des n vecteurs Y par CX .

A.1 Analyse en Composantes Principales

L'objectif est de déterminer la matrice C qui minimise l'erreur globale ε avec la seule connaissance de la matrice Y .

Considérons U , de taille $m \times 1$ l'une des p composantes recherchées, en d'autres termes une des colonnes de C . On cherche U qui minimise l'erreur ε . Une condition nécessaire est alors que la dérivée de ε par U s'annule :

$$\begin{aligned} & \frac{\partial \varepsilon}{\partial U} = 0 \\ \Leftrightarrow & \frac{\partial \text{tr}(E\{ZZ^t\})}{\partial U} = 0 \\ \Leftrightarrow & \frac{\partial \text{tr}(ZZ^t)}{\partial U} = 0 \\ \Leftrightarrow & \frac{\partial \text{tr}((Y - UX)(Y^t - X^tU^t))}{\partial U} = 0 \\ \Leftrightarrow & \frac{\partial \text{tr}(YY^t - YX^tU^t - UX^tY^t + UXX^tU^t)}{\partial U} = 0 \quad (\text{A.6}) \\ \Leftrightarrow & \frac{\partial (\text{tr}(YY^t) - 2\text{tr}(YX^tU^t) + \text{tr}(UXX^tU^t))}{\partial U} = 0 \\ \Leftrightarrow & \frac{\partial \text{tr}(YY^t)}{\partial U} - 2 \frac{\partial \text{tr}(YX^tU^t)}{\partial U} + \frac{\partial \text{tr}(UXX^tU^t)}{\partial U} = 0 \\ \Leftrightarrow & -2 \frac{\partial \text{tr}(YX^tU^t)}{\partial U} + \frac{\partial \text{tr}(UXX^tU^t)}{\partial U} = 0 \quad (\text{A.7}) \end{aligned}$$

On montre facilement par quelques calculs simples que :

$$\frac{\partial \text{tr}(YX^tU^t)}{\partial U} = YX^t \quad \text{et} \quad \frac{\partial \text{tr}(UXX^tU^t)}{\partial U} = 2UXX^t \quad (\text{A.8})$$

Il vient donc naturellement :

$$UXX^t = YX^t \quad (\text{A.9})$$

En combinant avec l'équation (A.3), on conclut que $ZX^t = 0 = E\{ZX^t\}$. Minimiser l'erreur ε revient donc à dire que X et Z sont décorrélés. On peut alors écrire dans le cas optimal à partir de l'équation (A.3) :

$$E\{YY^t\} = CE\{XX^t\}C^t + CE\{XZ^t\} + E\{ZX^t\}C^t + E\{ZZ^t\} = CE\{XX^t\}C^t + E\{ZZ^t\} \quad (\text{A.10})$$

Enfin, d'après le théorème spectral, $E\{YY^t\}$ étant une matrice hermitienne, elle est diagonalisable de la manière suivante :

$$E\{YY^t\} = VDV^t = CE\{XX^t\}C^t + E\{ZZ^t\} \quad \text{avec} \quad VV^t = 1 \quad (\text{A.11})$$

où V est la matrice des vecteurs propres mutuellement orthogonaux entre eux et D la matrice diagonale de ses valeurs propres $\{\lambda_i, i = 1..m\}$, réels positifs. On peut déjà remarquer que si $p = m$, c'est-à-dire le nombre de composantes p est égal à la dimension d'origine m , alors on peut choisir chaque composante de C comme un vecteur propre de V , et par identification $C = V$ et $E\{XX^t\} = D$, d'où $E\{ZZ^t\} = 0$ et enfin $\varepsilon = 0$, c'est-à-dire qu'on peut représenter les données avec une erreur nulle. Néanmoins, un des enjeux est de réduire la dimension de l'espace d'entrée, typiquement moins d'une dizaine de composantes en modélisation articulatoire, voire une ou deux dans le cas qui nous concerne. Dans ce cas $p < m$, $E\{ZZ^t\}$ n'est plus nulle et la somme de ses éléments diagonaux, qui constitue l'erreur globale, non plus. Exprimons alors la valeur de l'erreur globale dans ce cas. La matrice de vecteurs propres V peut être séparée en deux blocs : d'une part les p premières colonnes définissant les p vecteurs propres utilisés pour le modèle, c'est à dire définissant C , et d'autre part la matrice complémentaire B :

$$C = (V_{1..}V_p) \quad \text{et} \quad B = (V_{p+1..}V_m) \quad (\text{A.12})$$

où $\{V_i, i = 1..m\}$ désignent les m colonnes de V , autrement dit les m vecteurs propres de $E\{YY^t\}$. On a vu que l'erreur était nulle dans le cas $C = V$, on peut donc alors écrire :

$$Y = VX_{opt} = [C|B]X_{opt} \quad (\text{A.13})$$

De la même manière que pour V , la matrice X_{opt} est divisée en deux blocs, un premier X de taille $p \times n$ définit par les p premières lignes de X_{opt} et W de taille $(m-p) \times n$ son complémentaire. Alors :

$$Y = VX_{opt} = [C|B]X_{opt} = [C|B] \begin{bmatrix} X \\ W \end{bmatrix} = CX + BW \quad (\text{A.14})$$

et par identification avec l'équation (A.3), $Z = BW$ où B est composé des $(m-p)$ derniers vecteurs propres de la matrice V . En remplaçant sa valeur dans l'équation précédente et en continuant à la dériver :

$$YY^t = CXX^tC^t + ZZ^t \quad (\text{A.15})$$

et

$$E\{YY^t\} = VDV^t = CE\{XX^t\}C^t + E\{ZZ^t\} \quad (\text{A.16})$$

Or VDV^t peut se décomposer comme suit :

$$VDV^t = C \text{diag}(\lambda_1.. \lambda_p) C^t + B \text{diag}(\lambda_{p+1}.. \lambda_m) B^t \quad (\text{A.17})$$

où $\text{diag}(\lambda_k.. \lambda_l)$ définit la matrice diagonale composée des valeurs λ_k à λ_l . Par identification entre les deux équations précédentes, $E\{ZZ^t\} = B \text{diag}(\lambda_{p+1}.. \lambda_m) B^t$. Puisque les colonnes de B , constituées de colonnes issues de la matrice V , sont unitaires, il vient :

$$\varepsilon = \text{tr}(E\{ZZ^t\}) = \text{tr}(B \text{diag}(\lambda_{p+1}.. \lambda_m) B^t) = \text{tr}(\text{diag}(\lambda_{p+1}.. \lambda_m)) = \sum_{i=p+1}^m \lambda_i \quad (\text{A.18})$$

L'erreur globale de modélisation correspond donc à la somme des $(m - p)$ valeurs propres de $E\{YY^t\}$ non utilisées par le modèle. Pour minimiser cette erreur, les vecteurs propres sont toujours classés dans la matrice V par ordre d'importance de leur valeur propre correspondante et le modèle de la matrice C est toujours construit grâce aux p premiers vecteurs de V . Ainsi, les valeurs propres restantes dont la somme forme l'erreur globale sont les plus petites. Ceci revient à faire ressortir les composantes les plus importantes en termes articulatoires. Ce terme d'erreur peut être exprimé en *pourcentage* de l'erreur totale possible, soit la somme de toutes les valeurs propres. Or puisque $E\{YY^t\} = VDV^t$ et V étant unitaire, la variance globale des données Y , c'est-à-dire la trace de sa matrice de covariance, est égale à la trace de D , soit la somme des valeurs propres $\{\lambda_i, i = 1..m\}$ ou encore l'erreur totale possible précédemment définie. On peut donc parler par cette erreur de *pourcentage de variance des données non expliqué*, ou encore en considérant son complémentaire de *pourcentage de variance des données expliqué*.

A.2 Régression Linéaire

Il se peut qu'en plus de la matrice Y des données un paramètre X_0 de taille $1 \times n$ soit connu. On cherche alors le vecteur U , de taille $m \times 1$, associé à ce paramètre, qui sera un des vecteurs de la matrice C du modèle. On note que X_0 couvre bien l'ensemble des observations. Le critère d'ajustement pour trouver U est toujours celui de minimisation de l'erreur de modélisation exprimée à l'équation (A.5). On a vu qu'il était équivalent à l'équation (A.9) rappelée ici :

$$UX_0X_0^t = YX_0^t \quad (\text{A.19})$$

$X_0X_0^t$ étant un simple scalaire, le modèle est immédiatement calculé :

$$U = \frac{1}{X_0X_0^t} YX_0^t \quad (\text{A.20})$$

Il suffit alors de normaliser si besoin.

De même, il se peut qu'en plus de la matrice Y des données la matrice C du modèle soit connue. On cherche alors pour chaque observation Y_j , de taille $m \times 1$, le vecteur correspondant X_j , de taille $p \times 1$. Toujours en utilisant le même critère de minimisation de l'erreur de modélisation définie à l'équation (A.5), et en le dérivant de la même manière qu'aux équations (A.6), non plus par rapport aux vecteurs de C qui sont connus mais par rapport au vecteur de paramètres X , on arrive au critère suivant :

$$X^t C^t C = Y^t C \quad (\text{A.21})$$

C étant de plein rang (colonnes deux à deux indépendantes par construction), $C^t C$ est inversible et on calcule facilement X :

$$X^t = Y^t C (C^t C)^{-1} \quad (\text{A.22})$$

Publications

- [Serrurier et Badin, 2006] Serrurier, A. et Badin, P. (2006). A three-dimensional articulatory model of nasals based on MRI and CT data (submitted). *Journal of the Acoustical Society of America*.
- [Serrurier et Badin, 2005a] Serrurier, A. et Badin, P. (2005a). A Three-Dimensional Linear Articulatory Model of Velum Based on MRI Data. In *Proceedings of Interspeech*, pages 2161–2164, Lisbon, Portugal.
- [Serrurier et Badin, 2005b] Serrurier, A. et Badin, P. (2005b). Towards a 3D articulatory model of velum based on MRI and CT images. *ZAS Papers in Linguistics*, 40 :195–211.
- [Badin et Serrurier, 2006a] Badin, P. et Serrurier, A. (2006a). Three-dimensional linear modeling of tongue : Articulatory data and models. In *Proceedings of the 7th International Seminar on Speech Production*, page To appear, Ubatuba, Brazil.
- [Badin et Serrurier, 2006b] Badin, P. et Serrurier, A. (2006b). Three-dimensional modeling of speech organs : Articulatory data and models. In *IEICE Technical Report*, volume Vol. 106, No 177, SP2006-26, pages 29–34, Kanazawa, Japan. The Institute of Electronics, Information, and Communication Engineers.

Références bibliographiques

- [Amelot *et al.*, 2002] Amelot, A., Basset, P., Crevier-Buchman, L., et Roubeau, B. (2002). Nasalité en français spontané : Mesures aérodynamiques et fibroscopiques, études préliminaires. In *XXIVèmes Journées d'Études sur la Parole, Nancy, France, 24-27 Juin 2002*, pages 361–364.
- [Amelot *et al.*, 2003] Amelot, A., Crevier-Buchman, L., et Maeda, S. (2003). Observations of the Velopharyngeal Closure Mechanism in Horizontal and Lateral Directions from Fiberscopic Data. In *Proceedings of 15th ICPHS, Barcelona, Spain*, pages 3021–3024.
- [Badin, 2002] Badin, P. (2002). *Parole d'Homme - Parole de Clone : Vers une machine parlante anthropomorphique : Données et modèles en production de la parole*. Habilitation à diriger des recherches, Institut National Polytechnique de Grenoble, France.
- [Badin *et al.*, 2003] Badin, P., Bailly, G., Elisei, F., et Odisio, M. (2003). Virtual Talking Heads and audiovisual articulatory synthesis. In M.-J. Solé, D. R. et Romero, J., editors, *Proceedings of the 15th International Congress of Phonetic Sciences*, volume 1, pages 193–197, Barcelona, Spain.
- [Badin *et al.*, 1998] Badin, P., Bailly, G., Raybaudi, M., et Segebarth, C. (1998). A three-dimensional linear articulatory model based on MRI data. In Mannell, R. et Robert-Ribes, J., editors, *Proceedings of the 5th International Conference on Spoken Language Processing*, volume 2, pages 417–420, Sydney, Australia. Australian Speech Science and Technology Association Inc.
- [Badin *et al.*, 2002] Badin, P., Bailly, G., Revéret, L., Baciú, M., Segebarth, C., et Savariaux, C. (2002). Three-dimensional linear articulatory modeling of tongue, lips and face, based on MRI and video images. *Journal of Phonetics*, 30 :533–553.
- [Badin et Fant, 1984] Badin, P. et Fant, G. (1984). Notes on vocal tract computation. *Speech Transmission Laboratory - Quarterly Progress and Status Report - Stockholm, Sweden*, 25 :053–108.
- [Badin *et al.*, 1990] Badin, P., Perrier, P., Boe, L.-J., et Abry, C. (1990). Vocalic nomograms : Acoustic and articulatory considerations upon formant convergences. *Journal of the Acoustical Society of America*, 87 :1292–1300.
- [Badin et Serrurier, 2006] Badin, P. et Serrurier, A. (2006). Three-dimensional modeling of speech organs : Articulatory data and models. In *IEICE Technical Report*, volume Vol. 106, No 177, SP2006-26, pages 29–34, Kanazawa, Japan. The Institute of Electronics, Information, and Communication Engineers.

- [Baer *et al.*, 1991] Baer, T., Gore, J., Gracco, L., et Nye, P. (1991). Analysis of vocal tract shape and dimensions using magnetic resonance imaging : Vowels. *Journal of the Acoustical Society of America*, 90 :799–828.
- [Bailly *et al.*, 2007] Bailly, G., Badin, P., et Revéret, L. (2007). *Audiovisual Speech*, chapter Sensori-motor characteristics of speech production, page (in press). MIT Press.
- [Baken, 1987] Baken, R. (1987). *Clinical Measurement of Speech and Voice*, chapter Chapter 10 : Velopharyngeal Function, pages 393–431. Allyn and Bacon.
- [Beautemps *et al.*, 2001] Beautemps, D., Badin, P., et 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 :2165–2180.
- [Beautemps *et al.*, 1995] Beautemps, D., Badin, P., et Laboissière, R. (1995). Deriving vocal-tract area functions from midsagittal profiles and formant frequencies : A new model for vowels and fricative consonants based on experimental data. *Speech Communication*, 16(1) :27–47.
- [Bell-Berti, 1976] Bell-Berti, F. (1976). An electromyographic study of velopharyngeal function in speech. *Journal of Speech and Hearing Research*, 19 :225–240.
- [Bell-Berti, 1993] Bell-Berti, F. (1993). Understanding velic motor control : studies of segmental context. In *Phonetics and Phonology. Nasals, Nasalization, and the Velum*, pages 63–85. Huffman and Krakow, Academic Press Inc.
- [Bell-Berti et Baer, 1983] Bell-Berti, F. et Baer, T. (1983). Velar position, port size, and vowel spectra. In *Actes du 11ème Congrès International d'Acoustique*, pages 19–21, Paris, France.
- [Bell-Berti et Hirose, 1973] Bell-Berti, F. et Hirose, H. (1973). Patterns of Palatoglossus Activity and Their Implications for Speech Organization. In *Journal of the Acoustical Society of America : 85th meeting of the Acoustical Society of America*, volume 54, page 320.
- [Björk, 1961] Björk, L. (1961). *Velopharyngeal function in connected speech. Studies using tomography and cineradiography synchronized with speech spectrography*. PhD thesis, Department of Diagnostic Radiology (University Hospital, Uppsala) and Speech Transmission Laboratory (Royal Institute of Technology, Stockholm), Sweden.
- [Björk *et al.*, 1961] Björk, L., Nylén, B., Möller, A., et Fant, G. (1961). Velopharyngeal function in connected speech. *Speech Transmission Laboratory - Quarterly Progress and Status Report - Stockholm, Sweden*, 1 :13–14.
- [Bjuggren et Fant, 1964] Bjuggren, G. et Fant, G. (1964). The nasal cavity structures. *Speech Transmission Laboratory - Quarterly Progress and Status Report - Stockholm, Sweden*, 5 :005–007.
- [Boë *et al.*, 1994] Boë, L. J., Maeda, S., et Perrier, P. (1994). La modélisation articulatoire : un demi-siècle d'évolution entre fonctionnel, physique et biomécanique. In *XXèmes Journées d'Études sur la Parole*, pages 41–54, Trégastel, France.
- [Bouchet et Cuilleret, 1991] Bouchet, A. et Cuilleret, J. (1991). *Anatomie Topographique Descriptive et Fonctionnelle*. SIMEP, Paris.
- [Carton, 1974] Carton, F. (1974). *Introduction à la Phonétique du Français*. Bordas, Paris, France.

- [Castelli, 1989] Castelli, E. (1989). *Caractérisation acoustique des voyelles nasales du français. Mesures, modélisation et simulation temporelle*. PhD thesis, Institut National Polytechnique de Grenoble, Grenoble, France.
- [Chiba et Kajiyama, 1941] Chiba, T. et Kajiyama, M. (1941). *The Vowel : Its Nature and Structure*. Tokyo-Kaiseikan, Tokyo, Japan.
- [Clumeck, 1976] Clumeck, H. (1976). Patterns of soft palate movements in six languages. *Journal of Phonetics*, 4 :337–351.
- [Cohen *et al.*, 1998] Cohen, M. M., Beskow, J., et Massaro, D. W. (1998). Recent developments in facial animation : an inside view. In *Proceedings of Auditory-Visual Speech Processing*, Terrigal, Australia.
- [Coker, 1967] Coker, C. H. (1967). Synthesis by rule from articulatory parameters. In *Proceedings of the 1967 IEEE Conference on Speech Communication Processes*, pages 52–53, Boston, USA.
- [Crystal, 1997] Crystal, D. (1997). *A dictionary of linguistics and phonetics*. Blackwell Publishers, Cambridge, USA, 4th edition.
- [Csillag, 1999] Csillag, A. (1999). *Atlas d'Imagerie Médicale. Anatomie de l'Être Humain*. Köne-mann.
- [Dang et Honda, 1996] Dang, J. et Honda, K. (1996). An improved vocal tract model of vowel production implementing piriform resonance and transvelar nasal coupling. In *Proceedings of International Conference on Spoken Language Processing*, volume 2, pages 965–968, Philadelphia, PA.
- [Dang et Honda, 2004] Dang, J. et Honda, K. (2004). Construction and control of a physiological articulatory model. *Journal of the Acoustical Society of America*, 115(2) :853–870.
- [Dang *et al.*, 1994] Dang, J., Honda, K., et Suzuki, H. (1994). Morphological and acoustical analysis of the nasal and the paranasal cavities. *Journal of the Acoustical Society of America*, 96(4) :2088–2100.
- [Delattre, 1954] Delattre, P. (1954). Les attributs acoustique de la nasalité vocalique et consonantique. *Studia Linguistica, Revue de linguistique générale et comparée*, 8(2) :103–109.
- [Delattre, 1968] Delattre, P. (1968). The role of duration in the identification of French nasal vowels. *International Review of Applied Linguistics*, 6(3) :267–288.
- [Delvaux *et al.*, 2002a] Delvaux, V., Metens, T., et Soquet, A. (2002a). French Nasal Vowels : Acoustic and Articulatory Properties. In *Proceedings of International Conference on Spoken Language Processing, Denver, USA, Septembre 2002*, pages 53–56.
- [Delvaux *et al.*, 2002b] Delvaux, V., Metens, T., et Soquet, A. (2002b). Propriétés acoustiques et articulatoires des voyelles nasales du français. In *XXIVèmes Journées d'Études sur la Parole, Nancy, France, 24-27 Juin 2002*, pages 357–360.
- [Demolin *et al.*, 2003] Demolin, D., Delvaux, V., Metens, T., et Soquet, A. (2003). Determination of Velum Opening for French Nasal Vowels by Magnetic Resonance Imaging. *Journal of Voice*, 17 :454–467.

- [Demolin *et al.*, 1996a] Demolin, D., George, M., Lecuit, V., Metens, T., et Socquet, A. (1996a). Détermination, par IRM, de l'ouverture au velum des voyelles nasales du français. In *Actes des XXIèmes Journées d'Études sur la Parole*, pages 83–86, Avignon, France.
- [Demolin *et al.*, 1998] Demolin, D., Lecuit, V., Metens, T., Nazarian, B., et Soquet, A. (1998). Magnetic Resonance Measurements of the Velum Port Opening. In *Proceedings of the 5th International Conference on Spoken Language Processing*, volume 2, pages 425–428, Sydney, Australia.
- [Demolin *et al.*, 1996b] Demolin, D., Metens, T., et Soquet, A. (1996b). Three-dimensional measurements of the vocal tract by MRI. In *Proceedings of the 4th International Conference on Spoken Language Processing*, volume 1, pages 272–275, Philadelphia, PA, USA. University of Delaware and Alfred I. du Pont Institute.
- [Demolin *et al.*, 2000] Demolin, D., Metens, T., et Soquet, A. (2000). Real-time MRI and articulatory coordinations in vowels. In *Proceedings of the 5th Seminar on Speech Production*, pages 86–93, Kloster Seeon, Germany.
- [Dickson et Maue-Dickson, 1972] Dickson, D. R. et Maue-Dickson, W. (1972). Velopharyngeal anatomy. *Journal of Speech and Hearing Research*, 15 :372–381.
- [Dixit *et al.*, 1987] Dixit, R. P., Bell-Berti, F., et Harris, K. S. (1987). Palatoglossus Activity during Nasal/Nonnasal Vowels of Hindi. *Phonetica*, 44 :210–226.
- [Dunn, 1950] Dunn, H. K. (1950). The Calculation of Vowel Resonances, and an Electrical Vocal Tract. *Journal of the Acoustical Society of America*, 22(6) :740–753.
- [Engwall, 1999] Engwall, O. (1999). Modeling of the vocal tract in three dimensions. In *Eurospeech'99*, pages 113–116.
- [Engwall, 2000] Engwall, O. (2000). A 3D tongue model based on MRI data. In *Sixth International Conference on Spoken Language Processing 2000*, volume 3, pages 901–904, Beijing, China.
- [Engwall et Badin, 1999] Engwall, O. et Badin, P. (1999). Collecting and analysing two- and three-dimensional MRI data for swedish. In *TMH-QPSR3-4*, pages 11–38.
- [Engwall et Beskow, 2003] Engwall, O. et Beskow, J. (2003). Effects of corpus choice on statistical articulatory modeling. In *Sixth International Seminar on Speech Production*, pages 1–6, Sydney, Australia.
- [Ettema *et al.*, 2002] Ettema, S. L., Kuehn, D. P., Perlman, A. L., et Alperin, N. (2002). Magnetic Resonance Imaging of the Levator Veli Palatini Muscle During Speech. *Cleft Palate-Craniofacial Journal*, 39(2) :130–144.
- [Fabri, 2002] Fabri, A. (2002). 3D Meshing Software (<http://cgal.inria.fr/reconstruction>).
- [Fant, 1960] Fant, G. (1960). *Acoustic Theory of Speech Production*. Mouton, The Hague.
- [Fant, 1979] Fant, G. (1979). Vocal source analysis - a progress report. *Speech Transmission Laboratory - Quarterly Progress and Status Report - Stockholm, Sweden*, 20(3-4) :31–53.
- [Fant, 1985] Fant, G. (1985). The vocal tract in your pocket calculator. *Speech Transmission Laboratory - Quarterly Progress and Status Report - Stockholm, Sweden*, 26 :001–019.

- [Feng, 1986] Feng, G. (1986). *Modélisation acoustique et traitement du signal de parole - Le cas des voyelles nasales*. PhD thesis, Institut National Polytechnique de Grenoble, Grenoble, France.
- [Feng et Castelli, 1996] Feng, G. et Castelli, E. (1996). Some acoustic features of nasal and nasalized vowels : a target for vowel nasalisation. *Journal of the Acoustical Society of America*, 99 :3694–3706.
- [Feng et Kotenkoff, 2004] Feng, G. et Kotenkoff, C. (2004). Vers un nouveau modèle acoustique des nasales basé sur l'enregistrement bouche - nez séparé. In *Actes des XXVèmes Journées d'Etude sur la Parole, Fès, Maroc*, pages 213–216.
- [Feng et Kotenkoff, 2006] Feng, G. et Kotenkoff, C. (2006). New considerations for vowel nasalization based on separate mouth-nose recording. In *Proceedings of Interspeech'06*, pages 2242–2245, Pittsburgh, Pennsylvania, USA.
- [Ferguson *et al.*, 1975] Ferguson, C. A., Hyman, L. M., et Ohala, J. J., editors (1975). *Nasálfest - Papers from a Symposium on Nasals and Nasalization*. Department of Linguistics, Stanford University, Stanford, USA.
- [Finkelstein *et al.*, 1995] Finkelstein, Y., Shapiro-Feinberg, M., Talmi, Y. P., Nachmani, A., De-Rowe, A., et Ophir, D. (1995). Axial Configuration of the Velopharyngeal Valve and Its Valving Mechanism. *Cleft Palate-Craniofacial Journal*, 32 :299–305.
- [Flanagan, 1972] Flanagan, J. L. (1972). *Speech Analysis Synthesis and Perception*. Springer-Verlag, Berlin, Heidelberg, New-York.
- [Fritzell, 1969] Fritzell, B. (1969). The velopharyngeal muscles in speech. An electromyographic and cineradiographic study. *Acta Otolaryngologica*, Supplement 250 :1–81.
- [Fujimura et Lindqvist, 1971] Fujimura, O. et Lindqvist, J. (1971). Sweep-tone Measurements of Vocal-Tract Characteristics. *Journal of the Acoustical Society of America*, 49(2B) :541–558.
- [Gérard *et al.*, 2003] Gérard, J.-M., Wilhelms-Tricarico, R., Perrier, P., et Payan, Y. (2003). A 3D dynamical biomechanical tongue model to study speech motor control. *Research Developments in Biomechanics*, 1 :49–64.
- [Guérin et Mrayati, 1977] Guérin, B. et Mrayati, M. (1977). Nasal vowels study sensitivity functions. In Carré, R., D. R. et Wajskop, M., editors, *Articulatory Modelling and Phonetics*, pages 137–146. G.A.L.F.
- [Hardcastle, 1976] Hardcastle, W. J. (1976). *Physiology of Speech Production. An Introduction for Speech Scientists*. Academic Press London.
- [Hardcastle et Laver, 1997] Hardcastle, W. J. et Laver, J., editors (1997). *The Handbook of Phonetic Sciences*. Blackwell Publishers, Cambridge, Massachusetts, USA.
- [Harshman *et al.*, 1977] Harshman, R., Ladefoged, P., et Goldstein, L. (1977). Factor analysis of tongue shapes. *Journal of the Acoustical Society of America*, 62 :693–707.
- [Hashimoto et Suga, 1986] Hashimoto, K. et Suga, S. (1986). Estimation of the muscular tensions of the human tongue by using a three-dimensional model of the tongue. *Journal of Acoustic Society of Japan*, 7(1) :39–46.

- [Hawkins et Stevens, 1985] Hawkins, S. et Stevens, K. N. (1985). Acoustic and perceptual correlates of the non-nasal–nasal distinction for vowels. *Journal of the Acoustical Society of America*, 77(4) :1560–1575.
- [Hecker, 1962] Hecker, M. H. L. (1962). Studies of Nasal Consonants with an Articulatory Speech Synthesizer. *Journal of the Acoustical Society of America*, 34(2) :179–187.
- [Heinz et Stevens, 1965] Heinz, J. M. et Stevens, K. N. (1965). On the relations between lateral cineradiographs area functions, and acoustic spectra of speech. In *Proceedings of 5th International Congress on Acoustics*, page A44.
- [Hoole, 1999] Hoole, P. (1999). On the lingual organization of the German vowel system. *The Journal of the Acoustical Society of America*, 106(2) :1020–1032.
- [Hoole et Nguyen, 1997] Hoole, P. et Nguyen, N. (1997). Electromagnetic articulography in coarticulation research. *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation, München, Germany*, 35 :177–184.
- [Hoole et Nguyen, 1999] Hoole, P. et Nguyen, N. (1999). Electromagnetic articulography in coarticulation research. In Hardcastle, W. et Hewlett, N., editors, *Coarticulation : Theory, Data and Techniques*, pages 260–269. Cambridge University Press.
- [House et Stevens, 1956] House, A. S. et Stevens, K. N. (1956). Analog studies of the nasalization of vowels. *Journal of Speech and Hearing Disorders*, 21 :218–232.
- [Huber, 1999] Huber, D. (1999). Smoothmesh Software (<http://www.cs.cmu.edu/vmr/software/meshtoolbox/downloads.html>).
- [Huffman et Krakow, 1993] Huffman, M. K. et Krakow, R. A., editors (1993). *Phonetics and Phonology. Nasals, Nasalization, and the Velum. Vol. 5*. Academic Press, Inc., USA.
- [Kaburagi et Honda, 2001] Kaburagi, T. et Honda, M. (2001). Dynamic articulatory model based on multidimensional invariant-feature task representation. *The Journal of the Acoustical Society of America*, 110(1) :441–452.
- [Kent, 1997] Kent, R. D. (1997). *The Speech Sciences*. Singular Publishing Group, Inc.
- [Kiritani *et al.*, 1976] Kiritani, S., Miyawaki, K., Fujimura, O., et Miller, J. E. (1976). A computational model of the tongue. *Annual Bulletin of the Research Institute for Logopedics and Phoniatrics, Tokyo, Japan*, 10 :243–251.
- [Kitamura *et al.*, 2005] Kitamura, T., Takemoto, H., Honda, K., Shimada, Y., Fujimoto, I., Syakudo, Y., Masaki, S., Kuroda, K., Oku-uchi, N., et Senda, M. (2005). Difference in vocal tract shape between upright and supine postures : Observation by an open-type MRI scanner. *Acoustical Science and Technology*, 5 :465–468.
- [Kollia *et al.*, 1995] Kollia, H., Gracco, V., et Harris, K. (1995). Articulatory organization of mandibular, labial, and velar movements during speech. *Journal of the Acoustical Society of America*, 98(3) :1313–1324.
- [Krakow et Huffman, 1993] Krakow, R. A. et Huffman, M. K. (1993). Instruments and techniques for investigating nasalization and velopharyngeal function in the laboratory : an introduction. In Huffman, M. K. et Krakow, R. A., editors, *Phonetics and Phonology, Volume 5 : Nasals, Nasalization and the Velum*, pages 3–59. Academic Press, Inc.

- [Le Huche et Allali, 1984] Le Huche, F. et Allali, A. (1984). *La Voix, tome 1, Anatomie et Physiologie des Organes de la Voix et de la Parole*. Masson.
- [Lebart *et al.*, 1995] Lebart, L., Morineau, A., et Piron, M. (1995). *Statistique exploratoire multidimensionnelle*. Dunod, Paris, France.
- [Lherm, 1984] Lherm, B. (1984). *Réalisation d'un outil de simulation harmonique et de synthèse temporelle du type analogue du conduit vocal - Étude des voyelles nasales*. PhD thesis, Institut National Polytechnique de Grenoble, France.
- [Liang et Lauterbur, 2000] Liang, Z.-P. et Lauterbur, P. C. (2000). *Principles of Magnetic Resonance Imaging*. IEEE Press, USA.
- [Liljencrants, 1971] Liljencrants, J. (1971). A Fourier series description of the tongue profile. *Speech Transmission Laboratory - Quarterly Progress and Status Report - Stockholm, Sweden*, 12(4) :9–18.
- [Lindblom et Sundberg, 1971] Lindblom, B. E. F. et Sundberg, J. E. F. (1971). Acoustical Consequences of Lip, Tongue, Jaw, and Larynx Movement. *Journal of the Acoustical Society of America*, 50(4B) :1166–1179.
- [Lindqvist-Gauffin et Sundberg, 1972] Lindqvist-Gauffin, J. et Sundberg, J. (1972). Acoustic properties of the nasal tract. *Speech Transmission Laboratory - Quarterly Progress and Status Report - Stockholm, Sweden*, 13(1) :13–17.
- [Lonchamp, 1988] Lonchamp, F. (1988). *Études sur la production et la perception de la parole - Les indices acoustiques de la nasalité vocalique - La modification du timbre par la fréquence fondamentale*. Thèse d'État, Université de Nancy II, France.
- [Lubker *et al.*, 1970] Lubker, J., Fritzell, B., et Lindqvist-Gauffin, J. (1970). Velopharyngeal function : An electromyographic study. *Speech Transmission Laboratory - Quarterly Progress and Status Report - Stockholm, Sweden*, 11(4) :9–20.
- [Lubker *et al.*, 1972] Lubker, J., Lindqvist, J., et Fritzell, B. (1972). Temporal Characteristics of Velopharyngeal Muscle Function. In *Journal of the Acoustical Society of America : 83rd meeting of the Acoustical Society of America*, volume 52, page 125.
- [Mády *et al.*, 2002] Mády, K., Sader, R., Zimmermann, A., Hoole, P., Beer, A., Zeilhofer, H., et Hannig, C. (2002). Assessment of Consonant Articulation in Glossectomee Speech by Dynamic MRI. In *Proceedings of the 7th International Conference on Spoken Language Processing*, pages 961–964, Denver, USA.
- [Maeda, 1978] Maeda, S. (1978). Une analyse statistique sur les positions de la langue : Etude préliminaire sur les voyelles du français. In *Actes des 9èmes Journées d'Etude sur la Parole*, pages 191–199, Lannion, France.
- [Maeda, 1979] Maeda, S. (1979). Un modèle articulatoire de la langue avec des composantes linéaires. In *Actes des 10èmes Journées d'Etude sur la Parole*, pages 152–162, Grenoble, France.
- [Maeda, 1982] Maeda, S. (1982). The role of the sinus cavities in the production of nasal vowels. In *Proc. of IEEE ICASSP*, volume 2, pages 911–914.

- [Maeda, 1984] Maeda, S. (1984). Une paire de pics spectraux comme corrélat acoustique de la nasalisation des voyelles. In *Actes des 13èmes Journées d'Etude sur la Parole*, pages 223–224, Bruxelles, Belgique.
- [Maeda, 1990] Maeda, S. (1990). Compensatory articulation during speech : Evidence from the analysis and synthesis of vocal-tract shapes using an articulatory model. In *Speech Production and Speech Modelling (Hardcastle and Marchal Eds., Kluwer Academic)*, pages 131–149.
- [Maeda, 1991] Maeda, S. (1991). On articulatory and acoustic variabilities. *Journal of Phonetics*, 19 :321–331.
- [Maeda, 1993] Maeda, S. (1993). Acoustics of vowel nasalization and articulatory shifts in French nasal vowels. In *Phonetics and Phonology, Volume 5 : Nasals, Nasalization and the Velum (Huffman and Krakow Eds., Academic Press.)*, pages 147–167.
- [Majid, 1986] Majid, R. (1986). *Modélisation articulatoire du conduit vocal, exploration et exploitation, fonctions de macro-sensibilité paramétriques et voyelles du français*. PhD thesis, Institut National Polytechnique de Grenoble.
- [Matsumura, 1992] Matsumura, M. (1992). Measurement of three-dimensional shapes of vocal tract and nasal cavity using Magnetic Resonance Imaging technique. In *Second International Conference on Spoken Language Processing*, pages 779–782, Banff, Alberta, Canada.
- [Matsumura et al., 1994] Matsumura, M., Niikawa, T., Shimizu, K., Hashimoto, Y., et Morita, T. (1994). Measurement of 3D shapes of vocal tract, dental crown and nasal cavity using MRI : vowels and fricatives. In *Third International Conference on Spoken Language Processing*, pages 619–622, Yokohama, Japan.
- [Merlier, 1984] Merlier, B. (1984). *Études sur le rayonnement externe des conduit oral et nasal*. PhD thesis, Institut National Polytechnique de Grenoble, France.
- [Mermelstein, 1973] Mermelstein, P. (1973). Articulatory model for the study of speech production. *Journal of the Acoustical Society of America*, 53 :1070–1082.
- [Minifie et al., 1973] Minifie, F. D., Hixon, T. J., et Williams, F., editors (1973). *Normal Aspects of Speech, Hearing, and Language*. Prentice-Hall, Inc.
- [Moon et Canady, 1995] Moon, J. B. et Canady, J. W. (1995). Effects of Gravity on Velopharyngeal Muscle Activity During Speech. *Cleft Palate Craniofacial Journal*, 32(5) :371–375.
- [Moon et al., 1994] Moon, J. B., Smith, A. E., Folkins, J. W., Lemke, J. H., et Gartlan, M. (1994). Coordination of Velopharyngeal Muscle Activity during Positioning of the Soft Palate. *Cleft Palate-Craniofacial Journal*, 31 :45–55.
- [Mrayati, 1976] Mrayati, M. (1976). *Contribution aux études sur la production de la parole - Modèles électriques du conduit vocal avec pertes, du conduit nasal et de la source vocale - Études de leurs interactions - Relations entre disposition articulatoire et caractéristiques acoustiques*. Thèse d'État, Université Scientifique et Médicale de Grenoble, Institut National Polytechnique de Grenoble, France.
- [Narayanan et al., 2004] Narayanan, S., Nayak, K., Lee, S., Sethy, A., et Byrd, D. (2004). An approach to real-time magnetic resonance imaging for speech production. *Journal of the Acoustical Society of America*, 115(4) :1771–1776.

- [Netter, 1997] Netter, F. H. (1997). *Atlas of Human Anatomy*. Hoechstetter Printing Company Inc., USA, 2nd edition.
- [Parush et Ostry, 1986] Parush, A. et Ostry, D. J. (1986). Superior lateral pharyngeal wall movements in speech. *Journal of the Acoustical Society of America*, 80(3) :749–756.
- [Passavant, 1869] Passavant, G. (1869). Ueber die verschliessung des Schlundes beim Sprechen. *Virchows Archiv*, 46(1) :1–31.
- [Payan et Perrier, 1997] Payan, Y. et Perrier, P. (1997). Synthesis of V-V sequences with a 2D biomechanical tongue model controlled by the equilibrium point hypothesis. *Speech Communication*, 22(2-3) :185–205.
- [Perkell, 1974] Perkell, J. S. (1974). *A Physiologically-Oriented Model of Tongue Activity in Speech Production*. PhD thesis, MIT, Cambridge, USA.
- [Perkell et al., 1992] Perkell, J. S., Cohen, M. H., Svirsky, M. A., Matthies, M. L., naki Garabieta, I., et Jackson, M. T. T. (1992). Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements. *Journal of the Acoustical Society of America*, 92(6) :3078–3096.
- [Perrier et al., 1992] Perrier, P., Boë, L. J., et Sock, R. (1992). Vocal Tract Area Function Estimation From Midsagittal Dimensions With CT Scans and a Vocal Tract Cast : Modeling the Transition With Two Sets of Coefficient. *Journal of Speech and Hearing Research*, 35(1) :53–67.
- [Poppelreuter et al., 2000] Poppelreuter, S., Engelke, W., et Bruns, T. (2000). Quantitative Analysis of the Velopharyngeal Sphincter Function During Speech. *Cleft Palate-Craniofacial Journal*, 37 :157–165.
- [Ramsay, 2001] Ramsay, G. (2001). Communication personnelle (courriel).
- [Reenen, 1982] Reenen, P. (1982). *Phonetic Feature Definitions : Their integration into phonology and their relation to speech. A case study of the feature NASAL*, chapter Chapter 5 : A survey of data on nasality, pages 55–68. Foris Publications, Dordrecht, Holland - Cinnaminson, USA.
- [Rokkaku et al., 1986] Rokkaku, M., Hashimoto, K., Imaizumi, S., Niimi, S., et Kiritani, S. (1986). Measurement of the three-dimensional shape of the vocal tract based on the Magnetic Resonance Imaging technique. *Annual Bulletin of the Research Institute for Logopedics and Phoniatics, Tokyo, Japan*, 20 :47–54.
- [Rossato, 2000] Rossato, S. (2000). *Du son au geste, inversion de la parole : le cas des voyelles nasales*. Mémoire de doctorat, ICP - INPG, Grenoble, France.
- [Rossato et al., 2003] Rossato, S., Badin, P., et Bouaouni, F. (2003). Velar movements in French : An articulatory and acoustical analysis of coarticulation. In *Proceedings of the 15th International Congress of Phonetic Sciences*, pages 3141–3144, Barcelona, Spain.
- [Rubin et al., 1981] Rubin, P., Baer, T., et Mermelstein, P. (1981). An articulatory synthesizer for perceptual research. *Journal of the Acoustical Society of America*, 70(2) :321–328.

- [Shadle *et al.*, 1999] Shadle, C. H., Mohammad, M., Carter, J. N., et Jackson, P. J. (1999). Multi-planar dynamic Magnetic Resonance Imaging : new tools for speech research. In *Proceedings of the 14th International Congress of Phonetic Sciences*, pages 623–626, San Francisco, USA.
- [Skolnick *et al.*, 1973] Skolnick, M. L., McCall, G. N., et Barnes, M. (1973). The sphincteric mechanism of velopharyngeal closure. *Cleft Palate Journal*, 10 :286–305.
- [Sobotta, 2001] Sobotta, J. (2001). *Sobotta : Atlas of Human Anatomie. Volume 1 : Head, Neck, Upper Limb (R.Putz and R.Pabst Eds. - 13th English Edition)*. Lippincott Williams and Wilkins.
- [Soquet *et al.*, 2002] Soquet, A., Lecuit, V., Metens, T., et Demolin, D. (2002). Mid-sagittal cut to area function transformations : Direct measurements of mid-sagittal distance and area with MRI. *Speech Communication*, 36(3-4) :169–180.
- [Stefanuto et Vallée, 1999] Stefanuto, M. et Vallée, N. (1999). Consonant systems : from universal trend to ontogenesis. In *Proceedings of the International Congress of Phonetic Sciences*, volume 3, pages 1973–1976, San Francisco, USA.
- [Stevens, 1998] Stevens, K. N. (1998). *Acoustic Phonetics*. The MIT Press, Cambridge, USA, London, England.
- [Stevens *et al.*, 1953] Stevens, K. N., Kasowski, S., et Fant, C. G. M. (1953). An Electrical Analog of the Vocal Tract. *Journal of the Acoustical Society of America*, 25(4) :734–742.
- [Stone et Lundberg, 1996] Stone, M. et Lundberg, A. (1996). Three-dimensional tongue surface shapes of English consonants and vowels. *Journal of the Acoustical Society of America*, 99(6) :3728–3737.
- [Story *et al.*, 1996] Story, B., Titze, I., et Hoffman, E. (1996). Vocal tract area functions from Magnetic Resonance Imaging. *Journal of the Acoustical Society of America*, 100 :537–554.
- [Sundberg *et al.*, 1987] Sundberg, J., Johansson, C., Wilbrand, H., et Ytterbergh, C. (1987). From sagittal distance to area. A study of transverse, vocal tract cross-sectional area. *Phonetica*, 44 :76–90.
- [Szeliski et Lavallée, 1996] Szeliski, R. et Lavallée, S. (1996). Matching 3-D Anatomical Surfaces with Non-Rigid Deformations using Octree-Splines. *International Journal of Computer Vision*, 18(2) :171–186.
- [Takemoto *et al.*, 2004] Takemoto, H., Kitamura, T., Nishimoto, H., et Honda, K. (2004). A method of teeth superimposition on MRI data for accurate measurement of vocal tract shape and dimensions. *Acoustical Science and Technology*, 25 :468–474.
- [Teixeira *et al.*, 2003] Teixeira, A., Moutinho, L., et Coimbra, R. (2003). Production, acoustic and perceptual studies on european Portuguese nasal vowels height. In *Proceedings of the 15th International Congress of Phonetic Sciences*, pages 3033–3036, Barcelona, Spain.
- [Teixeira *et al.*, 2001] Teixeira, A., Vaz, F., Moutinho, L., et Coimbra, R. (2001). Articulatory synthesis of Portuguese. In *III Encontro do Forum Internacional de Investigadores Portugueses*, Institut EETA, Aveiro, Portugal.

- [Tiede *et al.*, 2000] Tiede, M. K., Masaki, S., et Vatikiotis-Bateson, E. (2000). Contrasts in speech articulation observed in sitting and supine conditions. In *Proceedings of the 5th Seminar on Speech Production : Models and Data - CREST Workshop on Models of Speech Production : Motor Planning and Articulatory Modelling*, pages 25–28, Kloster Seeon, Germany.
- [Vallée, 1994] Vallée, N. (1994). *Systèmes vocaliques : de la typologie aux prédictions*. PhD thesis, Université Stendhal, Grenoble, France.
- [Wilhelms-Tricarico, 1995] Wilhelms-Tricarico, R. (1995). Physiological modeling of speech production : Methods for modeling soft-tissue articulators. *Journal of the Acoustical Society of America*, 97(5) :3085–3098.
- [Wrench, 1999] Wrench, A. (1999). An investigation of sagittal velar movement and its correlation with lip, tongue and jaw movement. In *Proceedings International Congress of Phonetic Sciences, San Francisco, USA*, pages 2259–2262, Queen Margaret University College, Edinburgh, UK.
- [Yehia et Tiede, 1997] Yehia, H. et Tiede, M. (1997). A parametric three-dimensional model of the vocal-tract based on MRI data. In *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, pages 1619–1622, Munich, Germany.
- [Zemlin, 1968] Zemlin, W. R. (1968). *Speech and Hearing Science. Anatomy and Physiology*. Prentice-Hall, Inc.
- [Zerling, 1984] Zerling, J.-P. (1984). Phénomènes de nasalité et de nasalisation vocaliques : étude cinéradiographique pour deux locuteurs. *Travaux de l'Institut de Phonétique de Strasbourg*, 16 :241–266.

RÉSUMÉ

Titre Modélisation tridimensionnelle des organes de la parole à partir d'images IRM pour la production de nasales – Caractérisation articulatoire-acoustique des mouvements du voile du palais.

Ce travail a pour objectif la caractérisation articulatoire-acoustique de la nasalité : nature des mouvements du port vélopharyngé, caractéristiques acoustiques liées. La construction d'un modèle articulatoire linéaire 3D monosujet du conduit nasal à partir d'images IRM et CT a fait émerger deux degrés de liberté parmi les mouvements du voile du palais et de la paroi nasopharyngée. Le premier, prédominant, correspond à un mouvement conjoint vertical oblique du voile et horizontal de la paroi pharyngée, traduisant l'effet de sphincter du port vélopharyngé, et le second à un petit mouvement horizontal du voile seul, modifiant sensiblement l'aire de couplage nasal. L'espace des mouvements du modèle décrit exactement celui d'un point du voile obtenu par articulographie électromagnétique. Les fonctions d'aire réalistes du conduit nasal déduites du modèle ont permis de déterminer les fonctions de transfert acoustiques de sept voyelles et l'influence acoustique des mouvements du voile.

Mots-clés Production de la parole, Modélisation articulatoire, 3D, Voile du palais, Nasal, Acoustique, IRM, Conduit vocal

ABSTRACT

Title Three-dimensional modelling of the speech organs from MRI images for nasals production – Articulatory-acoustic characterization of the velum movements.

The goal of this work was the articulatory-acoustic characterization of nasality : the nature of velopharyngeal port movement and associated acoustic characteristics. The building of a 3D linear single-subject articulatory model of the nasal tract from MRI and CT images revealed two degrees of freedom for the movements of the velum and nasopharyngeal wall. The dominant parameter corresponds to a joint vertical oblique movement of the velum and a horizontal movement of the pharyngeal wall, expressing the sphincter effect of the velopharyngeal port. The second parameter corresponds to a smaller movement of the velum, significantly modifying the nasal coupling area. The space covered by the model's movements corresponds precisely to that described by a velum point measured by electromagnetic articulography. Realistic area functions of the nasal tract derived from the model allowed us to determine acoustic transfer functions for seven vowels and the acoustic influence of velar movements.

Keywords Speech production, Articulatory modeling, 3D, Velum, Nasal, Acoustics, MRI, Vocal tract
