

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université des Sciences et de la Technologie Houari Boumediene
Faculté d'Electronique et d'Informatique



THESE

Présentée pour l'obtention du diplôme de DOCTORAT D'ETAT
EN: ELECTRONIQUE
Spécialité : Electronique des systèmes

Par : Mme FALEK née ALI BENALI Djouher Leila

THEME

***Etude du contrôle temporel multi-contextuel
du signal de parole***

Soutenue le 18 décembre 2008, devant le jury composé de :

A.HOUACINE	PRESIDENT	Professeur, USTHB
A. DJERADI	DIRECTEUR DE THESE	Professeur, USTHB
M.GUERTI	EXAMINATRICE	Professeur, ENP
B.BOUDRAA	EXAMINATEUR	Maître de conférences, USTHB
M.DEBYECH	EXAMINATEUR	Maître de conférences, USTHB
F.MERAZKA	EXAMINATRICE	Maître de conférences, USTHB

A la mémoire de mon père, ancien élève de l'école polytechnique de Paris et ingénieur en électronique. Tombé au champ d'honneur pendant la guerre de libération, il n'a pas pu nous voir grandir.

A ma mère, qui a sacrifié toute sa jeunesse pour nous.

REMERCIEMENTS

Le travail présenté dans ce mémoire de thèse a été effectué au laboratoire de communication parlée et traitement des signaux (LCPTS) de la faculté d'électronique et d'Informatique de l'USTHB.

Que Monsieur le Professeur A.Djeradi, directeur du laboratoire de recherche LCPTS, soit assuré de ma vive et sincère reconnaissance, pour avoir accepté de diriger ce travail. Les conseils fructueux et les encouragements qu'il n'a cessé de me prodiguer m'ont été très précieux tant au cours du développement de mes recherches que dans la rédaction de cette thèse. J'ai été particulièrement touchée par son humanisme et la bienveillance qu'il m'a toujours prodiguée.

Je remercie vivement Monsieur le Professeur A.Houacine, chef d'équipe au LCPTS et enseignant chercheur de la faculté d'électronique et d'Informatique de l'USTHB, d'avoir accepté de présider ce jury et pour ses encouragements.

J'exprime ma profonde gratitude au professeur M.Guerti de l'Ecole Polytechnique d'El Harrach. Je suis très touchée de l'intérêt qu'elle a porté à ma recherche en acceptant d'examiner cette thèse. Le Professeur M.Guerti a été mon enseignante en graduation et aujourd'hui elle m'honore de sa présence pour clôturer ma post graduation. Qu'elle trouve ici toute ma reconnaissance et mon meilleur souvenir.

Mes vifs remerciements vont vers M. Debyeche, Maître de conférences, chercheur au LCPTS et enseignant de la faculté d'électronique et d'Informatique de l'USTHB, d'avoir accepté de participer au jury de cette thèse. Je tiens à lui exprimer toute ma reconnaissance pour m'avoir continuellement encouragé.

Monsieur B.Boudraa, Maître de conférences, chercheur au LCPTS et enseignant de la faculté d'électronique et d'Informatique de l'USTHB, m'a encouragé tout au long de ma recherche. Il m'a prodigué des conseils précieux qui m'ont beaucoup servi dans mes recherches. Qu'il trouve ici ma profonde reconnaissance d'avoir en plus accepté de faire partie du jury de cette thèse.

J'exprime toute ma reconnaissance à F.Merazka, Maître de conférences, chercheur au LCPTS et enseignant de la faculté d'électronique et d'Informatique de l'USTHB, d'avoir accepté d'examiner cette thèse. Je la remercie vivement aussi pour ses encouragements.

Je tiens à exprimer toute ma reconnaissance à Hocine Teffahi, Maître de conférences, chercheur au LCPTS et enseignant de la faculté d'électronique et d'Informatique de l'USTHB, pour avoir accepté d'entreprendre des travaux de recherche avec moi qui ont contribué à l'aboutissement de cette thèse. J'espère que notre collaboration continuera d'être fructueuse. Je le remercie vivement pour ses encouragements, ses conseils et son soutien moral.

Je n'oublierai pas de remercier le professeur Yves Laprie. Chercheur à LORIA- Nancy. De m'avoir aidée à exploiter mes données en mettant à ma disposition son logiciel winsnoori. Qu'il trouve ici toute ma reconnaissance

Mes recherches ont été facilitées par l'excellente ambiance de travail et de coopération existant au sein du laboratoire LCPTS. J'en sais gré à tous mes camarades chercheurs. Je n'oublierai pas d'adresser également mes remerciements à tous mes collègues et personnel de la faculté d'électronique et d'Informatique qui m'ont sans cesse encouragée et aidée.

SOMMAIRE

INTRODUCTION.....	1
-------------------	---

Chapitre 1- Etat de l’art- Les expériences de perturbation de la production de la parole

1.1. Introduction	3
1.2. L’émergence des expériences perturbants la parole et le phénomène d’articulation compensatoire.....	4
1.3. La perception et les expériences perturbant la parole	4
1.3.1 Quatre théories de la perception de la parole	5
1.4 Les expériences de perturbation de la parole	6
1.4.1 Les expériences avec une plaque occlusale.....	6
1.4.2 Les expériences avec un tube labial	11
1.5 Le rôle du feedback dans la production de la parole.....	14
1.6 Conclusion	15
1.7 Bibliographie du chapitre 1	16

Chapitre 2 - Etude expérimentale : réalisation du matériau linguistique, analyse des données et détermination de patrons temporels des fricatives

2.1 Introduction	17
2.2 Les fricatives pharyngales de l'arabe.....	17
2.3 Le matériau linguistique.....	22
a) Particularité du corpus.....	22
b) Enregistrement des phrases	23
c) Analyse des segments de parole : durées relatives et absolues.....	23
2.4 Présentation des résultats	24
2.5 Analyse des durées par l'analyse de la variance.....	25
2.5.1 Méthode ANOVA	26
2.5.2 Principe général de la méthode	26
2.6 Analyse de l’influence du contexte vocalique par locuteur	31
a- Locuteur FE.....	31
a-1- fricative /a_/	31
a-2- fricative /h/.....	32
a-3 fricative /h_/	33
b- Locuteur CH.....	34
b-1- fricative /a_/.....	34
b-2- fricative /h/.....	35
b-3- fricative /h_/.....	46

c- Locuteur SA	37
c-1- fricative /a_/	37
c-2- fricative /h/	38
c-3- fricative /h_/	39
2.7 Etude de l'influence du contexte consonantique, pour un même contexte vocalique pour un même locuteur, sur les unités de parole analysées	40
a- Locuteur FE	40
a-1- contexte vocalique /a/	40
a-2- contexte vocalique /i/	41
a-3- contexte vocalique /u/	42
b- Locuteur CH	43
b-1- contexte vocalique /a/	43
b-2- contexte vocalique /i/	44
b-3- contexte vocalique /u/	45
c- Locuteur SA	46
c-1- contexte vocalique /a/	46
c-2- contexte vocalique /i/	47
c-3- contexte vocalique /u/	48
2.8 Discussion des résultats	49
2.8.1 Récapitulatif des résultats pour l'étude de l'influence du contexte vocalique sur les durées	49
2.8.2 Récapitulatif des résultats pour l'étude de l'influence du contexte consonantique sur les durées	51
2.8.3. Conclusion du chapitre	53
a. Influence du contexte vocalique	53
b. Influence du contexte consonantique	53
2.9 Bibliographie du chapitre 2	54

Chapitre 3- La variation de débits et le contrôle des gestes de la parole

3.1 Introduction	55
3.2. Méthodologie	55
3.2.1. Les unités de parole analysées:	56
3.3. Analyse des durées relatives d(CjViL)/ d(phrase)	56
3.3.1 Résultats obtenus pour les moyennes des durées relatives	56
3.3.2 Interprétation des résultats	59
3.3.3. Discussions :	61
3.4. Analyses acoustiques des voyelles arabes Vi (i=/a/, /i/, /u/) des unités /CjVil/, en contrainte de débit.	62
3.4.1 Méthodologie	63
3.4.1.1 Tracé des ellipses de dispersion	63
3.4.1.2. Tracé des triangles vocaliques	73
3.4.1.3. Calcul de la surface des triangles vocaliques	73
3.4.1.4. Résultats des surfaces des triangles vocaliques	83
3.5 Calcul des variations des formants F1 et F2 en fonction du débit d'élocution	88
3.6 Conclusion	91
3.7 Bibliographie du chapitre 3	93

Chapitre 4- Simulation de phrases à débits variables par un modèle de production de la parole

4.1	Introduction	95
4.2	Le synthétiseur à formants de Klatt	97
4.2.1-	Description de la structure du SAF de KLATT	97
4.2.2	Réalisation du SAF	99
4.2.2.1	La commande du synthétiseur	99
a)	Matrice de paramètres de la source pour la simulation des voyelles	101
b)	Matrice de paramètres de la source pour la simulation des plosives :	101
c)	Matrice de paramètres de la source pour la simulation des nasales :	102
d)	Matrice de paramètres de la source pour la simulation des liquides	102
e)	Matrice de paramètres de la source pour la simulation des fricatives	103
f)	Matrice de paramètres de la source pour la simulation des fricatives sourdes	104
4.3	Etude de la production de la parole à débit variable par le SAF	109
4.3.1	Analyse fréquentielle d'une séquence de parole à débit variable	101
4.3.2	Production de parole à débit variable par le SAF	111
4.4	Etude du modèle à deux masses	112
4-4-1-	Description du modèle à deux masses	112
4-4-4-	Détermination des paramètres de commande du modèle à deux masses	116
4-4-5-	Variation de la fréquence fondamentale de vibration des cordes vocales en fonction des paramètres Pression et facteur masse/tension	118
4. 5	CONCLUSION	121
Bibliographie du chapitre 4		122
CONCLUSION GENERALE.....		124
Bibliographie Globale.....		126

TABLE DES FIGURES

Figure 1. 1 Comparaison des valeurs de déplacement (en mm) pour la lèvre supérieure, la lèvre inférieure et la mâchoire pour 3 conditions et 4 groupes d'âge.	9
Figure 1. 2: Corrélations des trajectoires cinématiques de la mâchoire (j), de la lèvre inférieure (ll) et de la lèvre supérieure (ul) par rapport à la moyenne des trajectoires du groupe adulte. ...	9
Figure 1. 3: Configuration expérimentale utilisée par Riordan (1977) afin d'étudier la protrusion des lèvres et le déplacement du larynx	12
Figure 2. 1: Tracé des films aux rayons X , fricative / ʃ / en traits pleins et / h / en discontinus, précédent respectivement les voyelles /a/, /i/, /u/ (Al Ani 70)	21
Figure 2. 2 : Tracé d'images obtenues par films aux rayons X, montrant une vue de profil d'un locuteur articulant des consonnes pharyngales et uvulaires en position initiale devant /a/, /i/, /u/. La figure à gauche de chaque rangée montre le moment de la constriction maximale. Les figures à droite de chaque rangée correspondent aux figures vocaliques (Delattre , 71).....	21
Figure 2. 3: Formes du conduit vocal produisant :/ ʃ / en traits discontinus , / h / en traits pleins, b) / ɣ / en points tirés et / ʁ / en traits pleins . c) /q/. Ces 5 consonnes articulées dans le mot /Caeli/ où C l'une de ces consonnes (Ghazali 77)	22
Figure 2. 4 : a) Modélisation de la fonction d'aire avec une constriction pharyngale (la glotte est à gauche) , b) représentation des 4 plus basses fréquences en fonction de d (longueur de la cavité arrière) et de l'aire à la constriction (ligne discontinue $AC=0.1cm^2$) (Klatt et Stevens 69).....	23
Figures 2.5.a. 1 : Influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /a_/ (en rouge : contexte /a/, en vert contexte /i/ en bleu : contexte /u/)	33
Figures 2.5.a. 2: Influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h/	35
Figures 2.5.a. 3: Influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h_/	36
Figures 2.5.b. 1 : Influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /a_/	38
Figures 2.5.b. 2: Influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h/	39
Figures 2.5.b. 3: Influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h_/	41

Figures 2.5.c. 1: Influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /a_ /	42
Figures 2.5.c. 2 : Influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h/	44
Figures 2.5.c. 3: Influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h_ /	45
Figures 2.6.a. 1: Influence des contextes consonantiques /a_/, /h/, /h_ / , en contexte vocalique /a/	47
Figures 2.6.a. 2: Influence des contextes consonantiques /a_/, /h/, /h_ / , en contexte vocalique /i/	48
Figures 2.6.a. 3: Influence des contextes consonantiques /a_/, /h/, /h_ / , en contexte vocalique /u/	50
Figures 2.6.b. 1:Influence des contextes consonantiques /a_/, /h/, /h_ / , en contexte vocalique /a/	51
Figures 2.6.b. 2: Influence des contextes consonantiques /a_/, /h/, /h_ / , en contexte vocalique /i/	52
Figures 2.6.b. 3: Influence des contextes consonantiques /a_/, /h/, /h_ / , en contexte vocalique /u/	54
Figures 2.6.c. 1: Influence des contextes consonantiques /a_/, /h/, /h_ / , en contexte vocalique /a/	55
Figures 2.6.c. 2: Influence des contextes consonantiques /a_/, /h/, /h_ / , en contexte vocalique /i/	57
Figures 2.6.c. 3: Influence des contextes consonantiques /a_/, /h/, /h_ / , en contexte vocalique /u/	58
Figures 3. 1. Représentation des Variation des durées relatives des CjVil/phrasede	68
Figures 3. 2. Valeurs des F du test statistique en fonction du contexte consonantique pour chaque locuteur	71
Figures 3. 3. Axe de production de la parole selon la contrainte du débit	72
Figures 3. 4. Ellipses de dispersion relatives à chaque locuteur, chaque contexte consonantique et chaque débit.	76
Figures 3. 5. Triangles vocaliques tracé à partir des ellipses de dispersion, en fonction du débit pour chaque locuteur et chaque contexte consonantique	80
Figures 3. 6. Représentation du triangle vocalique pour le calcul de sa surface.	81
Figures 3. 7. Surface des triangles vocaliques	85
Figures 3. 8. Triangle vocalique neutre et selon	85
Figures 3. 9. Exemple de centralisation des patrons articulatori-acoustiques lors de la production des voyelles à l'étude	86
figure 4. 1 Diagramme bloc du synthétiseur de Klatt.	95
figure 4. 2 Schéma simplifié su SAF	96
figure 4. 3 Résonateur numérique	97
figure 4. 4 Onde de voisement et son spectre simulés sur Matlab (pour les voyelles).	100
figure 4. 5 Onde de voisement et son spectre simulés sur Matlab (pour les plosives).	100
figure 4. 6 Onde de voisement et son spectre simulés sur Matlab (pour les nasales).	100
figure 4. 7 Onde de voisement et son spectre simulés sur Matlab (pour les liquides).	101

figure 4. 8 source de friction (non voisée) et son spectre simulés sur Matlab (pour les liquides).....	101
figure 4. 9 Signaux de la source d'excitation simulés sur Matlab, pour les fricatives sonnantes (en haut : source de voisement et son spectre, en bas : source de friction et son spectre). ...	103
figure 4. 10 Signaux de la source d'excitation simulés sur Matlab (pour les fricatives sourdes).	109
figure 4. 11 $F1(t)$ et $F2(t)$ pour le locuteur	109
figure 4. 12 $F1(t)$ et $F2(t)$ pour le locuteur SA	109
figure 4. 13 $F0(t)$ à débit variable pour les deux locuteurs	113
figure 4. 14 Modèle à deux masses des cordes vocales.	115
figure 4. 15. Modèle électrique équivalent de la source vocale à deux masses couplé au conduit vocal et représentation en concordance de la distribution de la pression dans la glotte d'après Guérin.	118
figure 4. 16 Exemples de forme d'onde $U_g(t)$	118
figure 4. 17 F_0 en fonction de P_s pour différentes valeurs de Q (résultats de 68 simulations)	119
figure 4. 18 F_0 en fonction de Q pour différentes valeurs de P_s (résultats de 32 simulations)	120

TABLEAUX

Tableau 1. 1: Résumé des résultats des expériences à l'aide d'une plaque occlusale (sujets adultes)	10
Tableau 1. 2 : Résumé des résultats des expériences à l'aide d'une plaque occlusale (sujets enfants).....	11
Tableau 2. 1. Exemple de durées et de formants relevés sur les unités d'analyse considérées.	28
Tableau 2. 2 : Durées relatives des CVC sur la phrase	30
Tableau 2. 3. Vérification de la normalité à l'aide du Test Shapiro-Wilk	31
Tableau 2. 4 : Valeurs de F et de p dans l'ANOVA.....	31
Tableau 2. 5. exemple des résultats obtenus à l'aide du test de Scheffé à l'aide du logiciel ORIGINPRO70.....	32
Tableau 2.6.a. 1. Influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /a_/	34
Tableau 2.6.a. 2. Résultats obtenus pour l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h/	35
Tableau 2.6.a. 3. Résultats obtenus pour l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h/	37
Tableau 2.6.b. 1. Résultats obtenus pour l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h/	38
Tableau 2.6.b. 2: Résultats obtenus pour l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h/	40
Tableau 2.6.b. 3 Résultats obtenus pour l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h_/	41
Tableau 2.6.c. 1. Résultats obtenus pour l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /a_/	43
Tableau 2.6.c. 2. Résultats obtenus pour l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h/	44
Tableau 2.6.c. 3. Résultats obtenus pour l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h_/	46
Tableau 2.7.a. 1 Résultats obtenus pour l'influence des contextes consonantiques /a_/, /h/, /h_/ , en contexte vocalique /a/	47
Tableau 2.7.a. 2: Résultats obtenus pour l'influence des contextes consonantiques /a_/, /h/, /h_/ , en contexte vocalique /i/	49
Tableau 2.7.a. 3: Résultats obtenus pour l'influence des contextes consonantiques /a_/, /h/, /h_/ , en contexte vocalique /u/.....	50

Tableau 2.7.b. 1: Tableau des résultats obtenus pour l'influence des contextes consonantiques /a_/, /h/, /h_/, en contexte vocalique /a/.....	52
Tableau 2.7.b. 2: résultats obtenus pour l'influence des contextes consonantiques /a_/, /h/, /h_/, en contexte vocalique /i/.....	53
Tableau 2.7.b. 3: Tableau des résultats obtenus pour l'influence des contextes consonantiques /a_/, /h/, /h_/, en contexte vocalique /u/.....	54
Tableau 2.7.c. 1: résultats obtenus pour l'influence des contextes consonantiques /a_/, /h/, /h_/, en contexte vocalique /a/.....	56
Tableau 2.7.c. 2: Résultats obtenus pour l'influence des contextes consonantiques /a_/, /h/, /h_/, en contexte vocalique /i/.....	57
Tableau 2.7.c. 3: résultats obtenus pour l'influence des contextes consonantiques /a_/, /h/, /h_/, en contexte vocalique /u/.....	59
Tableau 2.8.1. 1:Récapitulatif des résultats de l'influence du contexte vocalique sur les durées en présence de la fricative /a_/.....	59
Tableau 2.8.1. 2:Récapitulatif des résultats de l'influence du contexte vocalique sur les durées en présence de la fricative /h/.....	60
Tableau 2.8.1.3 :Récapitulatif des résultats de l'influence du contexte vocalique sur les durées en présence de la fricative /h_/.....	60
Tableau 2.8.2.1: Récapitulatif des résultats de l'influence du contexte consonantique sur les durées en présence de la voyelle /a/.....	61
Tableau 2.8.2. 2:Récapitulatif des résultats de l'influence du contexte consonantique sur les durées en présence de la voyelle /a/.....	62
Tableau 2.8.2. 3:Récapitulatif des résultats de l'influence du contexte consonantique sur les durées en présence de la voyelle /a/.....	63
Tableau 3. 1. Variations moyennes des durées (dCjVil/dphrase).....	67
Tableau 3. 2 : Résultats de l'analyse de la variance (ANOVA) des durées relatives par locuteur.....	70
Tableau 3. 3. Calcul des surfaces vocalique en fonction du contexte consonantique , du locuteur et du débit.....	84
Tableau 3. 4. Variation des formants avec le débit, en fonction du contexte consonantique, pour chaque locuteur et centralisation.....	89
Tableau 3.5 : Nombre de cas de centralisation des voyelles (/a/, /i/, /u/) en fonction du locuteur.....	90
Tableau 4. 1 Les différents paramètres de commande du SAF.....	98
Tableau 4. 2 Paramètres de commande de la source pour la simulation des voyelles.....	99
Tableau 4. 3 Paramètres de commande de la source pour la simulation des plosives.....	99
Tableau 4. 4: Paramètres de commande de la source pour la simulation des nasales.....	100
Tableau 4. 5 : Paramètres de commande de la source pour la simulation des liquides.....	101
Tableau 4. 6: Paramètres de commande de la source pour la simulation des fricatives sonnantes.....	101
Tableau 4. 7 : Paramètres de commande de la source pour la simulation des fricatives sourdes.....	102
Tableau 4. 8: Matrice de commande pour la génération des trois voyelles de l'arabe: /a/, /i/, /u/.....	105

Tableau 4. 9 : Matrice de commande la génération de la syllabe /ba/	106
Tableau 4. 10: paramètres acoustiques (F1, F2 et F0) pour les débit normal (n) et rapide (r)	108
Tableau 4. 11: Domaine d'évolution des paramètres de commande	117
Tableau 4. 12 : Relations entre les deux paramètres de commande P_s et Q et la variable de sortie F_0	119

INTRODUCTION

Le temps est un aspect fondamental du contrôle moteur de la parole. C'est un paramètre très important dans les systèmes de synthèse et de reconnaissance automatique de la parole.

Les avis des chercheurs par rapport à la conception de la parole et à son organisation temporelle sont souvent divergents. Certains pensent que l'organisation temporelle provient d'interactions de structures articulatoires périphériques et non d'un contrôle neurocognitif central. Les évidences et les contre-arguments relatifs à cette conception sont soumis à un examen critique. Il est indiqué que les évidences neurolinguistiques appuient minimalement une implication du cortex et du cervelet dans l'organisation temporelle de la parole.

Ensuite il est montré comment une telle notion de contrôle central peut être incorporée à une conception multifactorielle du contrôle moteur de la parole.

Cette approche est illustrée en regard de la distinction établie par Lindblom entre parole «hypo» et «hyper».

Sous une forme ou sous une autre, la plupart des investigations concernant la production de la parole traite du temps ou de l'organisation temporelle. Au niveau sémiotique le plus élémentaire, la parole implique une sériation temporelle d'éléments linguistiques. Au niveau phonémique, les voyelles, les fricatives, les occlusives ainsi que les autres segments occupent des espaces-temps spécifiques. Et sur le plan de l'analyse de la production phonétique, les débuts et les fins de différents segments de l'énoncé sont chronométrés selon des interdépendances coarticulatoires complexes qui peuvent changer considérablement selon les occasions, les locuteurs, les débits ou encore selon les dialectes. Dans son ensemble, la production de la parole semble être sujette à un schème temporel particulier, voire complexe.

Notre but dans cette étude est d'analyser dans la parole continue naturelle, les paramètres du timing de la production des sons (ici trois fricatives de l'arabe).

Les fricatives de l'arabe analysées seront prises en contexte vocalique arabe (/a/, /i/, /u/).

Notre intérêt est de comprendre les phénomènes mis en jeu dans la production de la parole, qui relèvent de l'inné ou simplement de l'habilité du locuteur.

Nous allons voir qu'en faisant varier le débit, nous pourrions relever sur le comportement du locuteur, les variations des principales caractéristiques des gestes articulatoires et définir ainsi l'espace vocalique contraint.

Notre étude consiste en une analyse des phénomènes mis en jeu lors de la production de la parole contrainte. Le paradigme utilisé est la variation de débit de locution.

Ce dernier permettra d'opposer la stratégie du locuteur relevant des gestes appris de ceux construits.

Pour cela, nous avons enregistré un corpus de 1620 phrases en arabes littéral, renfermant trois fricatives en contexte vocalique variable, et d'élocution à débit variable. Ces phrases ont été produites par six locuteurs différents par leurs origines géographiques : deux libanais, deux algérois et deux kabyles.

Par ailleurs, dans le souci d'une étude plus approfondie sur les effets du débit en production de la parole, nous avons entrepris de produire de la parole à débit variable artificiellement.

Nous avons mis en œuvre un modèle acoustique de production de la parole, le modèle à formants de Klatt.

Les paramètres de commandes de ce dernier sont principalement la fréquence fondamentale et les formants. Afin de pouvoir générer une mélodie liée à un paramètre très présent en variation de vitesse d'élocution (la pression sub-glottique), nous avons remplacé le modèle de source de Klatt par ce lui d'Ishizaka (le modèle à deux masses des cordes vocales).

L'étude et l'analyse de ce dernier nous ont permis d'obtenir une dynamique de la fréquence fondamentale en fonction de paramètres très réalistes liés aux cordes vocales comme la masse, l'étirement des cordes vocales et la pression sub-glottique de l'air en provenance des poumons.

Enfin, il s'agirait alors, d'introduire dans la commande du synthétiseur de Klatt la pression sub-glottique et attaquer le conduit vocal par l'onde de débit générée par le modèle à deux masses.

C'est ainsi que le chapitre 1 relate l'état de l'art de notre travail. C'est une introduction aux études sur la compréhension de la production de la parole et aux expériences qui permettent d'apporter des éléments nouveaux dans ce domaine. Il permet de poser la problématique et le but recherché.

Dans le chapitre 2, nous décrivons notre étude expérimentale du phénomène suivie d'une analyse des durées d'unités de parole contenant les fricatives arabes considérées en contexte vocalique. Il s'agit alors de voir l'impact de ces sons sur le timing de la parole en condition normale de débit avant de passer à la variation de la vitesse d'élocution.

Dans le chapitre 3, nous donnerons des notions de débit dans le contrôle des gestes de la parole. Nous expliquons comment par le biais de la contrainte de débit d'élocution nous arrivons à dégager des patrons temporels et acoustiques qui ont permis aux différents locuteurs d'atteindre leurs cibles acoustiques.

Chapitre 4 : Ce chapitre intitulé « Simulation de phrases à débits variables par un modèle de production de la parole » est une ébauche à une étude entreprise dans le but de simuler les différents gestes articulatoires en contrainte de débit par le biais d'un modèle acoustique de production.

Conclusion Générale de l'étude : Nous relatons ici les différentes étapes de notre travail suivies d'une discussion autour des résultats et des perspectives pour compléter cette étude.

Chapitre 1

Etat de l'art

Les expériences de perturbation de la production de la parole

1.1. Introduction

L'activité de parole est un processus dynamique complexe qui implique l'intégration et la synchronisation de tout un ensemble de traitements psycholinguistiques, linguistiques, moteurs et perceptifs.

L'organisation temporelle de la parole est également conditionnée par les habitudes sociolinguistiques. Tout d'abord, chaque langue dispose d'un système de sons et d'un système prosodique qui lui sont propres. Par exemple, une des caractéristiques de chaque langue, voire de chaque dialecte, est celle des *durées* des sons. Les durées caractéristiques des sons pour l'anglais sont différentes de celles pour le français. Ensuite, au-delà de ces systèmes, des standards socioculturels tels que la vitesse de parole exercent une pression supplémentaire sur les performances des locuteurs (Miller et al. 1976 [1]; Street & Brady, 1982 [2])

La dimension temporelle de la parole est aussi affectée par les événements pragmatiques et affectifs. Dans ce domaine, Caelen-Haumont, (1991) [3], a suggéré comment, en lecture, le passage d'un groupe de mots à un autre, et la cohésion entre les mots dépendent en partie de la signification du texte, des intentions du locuteur, de la perception de la situation, et des besoins de l'auditeur. De plus, une intention affective spécifique peut altérer la dynamique temporelle de la parole, du moins momentanément, du fait de changements dans la tension et l'élasticité des cordes vocales, ou dans le rythme respiratoire (Scherer, 1984, 1986, 1989).

Une perturbation dans l'organisation temporelle de la parole ne se produit pas seulement lors de difficultés rencontrées dans la gestion des contraintes biopsychologiques. Elle peut aussi se manifester lorsqu'un locuteur ne maîtrise pas correctement, par exemple, une langue seconde. Dans ce cas, le rythme peut être bizarre, ou de nombreuses hésitations peuvent être produites. Enfin, des phénomènes de diffluences et de perte momentanée de la fluidité peuvent également apparaître lorsqu'un locuteur se trouve dans une situation de communication difficile.

Nous voyons que comprendre le phénomène de production de la parole n'est pas toujours chose aisée. Un paradigme de recherche s'est développé au cours des années dans le but d'étudier les différents modules qui composent le modèle interne de production de la parole. Il s'agit des expériences de perturbation, qui étudient le phénomène d'articulation compensatoire. Ce dernier est une manifestation de la capacité d'équivalence motrice de l'humain, celle-ci permettant à l'homme d'atteindre une même cible spatiale avec une variété de trajectoires différentes, en faisant appel à différents groupes musculaires.

En effet, l'observation de personnes présentant des troubles du langage (bégaiement, parole rapide), troubles de la parole liée à la maladie de Parkinson (parole lente), troubles liés à des lésions cérébrales (parole inintelligible) etc ..., a permis de faire le lien entre l'effet et la cause de la perturbation. Ainsi, beaucoup de chercheurs ont créé artificiellement une

perturbation de la parole afin d'analyser les effets de cette dernière sur certains paramètres pertinents de la parole (variation des formants, réduction vocalique, variation du timing du signal de parole, variation de la coarticulation etc.) qui ont permis de mieux comprendre le phénomène de production de la parole. Un autre aspect important ressortant de ces expériences de perturbation de la production de la parole est le phénomène compensatoire que développe le locuteur pour parer à cette perturbation.

Notre travail s'inscrit dans la compréhension du contrôle temporel de la parole. Le paradigme utilisé est la perturbation par la contrainte de débit d'élocution. Un aperçu sur les expériences perturbantes de la parole nous aidera à mieux comprendre le rôle des différents feedback (auditif ou sensoriel en général) dans le processus de la production de la parole.

Nous allons donc décrire quelques expériences qui vont nous montrer en quoi ces perturbations peuvent apporter des éléments dans la compréhension de la production de la parole.

1.2. L'émergence des expériences perturbant la parole et le phénomène d'articulation compensatoire.

Lors de l'apprentissage de sa langue maternelle, l'enfant établit son modèle interne de production de la parole. Il met alors en relation les configurations articulatoires qu'il adopte aux sons qu'il produit. Les informations somatosensorielles, auditives et proprioceptives récupérées par l'enfant aident ce dernier à optimiser son modèle afin de renforcer les relations articulatoire-acoustiques qui y sont présentes.

Les expériences de perturbation de la production de la parole permettent d'étudier les différents modules qui composent le modèle interne de production de la parole et le phénomène d'articulation compensatoire.

Ce dernier est une manifestation de la capacité d'équivalence motrice de l'humain (par équivalence motrice on entend la capacité du système moteur à accomplir un objectif donné par divers moyens), celle-ci permettant à l'Homme d'atteindre une même cible spatiale avec une variété de trajectoires différentes, en faisant appel à différents groupes musculaires.

L'objectif principal étant d'étudier les effets acoustiques et articulatoires de perturbations labiales sur la production de la parole chez les enfants et les adultes. Cela permettra de quantifier l'efficacité des stratégies adoptées par les locuteurs et de préciser le rôle du feedback auditif dans le processus de production de la parole.

1.3. La perception et les expériences perturbant la parole

Les expériences perturbant la parole aident à préciser le rôle des différents types de feedback dans le processus de production de la parole, contribuant ainsi à mieux définir l'objectif de ce même processus. Cependant, les théories ne s'accordent pas toutes quant à la nature de l'objet qui est perçu lors de la production de la parole.

Avant de résumer les résultats des expériences utilisant différentes perturbations, il est donc nécessaire de faire un survol des théories contemporaines de la perception de la parole, pour une meilleure compréhension du phénomène. Il sera alors possible, ensuite, de se pencher sur ce que ces expériences peuvent apporter à ces théories.

1.3.1 Quatre théories de la perception de la parole

La recherche d'invariance dans la parole a été et est encore un débat des plus importants. De ce débat, plusieurs théories ont émergé, s'opposant ainsi sur la nature et la présence même de corrélats physiques perceptivement décodables dans la parole.

Pour certaines de ces théories (*Acoustic Invariance Theory*, *Adaptive Variability Theory*), la nature de l'objet perçu serait de type acoustique, tandis que pour d'autres (*Motor Theory*, *Direct-Realist Theory*), elle serait plutôt articulatoire.

a- La première théorie : théorie de l'invariance acoustique (*Acoustic Invariance Theory*), proposée par Blumstein (1986) [4] et Stevens et Blumstein (1981) [5] repose sur deux hypothèses :

- la première étant que le signal acoustique lui-même posséderait des caractéristiques acoustiques invariantes dans sa structure.
- La deuxième hypothèse est que le système perceptif serait sensible à ces propriétés invariantes et qu'il les utiliserait dans le processus de traitement de la parole.

Cette invariance acoustique reposerait sur les caractéristiques phonétiques des langues naturelles et serait indépendante du locuteur :

b- La deuxième théorie défendant la primauté des caractéristiques acoustiques sur les caractéristiques articulatoires est la théorie de variabilité adaptative (*Adaptive Variability Theory*) de Lindblom (1988 ; 1996). [6] [7]

Voici les 5 hypothèses de cette théorie appuyant l'idée que la parole serait plutôt perçue à l'aide de certaines caractéristiques acoustiques :

- L'auditeur et la situation de communication contribuent significativement à définir la tâche du locuteur;
- Cette tâche comporte des variations à court et à long terme;
- Le locuteur s'adapte à cette tâche;
- Cette adaptation ne se fait pas dans le but de faciliter le recouvrement articulatoire mais plutôt afin de produire un signal auditif devant toujours posséder suffisamment d'informations discriminantes;
- Dans le cas idéal, le degré de coarticulation du locuteur ne dépassera pas le seuil critique de tolérance de l'auditeur.

Cependant, Lindblom rejette l'idée de l'existence d'un invariant, puisque selon lui il ne s'agit pas d'un problème d'ordre phonétique.

L'invariance des catégories linguistiques résiderait seulement sur le plan de la compréhension du locuteur.

Ce serait donc ce comportement adaptatif du locuteur qui serait la cause du manque d'invariance dans le signal de la parole.

Malgré cela, il y aurait tout de même une quantité suffisante d'informations discriminantes dans le signal acoustique, permettant ainsi à un auditeur de décoder ce dernier.

Selon la *Motor Theory* (Liberman et Mattingly, 1985) [8] et la *Direct Realist Theory* (Fowler 1991; 1996) [9] [10], l'objet de la perception de la parole serait de type articulatoire.

Le principal élément sur lequel ces théories diffèrent est que la première traite la perception de la parole comme étant un système spécial de perception, tandis que la seconde la considère à l'intérieur du système perceptif global.

Pour Liberman et Mattingly, l'objet invariant de la perception de la parole se trouverait dans les commandes motrices contrôlant les mouvements des articulateurs. La parole serait donc décodée en retrouvant les gestes articulatoires produits par le locuteur à l'aide du système de contrôle moteur du locuteur.

Fowler croit également que cette perception est directe, et que les objets perceptifs sont les configurations articulatoires du conduit vocal. Seulement, cette perception ne requiert pas d'association entre les configurations du conduit vocal et un quelconque système de contrôle moteur de la parole.

Cela constitue une autre différence importante entre la théorie de Liberman et Mattingly et celle de Fowler, puisque l'une suggère que l'objet perçu est le geste intentionné, et l'autre le geste réellement produit par le locuteur.

Ces quatre théories diffèrent sur plusieurs points majeurs, discutables. Certains résultats expérimentaux apportent des appuis directs ou constituent des arguments en défaveur des hypothèses postulées ou des points soulevés dans ces théories. C'est le cas des expériences étudiant le phénomène d'équivalence motrice dans la parole.

En regard de ces théories, ces expériences sont particulièrement pertinentes puisque le fait de perturber la parole force le locuteur à réaliser la tâche du processus de production de la parole, et par conséquent à préciser la nature de l'objet perçu.

Le phénomène d'articulation compensatoire est donc en lien direct avec les théories de perception de la parole. De plus, les différents types de perturbations utilisées dans ces expériences font en sorte que les résultats peuvent nous donner différentes interprétations sur le rôle du feedback dans la production de la parole.

En apportant davantage de précisions sur les propriétés de l'objet de la perception de la parole et en caractérisant les rôles des différents types de feedback, les études portant sur la perturbation de la parole constituent des éléments importants dans la compréhension du contrôle moteur de la parole.

1.4 Les expériences de perturbation de la parole

Ces dernières font intervenir un facteur externe qui va perturber un articulateur du conduit vocal (plaque occlusale et tube labial).

1.4.1 Les expériences avec une plaque occlusale

Il est possible de perturber la parole de plusieurs façons, en modifiant ou non la géométrie du conduit vocal ou en interrompant un certain type de feedback (tactile, auditif, etc.).

Une méthode répandue pour étudier l'articulation compensatoire est celle de la plaque occlusale (*bite-block*), qui consiste à perturber un articulateur sans pour autant altérer la géométrie du conduit vocal.

L'objectif de la plaque occlusale est de placer la mâchoire en position fixe, bloquant ses mouvements verticaux. En contraignant la mâchoire de cette façon, on force le locuteur à recourir à des articulateurs non perturbés pour élaborer des stratégies articulatoires qui diffèrent de ses stratégies habituelles en parole non perturbée.

Ce qui ressort des premières expériences avec des plaques occlusales (Lindblom et Sundberg, 1971[11]; Lindblom et al., 1979 [12]; Gay et al., 1981[13]) est le caractère instantané de la compensation faite par le locuteur adulte. En plus d'être instantanée, cette compensation semble être de nature sélective (Gay et al., 1981), puisque des contraintes biomécaniques empêchent la reproduction des fonctions d'aires idéales. De plus, il semble que cela résulte également d'un choix du locuteur de compenser à certains endroits et non à d'autres.

Les auteurs proposent alors que la cible d'une voyelle serait codée en terme de spatialité sensorielle, ce qui n'implique pas nécessairement que la cible serait de nature acoustique seulement.

Quelques années plus tard, McFarland et Baum (1995) [14] ont démontré qu'en plus de ne pas être immédiates, les compensations peuvent être incomplètes. La nature incomplète de ces compensations a également été soulevée dans Baum et al. (1996) [15].

Les auteurs suggèrent donc que les stratégies compensatoires sont susceptibles de se développer avec le temps, nécessitant peut-être des corrections de type essais et erreurs. Aussi, McFarland et al. (1996) [16] ont fait ressortir une différence importante en ce qui a trait aux expériences de perturbation de la parole : la production des voyelles se trouve plus affectée par une modification d'une fonction orale (telle que la fixation de la mâchoire par une plaque occlusale) que par une modification d'une structure orale (telle que l'ajout d'un palais artificiel).

Les premières études acoustiques sur des sujets enfants ont fait ressortir l'incapacité de ceux-ci à élaborer des stratégies compensatoires optimales.

En effet, les travaux de Gibson et McPhearson (1980) [17] et Oller et MacNeilage(1983) [18] montrent que, bien que les enfants aient été capables d'adopter des stratégies compensatoires, celles-ci ne sont pas aussi efficaces que celles adoptées par les adultes.

En effet, les enfants qui ont participé à leurs expériences n'ont pas été capables de compenser complètement pour les voyelles produites avec une plaque occlusale. Les résultats obtenus par Baum et Katz (1988) [19] divergent de ces deux précédentes études, puisqu'ils montrent que dès l'âge de 4 ans, les enfants sont capables de produire des voyelles en condition perturbée acoustiquement comparables à celles en condition normale, et ce dès la première vibration périodique.

Baum et Katz (1988) soulignent qu'à un stade du développement où la stabilité formantique des voyelles n'est pas encore atteinte, l'habileté à compenser pour l'ouverture de la mâchoire est tout de même présente. Aussi, les auteurs rappellent qu'il n'est pas rare pour un jeune enfant de parler avec un objet ou de la nourriture dans la bouche, ce qui veut dire que celui-ci a déjà une certaine expérience dans l'articulation compensatoire.

Les résultats obtenus par Edwards (1992) [20] diffèrent de ceux de Baum et Katz (1988). En effet, Edwards a trouvé que les six enfants qui ont participé à son étude, n'éprouvant pas de problème de langage, n'ont pas parfaitement compensé lorsque l'ouverture de leur mâchoire était fixe et contrôlée par une plaque occlusale. De plus, certains des sujets ayant des problèmes d'ordre phonologique ont éprouvé des difficultés systématiques autant dans la production des voyelles que dans la production des consonnes.

Sur la base de ces observations, Edwards suggère que l'équivalence motrice de la parole se développe plus lentement que la capacité de produire des segments phonologiques perceptivement corrects. Cependant, le fait que les enfants soient capables d'élaborer des

stratégies articulatoires, bien que celles-ci ne soient pas optimales, tend au contraire à montrer que ceux-ci font tout de même preuve d'une certaine équivalence motrice.

Campbell (1999) [21] s'est également penchée sur le phénomène d'articulation compensatoire à l'aide du paradigme de la plaque occlusale.

Les participants de son étude étaient des adultes et des enfants de 4 ans, avec et sans problème d'audition.

Ses analyses acoustiques ont révélé que les 6 adultes avec une audition normale avaient effectué une compensation immédiate et substantielle.

Cependant, l'auteur a remarqué une grande différence inter-sujets en ce qui a trait aux productions acoustiques: certains sujets sont de bons compensateurs, tandis que d'autres produisent des stratégies compensatoires moins efficaces.

Aussi, Campbell a trouvé que pour les locuteurs enfants, le feedback auditif ne joue pas de rôle dans la réorganisation articulatoire, ce qui est plutôt surprenant, compte tenu du fait que les enfants ne possèdent pas encore de modèle interne des relations articulatoires-acoustiques robustes. Elle suggère donc que les enfants de cet âge se comportent comme les adultes, mais avec une plus grande erreur résiduelle dans leur production parce que leur connaissance des relations articulatoires-acoustiques est incomplète.

Certains résultats de ces expériences, notamment ceux de Baum et Katz (1988), seraient en accord avec les données articulatoires de Smith et McLean-Muse (1987). Ces derniers ont montré que la vélocité des articulateurs externes (mâchoire et lèvres) était plus grande chez les enfants que chez les sujets adultes lorsque les locuteurs parlent avec une plaque occlusale entre leurs dents, faisant ainsi ressortir le potentiel des enfants à faire une compensation semblable à celle des adultes.

Cependant, de récentes études proposent des résultats différents de ceux de Smith et McLean-Muse. En effet, bien que ces derniers aient obtenus des valeurs similaires pour les enfants et les adultes en termes de déplacement, de la lèvre inférieure et de la lèvre supérieure en condition normale, une étude des trajectoires cinématiques des articulateurs (Green et al., 2002 [22]) montre des différences significatives entre les trajectoires des enfants et celles des adultes.

Ces différences varient non seulement en fonction de l'âge, mais aussi en fonction des articulateurs. Ces résultats suggèrent donc un développement séquentiel non linéaire des articulateurs externes. Selon Green et al. Le contrôle de la mâchoire serait acquis plus tôt que celui des lèvres, et ceci serait spécifiquement relié à l'âge, en raison d'une maturation des mouvements articulatoires différente pour chaque articulateur.

D'après Ménard et al. (2006) [24], le fait qu'il y ait des différences articulatoires entre les enfants et les adultes en condition normale (sans perturbation) pourrait donc suggérer une évolution progressive du potentiel compensatoire des enfants. Ce potentiel pourrait être en partie relié à la maturation du contrôle de chaque articulateur, et à la maturation physiologique de l'articulateur lui-même.

1.4.2 Les expériences avec un tube labial

Les expériences avec un tube labial agissent sur une zone de constriction différente de celle ciblée par une plaque occlusale et apportent des données supplémentaires pertinentes en ce qui a trait au phénomène d'articulation compensatoire.

Il est difficile de vérifier si les résultats des études avec un tube labial présentent la même inconstance puisque seulement trois études ont été faites à l'aide de ce type de perturbation : deux avec des locuteurs adultes (Savariaux et. al., 1995 [27], 1999; Cocusse 1996[29]) et une avec des locuteurs enfants (Ménard et al., 2004[30]).

Cependant, bien avant ces expériences, Riordan (1977) [31] avait déjà eu l'idée de perturber la fonction d'aire du conduit vocal en contraignant la protrusion des lèvres, et ce afin d'étudier le déplacement vertical compensatoire du larynx lors de la production des voyelles [i],[y],[u],[ə],[ɛ] et bien qu'à l'époque les équipements d'acquisition de mesures articulatoires n'étaient pas aussi sophistiqués que de nos jours, il est intéressant de voir la configuration expérimentale adoptée (figure 1.3).

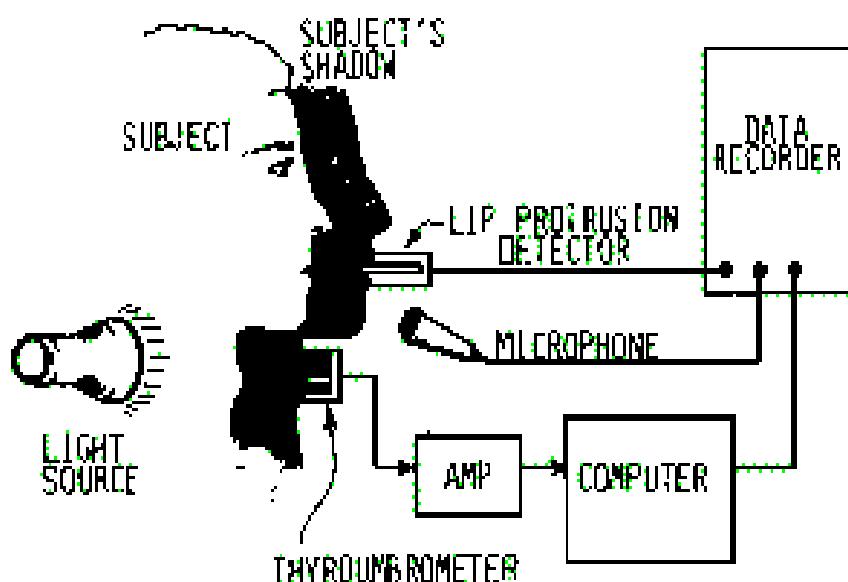


Figure 1. 1: Configuration expérimentale utilisée par Riordan (1977) afin d'étudier la protrusion des lèvres et le déplacement du larynx [31]

Riordan a utilisé un tube de métal de $\frac{1}{4}$ de pouce (0.63 cm) de diamètre pour forcer une protrusion lors de la production des voyelles.

Au lieu d'utiliser un tube labial de plus grand diamètre pour empêcher cette protrusion, il a placé une bande adhésive sur la lèvre supérieure.

Riordan a noté que la protrusion pour [u] était plus importante que pour [y] et qu'il n'y avait aucune différence significative entre les conditions sur le plan acoustique.

Les résultats de Riordan portent donc à croire que la position des lèvres n'est pas un élément crucial pour la production des voyelles arrondies. Suite à des simulations

informatiques, il a trouvé qu'un déplacement du larynx accompagné de changements dans la position de la langue pouvait compenser une perturbation d'aire aux lèvres.

C'est d'ailleurs ce qu'il a observé chez la majorité de ses locuteurs (4 sujets sur 6 abaissaient significativement le larynx pour compenser la perturbation).

Cependant, il indique que d'autres stratégies compensatoires peuvent être adoptées, comme par exemple l'ajustement de la position de la lèvre inférieure.

Certes les résultats obtenus par Riordan (1977) peuvent être contestables étant donné la technique utilisée pour modifier l'aire aux lèvres, mais il reste tout de même que ce fut une des premières études à utiliser un tube labial.

En utilisant un tube labial de 2.5cm de diamètre placé entre les lèvres, Savariaux et al.(1995,1999) ont fait prononcer la voyelle [u] dans cette condition de perturbation en recueillant des données acoustiques et articulatoires de 11 locuteurs adultes du français.

L'analyse cinéradiographique des données articulatoires couplées aux données acoustiques a permis de voir que les locuteurs n'ont pas tous compensé cette perturbation au même degré, et même qu'aucun locuteur n'a été capable de faire une compensation instantanée complète.

Cependant, à la suite d'une phase d'adaptation et d'une importante réorganisation articulatoire, un des locuteurs est parvenu à faire une compensation complète tandis que d'autres ont fait des améliorations marquées.

Les résultats des tests perceptifs montrent que même si la majorité des voyelles ont été bien identifiées, la qualité de celles-ci en condition perturbée a été jugée inférieure à celles produites en condition non perturbée.

Cependant, il semble que même si les locuteurs sont incapables de rencontrer les conditions nécessaires pour produire une certaine voyelle, ceux-ci ont tout de même une bonne représentation sensori-acoustique de celles-ci (Savariaux et al., 1999).

Dans le cas de la voyelle [u], cette représentation serait un espace acoustique composé de F0, F1 et F2. Les espaces acoustiques F2-F0 (en bark) et F1 seraient les valeurs qui caractériseraient le mieux la position haute et arrière de la langue lors de la production de cette voyelle par des adultes, permettant de bien faire ressortir le caractère grave (au sens fréquentiel : aigu et grave) de celle-ci.

Après avoir observé les résultats de Savariaux (1996)[27], notamment concernant les variations de la fréquence fondamentale en condition perturbée, Cocusse (1996)[29] a répété la même expérience en voulant stabiliser la fréquence fondamentale lors des productions.

Selon Cocusse, s'il y a variation de F0 pour compenser la perturbation, c'est un phénomène associé soit à la perturbation elle-même, ou alors associé à un contrôle réflexe au niveau du système central.

Parmi les 7 locuteurs ayant réussi à stabiliser leur fréquence fondamentale lors des répétitions, 3 ont réussi à compenser totalement à la suite d'une phase d'entraînement consistant en un guidage articulatoire à partir de la voyelle /o/. Le fait qu'un nombre relativement plus important de locuteurs ait réussi à compenser dans ces conditions par rapport à Savariaux et al. (1995, 1999) Suggère que la variation de la fréquence fondamentale serait un réflexe qui serait susceptible de nuire au locuteur dans l'élaboration de sa stratégie de compensation.

Donc, le fait d'effectuer un guidage articulatoire tout en stabilisant la fréquence fondamentale, a une influence directe sur l'efficacité compensatoire du locuteur, qui s'en trouve améliorée.

Ménard et al. (2004)[24] ont également guidé articulatoirement les sujets enfants de leur expérience lors de la production de la voyelle [u] avec un tube labial de 1.5 cm de diamètre.

Douze enfants ont participé à cette expérience. Même si les résultats des tests de production ne montrent pas de compensation totale pour la majorité des sujets, ceux des tests de perception indiquent que tous les enfants ont été capables de faire cette compensation puisque pour chaque sujet, au moins un des [u] produits en condition perturbée a été jugé comme étant de qualité égale ou supérieure aux [u] de la condition normale pré perturbation.

Cependant, l'utilisation du feedback faite par les enfants (productions de type essai/erreur) suggère que le modèle interne de ces derniers est incomplet. Toutefois, même si ces locuteurs n'ont pas terminé l'élaboration de leur modèle interne, il est intéressant de voir que quelques uns ont montré des signes d'effet post-perturbation.

Contrairement aux expériences avec une plaque occlusale, ces dernières études avec un tube labial montrent que, pour une perturbation qui affecte la fonction d'aire du conduit vocal, les compensations ne sont pas spontanées et requièrent un apprentissage.

Ces expériences renforcent l'hypothèse selon laquelle, lors de la production de la parole, les gestes articulatoires visent à atteindre une cible de nature acoustico-auditive (Perkell et al. 1993[33]; Perkell et al. 2000[34]).

1.5 Le rôle du feedback dans la production de la parole

Les études de compensations articulatoires lors de l'introduction d'un tube labial confirment l'importance du feedback auditif dans l'élaboration des stratégies articulatoires, contrairement à ce qui a été proposé par Campbell (1999).

Le rôle du feedback auditif a également été souligné par Jones et Munhall (2003) [34], qui ont modifié la géométrie du conduit vocal en augmentant la longueur des incisives supérieures à l'aide de prothèses dentaires.

Les productions de [s] et [ʃ] faites lorsque le feedback auditif était disponible ont été jugées meilleures que celles faites lorsque le feedback était masqué et un important effet d'apprentissage a été observé pour tous les locuteurs à toutes les sessions. Ces résultats démontrent que le signal acoustique aide les locuteurs à transposer les modifications nécessaires en termes de configuration articulatoire.

Tout comme Campbell (1999), Lane et al. (2005) [35] ont également utilisé le paradigme de la plaque occlusale pour étudier une population ayant des problèmes d'audition, c'est-à-dire les adultes atteints de surdité à un stade post-linguistique. Ils ont remarqué que la taille des espaces vocaliques (moyenne des distances euclidiennes entre toutes les paires possibles de voyelles) des locuteurs sourds avec un implant auditif étaient plus petits lorsque ce dernier était hors fonction que lorsqu'il fonctionnait.

Les mêmes observations ont été faites pour les conditions avec et sans plaque occlusale. Les résultats de Lane et al. (2005) confirment donc le rôle du feedback auditif et démontrent l'importance pour un modèle d'assigner un tel rôle au feedback auditif.

Ces résultats vont donc à l'encontre de Campbell (1999), qui minimise l'importance du feedback auditif dans la production de la parole.

Il n'y a pas seulement le feedback auditif qui joue un grand rôle dans ce modèle. La proprioception et les sensations somatosensorielles sont des éléments indispensables.

L'interruption de ces types de feedbacks est un moyen supplémentaire de perturber la parole et est l'élément sur lequel se sont penchées plusieurs études, entre autres Borden (1976) et Kelso et Tuller (1983) [39].

Borden (1976) [37] a étudié le rôle du feedback tactile chez les enfants en faisant une anesthésie partielle. Le nerf mandibulaire a ainsi été bloqué par anesthésie et les sensations tactiles de la langue ont été partiellement inhibées. Les résultats montrent que dès l'âge de 4 ans, les enfants détiennent un système de production de la parole pouvant fonctionner par moments sans même l'information tactile donnée par la langue, si l'information proprioceptive est disponible.

Kelso et Tuller (1983) [39] ont étudié les rôles des différents types de feedback lors de la production de la parole. Dans cette expérience, les locuteurs ont eu à prononcer des stimuli dans diverses conditions :

- 1 - avec une plaque occlusale entre les dents;
- 2 - sous anesthésie du joint mandibulaire, les privant ainsi de proprioception;
- 3 - sous anesthésie de la muqueuse orale, les privant ainsi de feedback tactile;
- 4 – sans feedback auditif.

Ce qu'ils ont observé peut surprendre, à savoir seulement 1 des 5 sujets n'a pas fait de compensation complète et immédiate, et ce, pour seulement une voyelle dans une seule condition. Les auteurs comparent cette capacité compensatoire à celle du système respiratoire, cette analogie rendant compte de la capacité compensatoire générative du corps humain.

1.6 Conclusion

Toutes ces études montrent bien que le feedback auditif n'est pas le seul à jouer un rôle dans le processus de production de la parole. Les autres types de feedback ne sont pas à négliger et sont des éléments dont il faut tenir compte lorsqu'il s'agit de faire un lien avec les théories de perception de la parole.

Cependant, une récente étude de Tremblay et al. (2002) a suggéré que le feedback auditif n'était pas l'élément principal qui guide le locuteur dans sa tâche.

En effet, à la suite de cette expérience où la mâchoire a été perturbée lors de la production de stimuli de parole et de stimuli non parole, les auteurs croient que leurs résultats représentent des différences dans la spécification des cibles dans le domaine articulatoire : la variation temporelle de l'information somatosensorielle lors des mouvements fait partie de l'objectif de la production de la parole, tandis que ce n'est pas le cas pour les mouvements en condition non-parole.

Ces observations confirment donc l'hypothèse d'une représentation articulatoire de l'objectif du processus de production de la parole dans le cerveau du locuteur.

De ces expériences, il est possible de constater qu'un seul type de feedback n'est pas suffisant pour aider le locuteur dans sa tâche et que la représentation de la tâche du locuteur n'est pas unimodale. Cette proposition va donc dans le sens de Perrier (2005) [40].

Les résultats de plusieurs expériences de perturbations avec une plaque occlusale ou un tube labial constituent des arguments de taille en faveur du feedback auditif comme étant le principal type de feedback. Aussi, tel que mentionné par Perrier (2005), aucune étude n'a

eu de résultats où les locuteurs acceptaient des changements acoustiques afin de préserver une certaine configuration articulatoire.

Finalement, pour conclure sur le rôle du feedback et son lien avec les théories de la perception, il serait bon de reporter les propos de Fowler 1996 [10], instigatrice de la *Direct Realist Theory*:

Ces différentes expériences montrent que le système nerveux est paré à toute perturbation des gestes de la parole. Le locuteur tentera de produire ses cibles acoustiques en empruntant des trajectoires différentes. Ces trajectoires sont liées à des paramètres divers qui peuvent être psycholinguistiques, linguistiques, moteurs, perceptifs ou autres. C'est ainsi que le locuteur arrive à se faire comprendre devant une contrainte.

1.7 Bibliographie du chapitre 1

- [1] Miller, N., Maruyana, G., Beaber, R.J. & Valone, K. (1976). Speed of speech and persuasion. *Journal of Personality and Social Psychology*, vol. 34. p 615-624.
- [2] Street, R. & Brady, R. (1982). Speech rate acceptance ranges as a function of evaluative domain, listener speech rate, and communication context. *Communication Monographs*, N° 4, p 290-308.
- [3] Caelen-Haumont, G. (1991). Stratégies des locuteurs et consignes de lecture d'un texte: Analyse des interactions entre modèles syntaxiques, sémantiques, pragmatique et paramètres prosodiques, Thèse d'Etat, Aix-en-Provence.
- [4] Blumstein, S. E. 1986. «On acoustic invariance in speech». In *Invariance and variability in speech processes*, sous la dir. de Joseph S. Perkell et Denis H. Klatt, p. 179-201. Hillsdale (NJ): Lawrence Erlbaum.
- [5] Stevens, K. N. et S. E. Blumstein. 1981. «The search for invariant acoustic correlates of phonetic features». In *Perspectives on the study of speech*, sous la dir. de Peter D. Eimas et Joanne L. Miller, p. 1-38. Hillsdale (NJ): Lawrence Erlbaum.
- [6] Lindblom, B. 1988. «Phonetic invariance and the adaptive nature of speech». In *Working Models of Human Perception*, sous la dir. de Ben A.G. Elsendoom & H. Bouma. p. 139- 173. London, UK: Academic Press.
- [7] Lindblom, B. 1996. «Role of articulation in speech perception: Clues from production». *Journal of the Acoustical Society of America*, vol. 99, no 3, p. 1683-792.
- [8] Liberman, A. M. et I. G. Mattingly. 1985. «The motor theory of speech perception revised». *Cognition*, vol. 21. p. 1-36.
- [9] Fowler, C. A. 1991. «Auditory perception is not special: We see the world, we feel the world, we hear the world». *Journal of the Acoustical Society of America*, vol. 89, no 6, p. 2910-2915.

- [10] Fowler, C. A. 1996. «Listeners do hear sounds, not tongues». *Journal of the Acoustical Society of America*, vol.99, no 3, p. 1730-1741.
- [11] Lindblom, B. et J. Sundberg. 1971. «Acoustical consequences of lip, tongue, jaw, and larynx movement». *Journal of the Acoustical Society of America*, vol. 50, no 4, p. 1166-1179.
- [12] Linblom, B., J. Lubker et T. Gay. 1979. «Formant frequencies of some fixed vowels and a model of speech motor programming by predictive simulation». *Journal of Phonetics*, vol. 7, p. 147-161.
- [13] Gay, T., B. Lindblom et J. Lubker. 1981. «Production of bite-block vowels : Acoustic equivalence by selective compensation». *Journal of the Acoustical Society of America*, vol. 69, no 3, p. 802-810.
- [14] D. H. McFarland, et S. R. Baum. 1995. «Incomplete compensation to articulatory perturbation». *Journal of the Acoustical Society of America*, vol. 97, no 3, p. 1865-1873.
- [15] S, R.Baum, D. H. McFarland et M. Diab. 1996. «Compensation to articulatory perturbation: Perceptual data». *Journal of the Acoustical Society of America*, vol. 99, no 6. p. 3791-3794.
- [16] D. H. McFarland, S. R. Baum et C. Chabot. 1996. «Speech compensation to structural modifications of the oral cavity». *Journal of the Acoustical Society of America*, vol. 100, no 2. p. 1093-1104.
- [17] A.Gibson et L. McPhearson 1980. «Production of bite-block vowels by children». *PERILUS Report*, vol. 2, p. 26-43
- [18] D. K.Oller, et P. MacNeilage. 1983. «Development of speech production : Perspectives from natural and perturbed speech». In *The production of speech, sous la dir. de Peter MacNeilage*, p.91-108. New-York: Springer-Verlag.
- [19] S. R.Baum et W. F. Katz. 1988. «Acoustic analysis of compensatory articulation in children». *Journal of the Acoustical Society of America*, vol. 84, no 5, p. 1662-1668.
- [20] J.Edwards, 1992. «Compensatory speech motor abilities in normal and phonologically disordered children». *Journal of Phonetics*, vol. 20. p. 189-207.
- [21] M. M.Campbell, 1999. «Articulatory compensation for a bite block, with and without auditory feedback, in hearing and hearing impaired speakers». Thèse de doctorat, New-York, City University of New-York, 523 p.
- [22] J.R.Green, A. Moore et K.J.Reilly, 2002. «The sequential development of lip band jaw control for speech». *Journal of Speech, Language, and Hearing Research*, vol.45, p. 66-79.
- [23] B. L Smith et A. McLean-Muse. 1987. «Effects of rate and bite-block manipulations on kinematic characteristics of children's speech». *Journal of the Acoustical Society of America*, vol. 81, no 3, p. 747-754.

- [24] L.Ménard, H. Loevenbruck et C. Savariaux. (2006). « Articulatory and acoustic correlates of contrastive focus in French: a developmental study ». In Towards a better understanding of speech production processes, sous la dir. de Harrington, J. and Tabain, M. Psychology Press : New York.
- [25] A.Gibson et L. McPhearson 1980. «Production of bite-block vowels by children». University of Stockholm. PERILUS Report, vol. 2, p. 26-43.
- [26] R.Netsell, 1985. «Construction and use of a bite-block for the evaluation and treatment of speech disorders». Journal of Speech and Hearing Disorders, vol. 50, p. 103-109.
- [27] C.Savariaux, P. Perrier et J.-P. Orliaguet. 1995. «Compensatory strategies for the perturbation of the rounded vowel [u] using a lip tube : A study of the control space in speech production». Journal of the Acoustical Society of America, vol. 98, no 5, p. 2428- 2442.
- [28] C.Savariaux, P. Perrier, J.-P. Orliaguet et J.-L. Schwartz. 1999. «Compensation strategies for the perturbation of French [u] using a lip tube. II. Perceptual analysis». Journal of the Acoustical Society of America, vol. 106, no 1, p. 381-393.
- [30] L.Ménard, P. Perrier et C. Savariaux. 2004. « Exploring production-perception relationships during speech development : a study of compensation strategies to a lip-tube perturbation », 147th Congress of the Acoustical Society of America, vol. 115, no 5, p. 2629.
- [31] C.J. Riordan, 1977. «Control of vocal-tract length in speech». Journal of the Acoustical Society of America, vol. 62, no 4, p. 988-1002.
- [32] J. S.Perkell, F. H. Guenther, H. Lane, M. Matthies, P. Perrier, J. Vick, R. Wilhelms-Tricarico et M. Zandipour. 2000. «A theory of speech motor control and supporting data from speakers with normal hearing and with profound hearing loss». Journal of Phonetics, vol. 28, p. 233-272.
- [33] J. S.Perkell, M. L. Matthies, M. A. Svirsky et M. I. Jordan. 1993. «Trading relations between tongue-body raising and lip rounding in production of the vowel /u/: A pilot “motor equivalence” study». Journal of the Acoustical Society of America, vol. 93, no 5, 2948-2961.
- [34] J. A. Jones et K. G. Munhall. 2003. «Learning to produce speech with an altered vocal tract: The role of auditory feedback». Journal of the Acoustical Society of America, vol.113, no 1, p. 532-543.
- [35] H.Lane, M. Denny, F. H. Guenther, M. Matthies, L. Ménard, J. S. Perkell, E. Stockmann, M. Tiede, J. Vick et M. Zandipour. 2005. «Effects of bite blocks and hearing status on vowel production». Journal of the Acoustical Society of America, vol.118, no 3, p. 1636- 1646.
- [36] F. H.Guenther, & J. S. Perkell. 2004. «A neural model of speech production and its application to studies of the role of auditory feedback in speech». In Speech Motor Control in Normal and Disordered Speech, sous la dir. de B. Maassen, R. Kent, H.Peters, P. Van Lieshout, et W. Hulstijn, p. 29-49. Oxford: Oxford University Press.

[37] J. S.Perkell, M. Zandipour, M. Matthies, et H. Lane. 2002. « Economy of effort in different speaking conditions. I. A preliminary study of intersubject differences and modelling issues», *Journal of the Acoustical Society of America*, vol. 112, p. 1627-41.

[38] G. J. Borden, 1976. «The effect of mandibular nerve block upon the speech of four-year-old boys». *Language and Speech*, vol. 19, 173-178.

[39] J. A. S.Kelso et B. Tuller. 1983. «Compensatory articulation under conditions of reduced afferent information : a dynamic formulation». *Journal of Speech and Hearing Research*, vol. 26, p. 217-224.

[40] P.Perrier 2005. «Control and representations in speech production». *ZAS Papers inLinguistics*, vol. 40, p. 109-132.

Chapitre 2 - Etude expérimentale : réalisation du matériau linguistique, analyse des données et détermination de patrons temporels des fricatives

2.1 Introduction

L'analyse temporelle et acoustique du signal de parole est réalisée sur de la parole produite dans des conditions contrôlées. L'étude porte sur les fricatives pharyngales de l'arabe /ħ/ (/ħ/) , /h/ (/h/) et /ħ/ (/ħ/) , placées dans les contextes vocaliques /a/, /i/, /u/. Les phrases sont interrogatives, ce qui permet de maintenir constant le contour prosodique de la phrase. Par ailleurs, d'une phrase à l'autre nous varierons soit la fricative et on garde le contexte vocalique constant, soit à l'inverse, on variera le contexte vocalique pour chaque fricative. Nous pensons ainsi pouvoir observer les variations temporelles relatives à la variation de contextes vocaliques d'une part et à la variation du lieu d'articulation d'autre part.

Dans une deuxième étape, nous allons faire une lecture des mouvements des articulateurs concernés par les variations des durées des unités de parole considérées, nous pensons que ceci nous aidera à mieux comprendre les phénomènes de contrôle mis en jeu dans la production des pharyngales arabes.

2.2 Les fricatives pharyngales de l'arabe

Une consonne pharyngale est une consonne dont le lieu d'articulation se situe dans la zone du pharynx ; Parmi ces sons, nous avons des occlusives et des fricatives : /ħ/ (/ħ/) est une fricative voisée et /ħ/ (/ħ/) et /h/ (/h/) sont des fricatives non voisées spécifiques de l'arabe. Il y a eu plusieurs travaux :

Al.Ani (1970) [1], dans son étude acoustique et physiologique des consonnes pharyngales de l'iraquien, a présenté le tracé des coupes sagittales obtenues par la méthode des rayons X (figure 2.1). Ces tracés sont obtenus sur un locuteur iraquien, produisant les fricatives /ħ/ et /ħ/ situées devant les voyelles de l'arabe /a/, /i/, /u/. Ces figures ne montrent pas de façon distincte la position de la langue et il est difficile de conclure que ces deux consonnes ont un même lieu d'articulation. En 1985, AL Ani [1] indique que le lieu d'articulation de la pharyngale non voisée /ħ/ est plus haut dans le pharynx que celui de la pharyngale voisée /ħ/. Il précise dans le cas du /ħ/ que la constriction est formée par la racine de la langue et les parois postérieures du pharynx et que ce son est non voisé. Il ajoute que cette constriction peut être voisée dans le cas où elle se trouve dans une position intervocalique. La consonne /ħ/ est variable selon le contexte, il affirme qu'à l'initial, elle peut être voisée, c'est « un glide à l'intervocalique » mais l'auteur trouve que /ħ/ a en général les caractéristiques d'une « occlusive non voisée ».

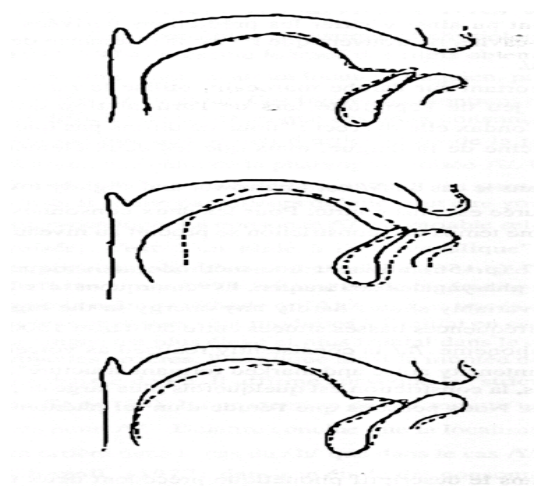


Figure 2. 1: Tracé des films aux rayons X, fricative /ʕ/ en traits pleins et /ħ/ en discontinus, précédant respectivement les voyelles /a/, /i/, /u/ (Al Ani 70)

Delattre, en 1971[1] fait une étude ciné radiographique des articulations pharyngales du libanais avec la même méthode qu'AL.Ani. Les tracés de la figure 2.2 sont faits sur un locuteur libanais et les consonnes sont produites à l'initial en contexte vocalique /a/, /i/, /u/. Il note que le dos de la langue est plus élevé et plus frontal dans le cas des consonnes pharyngales que dans le cas des voyelles antérieures, il explique ainsi la montée de F1. Enfin, l'élévation du « dorsal bulge » causerait une baisse de F3. Il affirme aussi que la « structure » de /ħ/ est plus basse et plus étroite que celle du /ʕ/. La production de /ħ/ montre un bruit durant la transition de F1, ce qui ne se voit pas pour /ʕ/. Delattre conclue que la localisation de la constriction durant l'articulation est plus en arrière dans le cas du /ħ/ que dans le cas du /ʕ/.

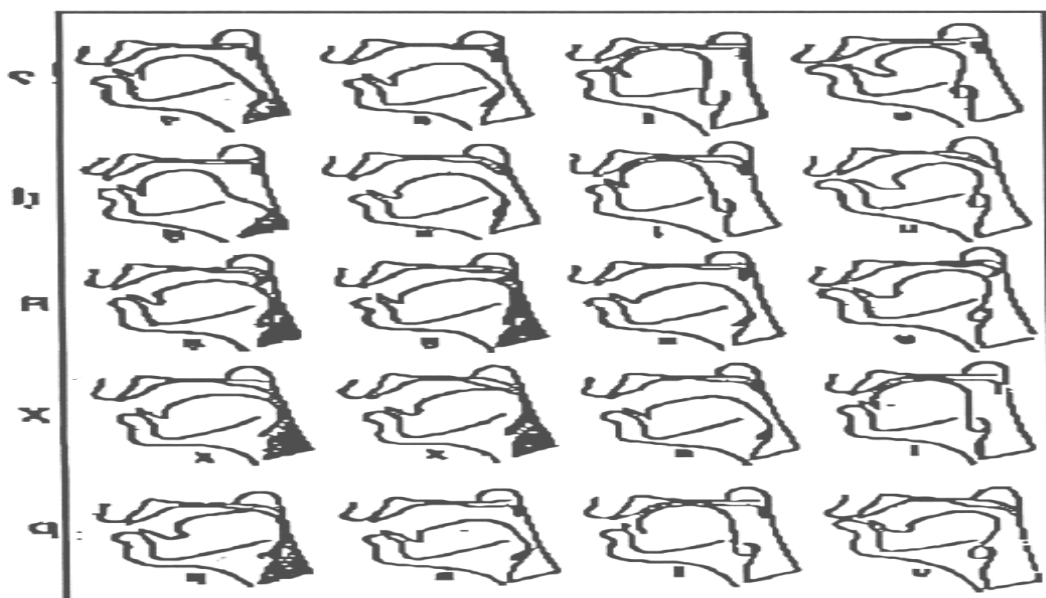


Figure 2. 2 : Tracé d'images obtenues par films aux rayons X, montrant une vue de profil d'un locuteur articulant des consonnes pharyngales et uvulaires en position initiale devant /a/, /i/, /u/. La figure à gauche de chaque rangée montre le moment de la constriction maximale. Les figures à droite de chaque rangée correspondent aux figures vocaliques (Delattre, 71)

Ghazali en 1977 [1], dans son étude des consonnes arrières du tunisien, arrive, après une étude avec le cinéfluographe, sensiblement aux mêmes conclusions que celles de Delattre. Par contre la mesure de F0 durant l'articulation de /**ʕ**/ indique la présence d'un « creek ». En utilisant aussi une méthode d'analyse aux rayons X (figure 2.3), sur un locuteur tunisien, il trouve que le lieu d'articulation des deux pharyngales est similaire. Le centre de la constriction est situé au centre de l'épiglotte, approximativement à 3.5cm de la glotte.

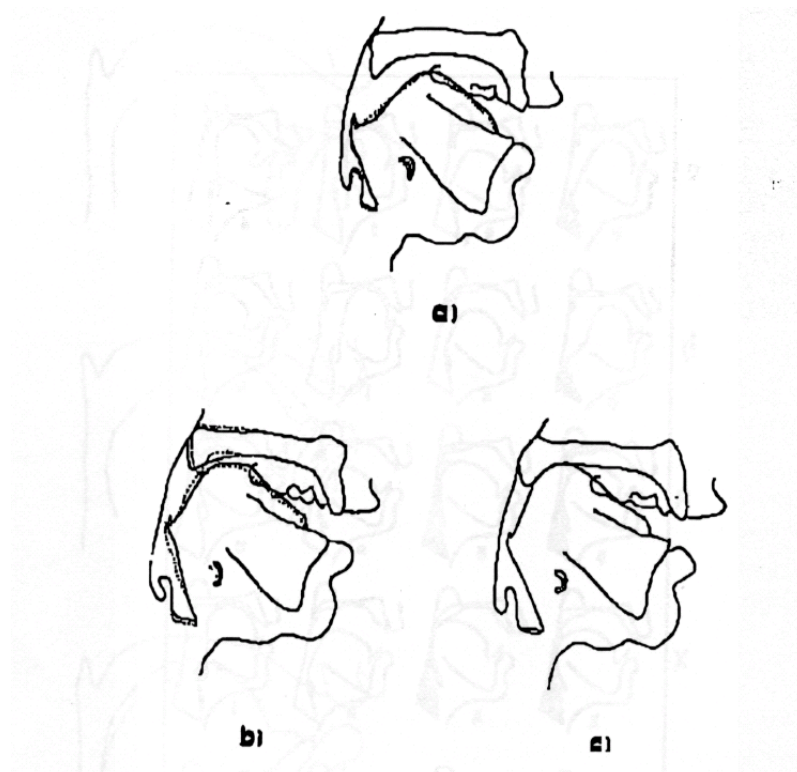


Figure 2. 3: a) Formes du conduit vocal produisant :/ʕ/ en traits discontinus , /ħ/ en traits pleins, b) /q/ en points tirés et [ħ] en traits pleins . c) /q/.

Ces 5 consonnes articulées dans le mot /Caeli/ où C l'une de ces consonnes (Ghazali 77)

Cependant cette constriction est plus étroite dans le cas du /ħ/ que dans le cas du /ʕ/ (3mm v.s. 4mm). Le larynx est relevé approximativement de 9mm par rapport à sa position neutre et de 7mm par rapport à la position active (7mm from its position during speech). Par ailleurs Ghazali note une autre constriction de l'ordre de 8mm, à 6cm des lèvres, celle-ci étant formée par la région située entre la racine, le dos de la langue et le palais dur.

Klatt et Stevens en 1974 [2], introduisent un modèle théorique de la fonction d'aire du conduit vocal des consonnes pharyngales, en régime de production. Un modèle idéal est utilisé pour examiner l'effet acoustique causé par la présence d'une constriction étroite située dans la région pharyngale (figure 2.4.a), dans ce cas (constriction postérieure), la longueur de la cavité arrière d est prise entre 3 et 4cm. Klatt et Stevens ont calculé les 4 plus basses fréquences du modèle en fonction de « d » et de l'aire à la constriction (figure 2.4.b). Les résultats obtenus sont conformes à ceux obtenus sur des spectrogrammes et ils ont pu vérifier les prédictions dérivées en ce qui concerne l'affiliation formant-cavité. Ils trouvent que F3 est la résonance de la cavité avant dans le cas des pharyngales /ʕ/ et /ħ/.

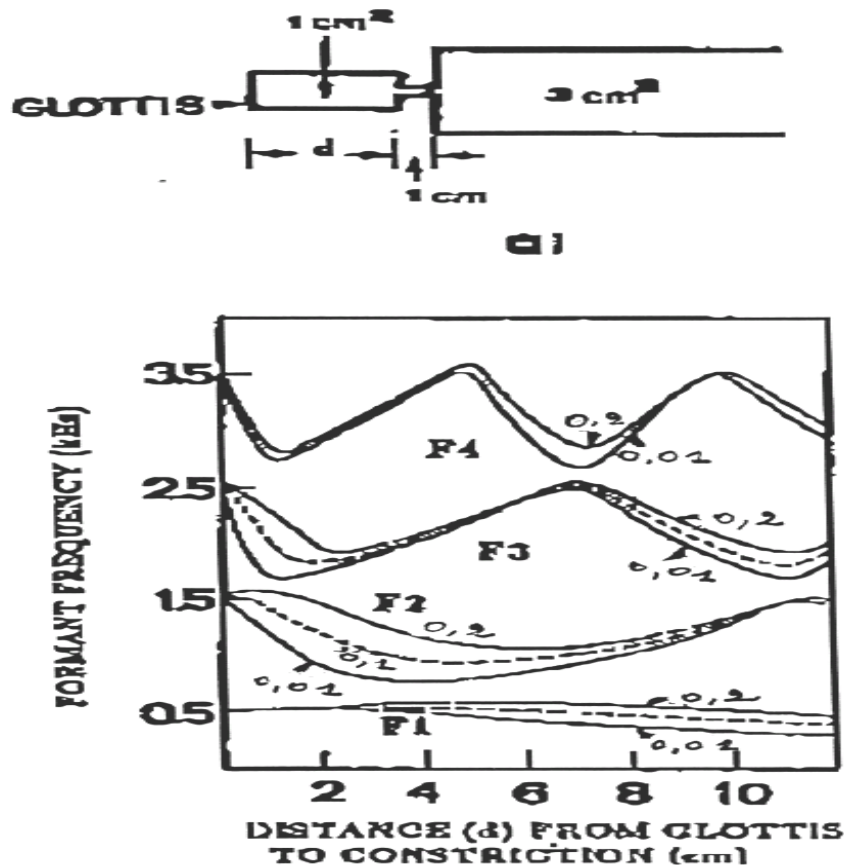


Figure 2. 4 : a) Modélisation de la fonction d'aire avec une constriction pharyngale (la glotte est à gauche) , b) représentation des 4 plus basses fréquences en fonction de d (longueur de la cavité arrière) et de l'aire à la constriction (ligne discontinue $AC=0.1\text{cm}^2$) (Klatt et Stevens 69)

Bohh Dkissi 1983[1], lors de son étude portant sur l'arabe marocain, utilise la méthode cinéfluographique pour observer la mise en jeu de l'épiglotte lors de l'articulation des deux consonnes / ʕ / et / h /. A l'inverse Laufer et Condax, elle dit ceci : « nous ne dirons pas que ce mouvement est indépendant de celui de la racine de la langue, mais que les deux éléments participent d'une même et unique contraction ».

Butcher en 1987, trouve que la constriction du / ʕ / est « narrower and slightly lower in the pharynx » que dans le cas du / h / et la durée est plus courte. Pour les deux consonnes le dos de langue est relevé vers le velum en même temps une constriction se produit au niveau du pharynx.

Butcher et Ahmed, dans leur article de 1987 [3], utilisent une méthode acoustique et aérodynamique pour caractériser les consonnes pharyngales de l'iraquien, ils remarquent : « when / ʕ / is released as a stop, the resulting burst invariably shows hardly an energy in the higher frequency region. ». On note aussi du bruit aux fréquences basses situées entre 500 et 1500Hz durant la réalisation des non voisées : « the phoneme / h / is consistly released as voiceless continuant sound with high rates airflow, high intensity voice and marked formant structure ». Ils remarquent qu'au voisinage des voyelles brèves, la constriction est quelquefois plus large et plus variable que dans le cas des voyelles longues.

A.Djeradi en 1992 [1], s'est intéressé à la production des fricatives arrières de l'arabe. Il a étudié les spécificités acoustiques de ces sons en développant un banc de mesures des fonctions de transfert du conduit vocal permettant de procéder à des mesures rapides et en cours de phonation, qui sont deux points capitaux dans l'analyse des fricatives.

Pour pouvoir établir les relations articulatoire-acoustiques, Il a aussi pu élaborer un modèle 3D du conduit vocal pour chaque fricative étudiée, qui peut être utilisé dans le modèle analogue du conduit vocal, pour simuler la production des sons analysés. Les résultats obtenus ont permis de mettre en évidence les effets de couplage des parties subglottiques des sources de friction

Pour la fricative /h/, il dit que dans la littérature, cette consonne est décrite comme une constrictive laryngale. Elle existe dans plusieurs langues (Anglais, Allemand, Tchèque Arabe....)

Comme l'ont remarqué certains auteurs, la radiographie relative à cette consonne montre que le lieu d'articulation n'est manifestement pas dans la cavité buccale, dans laquelle la langue présente une position neutre et laisse ainsi le libre passage à l'écoulement de l'air. Par comparaison avec les radiographies des autres pharyngales /ħ / et / h /, A.Djeradi ajoute qu'il n'a pu relever aucune similitude avec la contraction pharyngale de l'une ou de l'autre consonne. Comme le pharynx est complètement ouvert seul le rétrécissement laryngale qui est d'ailleurs similaire aux autres configurations, reste comme cause plausible de la phonation de cette consonne, et plusieurs auteurs affirment ne voir aucun mouvement musculaire se produire au cours de l'émission de cette consonne. C'est aussi le point de vue de Straka qui dit que celle-ci " ne se forme pas dans la cavité buccale, mais uniquement dans le larynx, et pendant son émission le canal buccal est largement ouvert sans qu'il se produise, au niveau de la langue, du maxillaire ou des lèvres aucun effort semblable aux autres consonnes".

A.Djeradi [1] ajoute que le comportement de cette consonne est de type vocalique sous l'effet du renforcement articulatoire puisqu'elle s'ouvre au lieu de se fermer.

Selon Lindqvist, parmi les constrictives sourdes, la consonne /h/ présente le plus faible mouvement d'abduction glottal, la consonne /h/ est donc caractérisée par la petitesse de l'ouverture de la glotte.

C.Zeroual et al. (99, 2000, 2002, 2003, 2004) [4, 5, 6, 7, 8] ont aussi présenté des travaux sur les pharyngales du parlé marocain. Ils se sont intéressés à la description et caractérisation des gestes laryngaux observés, par naso-fibroscopie, durant la production des gutturales [ħ ʁ h ʔ ħ] et de l'émphatique non-occlusive [S] de l'arabe marocain (l'AM). Cette analyse complète une étude antérieure faite aussi par naso-fibroscopie et dont les résultats majeurs ont été présentés dans plusieurs travaux.

Ces différentes études font apparaître, dans le descriptif phonétique, deux faits essentiels, la consonne /ħ / possède plusieurs allophones, ce qui est du probablement à son mécanisme de production complexe, et les controverses de la localisation du lieu d'articulation des différentes consonnes. La consonne /h/ ne fait intervenir aucun muscle lors de sa production, elle donc considérée comme neutre. Elle peut être considérée comme pharyngale semi voisée. Nous voyons que concernant les pharyngales arabes, beaucoup de points restent encore en suspend. La complexité de ces sons fait que de nos jours beaucoup de chercheurs continuent à s'y intéresser.

2.3 Le matériau linguistique

a) Particularité du corpus

Les pharyngales que nous étudions ici sont des fricatives de la langue arabe. Les phrases du corpus ont un sens et sont interrogatives. Ceci permet aux locuteurs de garder la même expression d'élocution. On relèvera les durées absolues des unités CVC puis des durées de la phrase complète. Les enregistrements sont fait pour trois débits différents d'élocution; lent, normal et rapide.

Nous n'imposerons pas un rythme mais une durée fixe de la phrase pour chaque débit. Les locuteurs devront faire leur phase d'apprentissage dans une chambre sourde sans la présence des autres locuteurs. Les phrases du corpus seront lues sur l'écran d'un ordinateur.

Afin de pouvoir faire une statistique de notre étude, le locuteur a répété plusieurs fois la même phrase. Nous choisirons 10 répétitions par phrase de corpus. Nous prendrons six locuteurs (3 régions géographiques)

Les six locuteurs sont tous étudiants. Ils ont des différences au niveau de l'apprentissage de la langue par le milieu maternel, ce sont :

- CH et LI (LI: locutrice), arabe du Liban, de langue maternelle libanaise,
- FE et NA sont algérois, de langue maternelle d'Alger: dialecte algérois (centre d'Algérie),
- SA et MA sont kabyles, de langue maternelle kabyle (très différente de l'arabe littéral) et vivant à Tizi Ouzou (est d'Algérie). Ils ont appris l'arabe à l'école pour la première fois.

Les phrases renferment trois fricatives pharyngales, spécifiques de l'arabe : le /ħ / (/ħ/ ; /h/ (/ħ/) ; /ħ / ou (ح), afin de mettre le locuteur en contexte consonantique variable de langue arabe. Notons que ces trois fricatives font partie aussi de l'alphabet kabyle et du dialecte algérois

Pour plus de facilité en écriture, nous représenterons :

le /ħ / (/ħ/ par /a_ /;
Le /h/ (/ħ/) par /h/ ;
Le /ħ / ou (ح) par /h_ /.

Les phrases sont au nombre de 9, trois pour chaque fricative, mise en contexte vocalique différent (trois contextes vocaliques arabes (/a/, /i/, /u/)

Pour le /ħ / (/ħ/ représenté par /a_ / :
/ men saa_ala ? / ; / men saa_ila ? / ; / men saa_ula ? /

Pour le /h/ (/ħ/) représenté par /h/ :
/ men sahala ? / ; / men sahila ? / ; / men sahula ? /

Pour le / h / ou (ح) représenté par /h_ / :
/ men sah_ala ? / ; / men sah_ila ? / ; / men sah_ula ? /.

Avec 10 répétitions par locuteur, dans trois débits différents: normal, rapide, lent. Le corpus comporte alors : $9 \times 10 \times 3 \times 6 = 1620$ phrases pour les six locuteurs.

b) Enregistrement des phrases

Les phrases du corpus sont écrites sur l'écran d'un ordinateur, il est demandé dans un premier temps au locuteur, de les lire dans trois débits différents : le débit normal du locuteur, un débit assez rapide par rapport à son débit normal, puis un débit relativement lent par rapport à son débit normal.

Dans un deuxième temps, des durées ont été fixées pour les débits artificiels (rapides et lents) puis il a été demandé aux locuteurs d'essayer de les respecter dans la mesure du possible, pour mettre davantage le locuteur dans la contrainte.

Les durées fixées par débit sont :

Environ 0.5s pour la durée de la phrase en débit rapide
Environ 1.5s la durée de la phrase en débit lent.

Pour arriver à respecter approximativement les durées fixées et éviter des influences mutuelles, chaque locuteur a suivi seul une phase d'apprentissage, dans une chambre sourde, en lisant les phrases du corpus sur l'écran d'un ordinateur.

Notons donc que chaque locuteur a été enregistré isolé des autres.

Les phrases ont été enregistrées dans la chambre sourde du laboratoire d'acoustique de la faculté d'électronique, sur un PC (P4), muni d'une carte son de type sound Blaster, à l'aide d'un microphone professionnel. La fréquence d'échantillonnage a été fixée à 16kHz à cause de la présence des fricatives dans le corpus de phrases. Le logiciel d'enregistrement utilisé est le logiciel PRAAT

c) Analyse des segments de parole : durées relatives et absolues

Les unités à analyser étant les CVC, prélevés dans les phrases du corpus. Il s'agit alors de segmenter les signaux acoustiques obtenus pour prélever les unités de parole concernées. La segmentation a été faite manuellement à l'aide du logiciel Praat (la segmentation automatique ne donnant pas toujours les résultats escomptés).

Il s'agit à présent de choisir l'espace de représentation des unités temporelles segmentées. Le choix de l'espace de représentation des temps est important si l'on veut éliminer les différences liées à l'articulation en absolue.

En effet, nous avons constaté que les durées absolues des phrases et des CVC sont variables pour un même locuteur, intra débit, nous avons donc opté pour l'analyse de durées relatives des CVC dans la phrase concernée, pour éliminer les différences liées à des allongements ou des rétrécissements des phrases dans un même débit.

Les segments temporels de parole retenus dans notre étude seront donc : les durées absolues de la phrase et les durées absolues des /CjVil/ concernées. Nous avons ensuite déterminé les durées relatives des CVC dans la phrase concernée : $d(\text{CVC})/d(\text{phrase})$.

Par ailleurs, le but fixé étant la recherche de patrons acoustiques de production de la parole en contrainte de débit pour ces six locuteurs, une étude de l'espace vocalique en contrainte de débit est nécessaire. Le calcul des fréquences formantiques F1 et F2 de la voyelle dans le CVC est imminent. Les fréquences formantiques ont été prélevées au cœur de la voyelle concernée, à l'aide du logiciel PRAAT.

Nous obtenons les unités temporelles /CjVil/, Vi étant les trois voyelles de l'arabe (i=/a/, /i/, /u/), en contexte consonantique /Cj/: j=/a_/, /h/, /h_/, pour chaque locuteur, soit :

/a_al/, /a_il/, /a_ul/, /hal/, /hil/, /hul/, /h_al/, /h_il/, /h_ul/, et l: la consone liquide /l/, et les formants F1 et F2 relatifs aux trois voyelles de l'arabe Vi (i=/a/, /i/, /u/).

2.4 Présentation des résultats

Les résultats obtenus sont regroupés dans des tableaux par débit, par locuteur, par fricative et par contexte vocalique. Nous trouvons dans ces tableaux :

- le nombre de mesures par échantillons qui est ici de dix (répétitions) par fricative, par contexte vocalique et par locuteur.
- la durée absolue de la phrase en milliseconde
- la durée absolue du CVC en milliseconde
- la durée relative du CVC sur la phrase
- les formants F1 et F2 de la voyelle V dans le CVC, en Hz.

Un exemple est donné par le tableau 2.1.

Consonne /a_/_	Durée de la Phrase en ms	Durées CVC en ms	Durées relatives	F1 En Hz	F2 En H
Echantillons	d(ph)		D(CVC)/d(phrase)	F1(a)	F2(a)
Numéro de l'échantillon	d(men saa_ala)	d(a_al)	d(a_al)/ d(men saa_ala)	F1(a)	F2(a)
1	876	246	0.2808	833	1322
2	896	261	0.2913	716	1357
3	900	270	0.3000	741	1461
4	900	290	0.3222	746	1479
5	928	256	0.2759	781	1479
6	988	280	0.2834	798	1461
7	900	272	0.3022	816	1409
8	904	285	0.3153	730	1385
9	934	285	0.3051	695	1412
10	904	269	0.2976	689	1442
moyenne	913	271,4	0,29738	754,5	1420,7
Écart type	30,94	14,09	0,015	50,2	53,5

Tableau 2. 1. Exemple de durées et de formants relevés sur les unités d'analyse considérées.

Des tableaux similaires ont été remplis pour chaque débit, chaque fricative, chaque contexte vocalique et chaque locuteur.

Nous avons analysé le timing des phrases des locuteurs. Nous avons constaté que les durées absolues des phrases intra débit étaient en effet assez différentes d'un locuteur à un autre. Les différents intervalles de durées par débit ont été relevés sur les enregistrements pour les durées de phrase:

0.75s<Durée de la phrase du corpus à débit normal <1s
0.45s<Durée de la phrase du corpus à débit rapide <0.55s
1.2s<Durée de la phrase du corpus à débit lent <1.8s

Notons que l'intervalle de durée le plus grand correspond au débit lent. Le débit lent n'est pas un débit naturel ! Il est sans doute plus difficile à réaliser que le débit rapide. Les locuteurs ont eu plus de difficultés dans la réalisation des débits lents.

2.5 Analyse des durées par l'analyse de la variance

Les différentes durées obtenues sont :

- les durées de phrases : d(ph)
- les durées des CVC prélevés dans les phrases : d(CjVil), (Vi étant les trois voyelles de l'arabe (i=/a/, /i/, /u/)), en contexte consonantique /Cj/: j=/a_/_/, /h/, /h_/_/,
- les durées relatives sont obtenues en faisant le rapport des durées de CVC et de la phrase entière contenant celui ci : d(CjVil)/d(ph).

Il s'agit alors de faire une étude comparative des moyennes de ces durées. L'outil le plus approprié pour la comparaison des moyennes d'échantillons est l'ANOVA.

2.5.1 Méthode ANOVA

L'analyse de la variance est une technique statistique [9] permettant de comparer les moyennes de plus de deux populations. Cette méthode permet d'étudier les différences de moyenne entre populations (par exemple, trois populations ont-elles la même moyenne? ou autrement dit, les différences de moyenne entre les trois populations sont-elles significatives?). Nous utiliserons cette méthode pour déterminer le caractère significatif, ou non, des différences de moyenne mesurées sur les populations.

Il existe plusieurs types d'ANOVA, ANOVA à un facteur lorsque les groupes analysés se distinguent par un seul facteur qualitatif (par exemple, contexte vocalique), ANOVA à deux facteurs si les groupes se distinguent par deux facteurs qualitatifs (par exemple, contexte vocalique /voyelle /a/ et débit /lent) ...etc.

Dans les cas simples (ANOVA à un facteur fixe), l'hypothèse nulle H_0 est : les moyennes des différents groupes sont égales. Il s'agit de comparer des moyennes de différents groupes et de dire si, parmi l'ensemble, au moins une d'entre elles diffère des autres, mais on ne sait ni laquelle ni combien d'entre elles.

2.5.2 Principe général de la méthode

Soit k , le nombre d'échantillons dont on doit comparer les moyennes, le nombre de mesures par échantillon est désigné par n et le nombre total de mesures, $k*n$.

L'analyse de variance à un facteur, consiste à chercher le rapport entre : la variance entre les groupes (V. inter-groupe) et la variance à l'intérieur des groupes (V. intra-groupe). La valeur de ce rapport notée F [attention : ce F n'a rien à voir avec le F du test de vérification de l'homogénéité des variances] est comparée à celle d'une table de F de Snedecor, ou Fisher-Snedecor, table à double entrée avec pour numérateur le [nombre d'échantillons (k) moins un] soit $(k-1)$ et pour dénominateur le [nombre total de mesures - k] soit $(k*n-k)$. La valeur obtenue de F nous donnera ainsi la valeur de p : probabilité de risque d'erreur (obtenu à l'aide dans la table de Fisher-Snedecor).

La décomposition de l'analyse de la variance est toujours valable, quelle que soit la distribution des variables étudiées. Cependant, lorsqu'on réalise le test final (de F), on admet la normalité de ces distributions (puisque le F est le rapport de deux khi-deux, qui sont des sommes de carrés de lois normales). L'ANOVA fait donc l'hypothèse de normalité. Elle est cependant assez robuste à la non normalité, ce qui permet de l'utiliser dans une grande variété de conditions.

Pour vérifier la normalité de nos échantillons nous avons opté ici pour le test de Shapiro-Wilk.

La méthode développée par Shapiro-Wilk est dans bien des cas, la plus puissante, en particulier lorsque l'échantillon provient d'une distribution asymétrique. Cette méthode implique l'emploi de tables, actuellement calculées pour une taille d'échantillon comprise entre 5 et 50. ($5 \leq n \leq 50$). Comme les autres tests, il faudra déterminer à l'avance un risque de rejeter l'hypothèse nulle alors que celle-ci est vraie (α).

Le test de Shapiro-Wilk est basé sur le rapport W de deux estimations liées à la variance de la population dont provient l'échantillon : l'une, fonction des étendues partielles : $x_2 - x_1, x_3 - x_2, \dots$, que l'on peut trouver à partir de la suite ordonnée des n observations indépendantes d'un échantillon de taille n , l'autre, $(n-1)S^2$ fonction des carrés des écarts (SCE) à la moyenne $\bar{x}$ des observations S^2 :

$$SCE = \sum (x_i - \bar{x})^2 = (n-1)S^2$$

$$d_1 = x_2 - x_1$$

$$d_2 = x_3 - x_2$$

$$d_3 = x_4 - x_3$$

$$\dots$$

$$b = \sum a_i d_i$$

$$W = \frac{b^2}{SCE}$$

Le rapport W sera comparé à une valeur théorique $W_{1-\alpha, n}$ et dans le cas où $W \geq W_{1-\alpha, n}$ nous pourrions affirmer avec un risque d'erreur α que la distribution suit une loi normale.

W est interprété comme un coefficient de corrélation quadratique entre la distribution observée et une distribution normale. Plus la distribution observée est proche d'une loi normale, plus la valeur de W tend vers 1.

• Calcul de l'ANOVA

Nous avons appliqué l'ANOVA dans notre étude pour la comparaison des moyennes des différentes durées des unités de parole considérées, qui sont les durées:

- des phrases : dph
- des CVC prélevés dans les phrases : dCjVil, (Vi étant les trois voyelles de l'arabe (i=/a/, /i/, /u/)), en contexte consonantique /Cj/: j=/a_/, /h/, /h_/,
- des durées relatives de ces CVC sur les phrases : d(CjVil)/d(ph)

L'ANOVA a un facteur intra (trois contextes vocaliques). Nous avons mélangé les données de ces trois voyelles, avec 10 répétitions par contexte vocalique, nous obtenons 30 stimuli. Nous rechercherons ainsi le paramètre $F(2,27)$ et la probabilité d'erreur « p » qui lui est associée, pour un locuteur donné, et une fricative donnée, pour un facteur de risque de 0.05.

Les groupes de mesures sont définis par $K=3$ échantillons comprenant $n=10$ individus dont les valeurs sont données dans le tableau ci-après : il s'agit des durées relatives des « CVC » dans le phrase « ph » porteuse.

Calculer l'ANOVA revient à déterminer la valeur de F qui sera comparée à celle d'une table de F de Fisher-Snedecor, avec pour numérateur ($k-1=2$) et pour dénominateur le [nombre total de mesures - k] soit ($kn-k=30-3=27$). Nous écrirons donc ANOVA avec $F(2,27)$.

Le but est de comparer les moyennes de ces durées relatives avec le risque de commettre une erreur de 5%. Plus le risque d'erreur est élevé, plus le test sera robuste.

Dans le tableau 2.2 nous avons 3 échantillons et 10 mesures par échantillon.

dCVC/dph(a)	dCVC/dph(i)	dCVC/dph(u)
0.2808	0.3003	0.2711
0.2913	0.2634	0.2703
0.3000	0.2742	0.2911
0.3222	0.2887	0.2904
0.2759	0.2892	0.2819
0.2834	0.2932	0.2897
0.3022	0.2830	0.2806
0.3153	0.2780	0.2850
0.3051	0.2671	0.2963
0.2976	0.2598	0.2726

Tableau 2. 2 : Durées relatives des CVC sur la phrase

Avant d'effectuer l'ANOVA, il faut vérifier le test de la normalité. Cette opération sera effectuée à l'aide du logiciel « ORIGINPRO70 » qui nous donne les résultats obtenus (tableau 2.3) à l'aide du test de Shapiro-Wilk :

Normality Test (Shapiro-Wilk)

Dataset	N	W	P Value	Decision
dCVC/dph(a)	10	0,972	0,902	Normal at 0,05 level
dCVC/dph(i)	10	0,965	0,836	Normal at 0,05 level
dCVC/dph(u)	10	0,926	0,402	Normal at 0,05 level

Tableau 2. 3. Vérification de la normalité à l'aide du Test Shapiro-Wilk

Toutes les probabilités « p » obtenues sont supérieures à 0.05, donc tous les échantillons suivent une loi normale par conséquent la normalité est vérifiée pour les 3 groupes.

- **Analyse de la variance :**

Nous avons donc appliqué l'ANOVA pour déterminer les paramètres du tableau 2.4:

Source de variation	Degré de liberté ou DDL	Somme des carrés	Moyenne des carrés ou variance	Test-F	P
V. Intergroupes	k-1=2	0,00177	8,881E-4	5,447	0,0102
V. Intra-groupes	K*n-k=27	0,00440	1,6303E-4		

Tableau 2. 4 : Valeurs de F et de p dans l'ANOVA

Dans cet exemple, la valeur de $F(k-1, kn-k)=F(2,27)$ est très inférieure au seuil de signification de 5% pour considérer l'hypothèse nulle. La probabilité exacte calculée est de 0,01029.

L'hypothèse nulle est donc rejetée. On considère que les 3 échantillons diffèrent significativement (la probabilité de se tromper est de $0,01029 < 0,05$).

- La valeur de p est l'erreur alpha, soit la probabilité ou le risque de commettre une erreur, en déclarant qu'il existe une différence entre les trois groupes ou les trois mesures.
- C'est ce résultat qui permet de confirmer ou d'infirmer notre hypothèse alternative (H1).
- Rappelons qu'en sciences humaines, le seuil de signification utilisé est en général de 0,05.
- Si la valeur de p est supérieure à 0,05, nous devons accepter l'hypothèse nulle et conclure qu'il n'y a pas de différence significative entre nos deux groupes (ou mesures).
- Si la valeur de p est inférieure à 0,05, il faut faire deux choses:
 1. d'abord rejetez l'hypothèse nulle et conclure qu'il y a une différence significative entre les trois groupes;
 2. procédez à un test post-hoc, afin de savoir laquelle des comparaisons de groupes, pris deux à deux, est significativement différente
- Il existe plusieurs types de test post-hoc

Le test de Scheffé est le plus utilisé en sciences humaines. C'est le test le plus conservateur.

Le F de Scheffé permet de déterminer si, après une ANOVA significative, les moyennes de 2 des groupes de la variable indépendante diffèrent.

Le test de Scheffé ne demande pas que tous les échantillons utilisés dans l'ANOVA aient la même taille.

L'application du test de Scheffé, correspondant à l'exemple précédent, avec le logiciel « ORIGINPRO70 » nous donne le tableau 2.5 :

Means Comparison using Scheffe' Test					
Dataset	Mean	Difference between Means	Simultaneous Confidence Intervals		Significant at 0,05 Level
			Lower Limit	Upper Limit	
dCVC/dph(a)	0,29738				
dCVC/dph(i)	0,27969	0,01769	0,0029	0,03248	Yes
dCVC/dph(u)	0,2829	0,01448	-3,0987E-4	0,02927	No
dCVC/dph(i)	0,27969				
dCVC/dph(u)	0,2829	-0,00321	-0,018	0,01158	No

Tableau 2. 5. Exemple des résultats obtenus à l'aide du test de Scheffé à l'aide du logiciel ORIGINPRO70.

Le tableau 2.5 montre une comparaison des durées relatives en contexte vocalique /a/ par rapport aux contextes vocaliques /i/ puis /u/ .

Ensuite une comparaison des durées relatives en contexte vocalique /i/ par rapport au contexte vocalique /u/ .

Sur la base des moyennes, des différences entre moyennes, de l'intervalle de confiance, pour pouvoir ensuite conclure par rapport à la probabilité de risque fixée « p »

Nous savons à présent que les trois moyennes sont significativement différentes globalement. Le test post-hoc va nous indiquer quels groupes, pris deux à deux, sont différents. Le test post hoc utilisé est le test de Scheffé, qui permet de comparer les moyennes des unités :

dCVC/dph(a) avec dCVC/dph(i)
dCVC/dph(a) avec dCVC/dph(u)
dCVC/dph(i) avec dCVC/dph(u)

Les résultats obtenus (à l'aide du logiciel originpro70) montrent que les moyennes des durées relatives :

dCVC/dph(a) et dCVC/dph(i) : sont significativement différentes à $p=0.05$
dCVC/dph(a) avec dCVC/dph(u) : ne sont pas significativement différentes à $p=0.05$
dCVC/dph(i) avec dCVC/dph(u) : ne sont pas significativement différentes à $p=0.05$

Cela voudrait dire que nous pouvons considérer que :

$dCVC/dph(a) \neq dCVC/dph(i)$
 $dCVC/dph(a) = dCVC/dph(u)$
Et que $dCVC/dph(i) = dCVC/dph(u)$

Si nous regardons les valeurs numériques des moyennes des durées relatives différentes, nous pouvons écrire que : $dCVC/dph(a) > dCVC/dph(i)$.

Nous dirons donc qu'en contexte vocalique /a/, les durées relatives sont plus élevées qu'en contexte /i/ et /u/

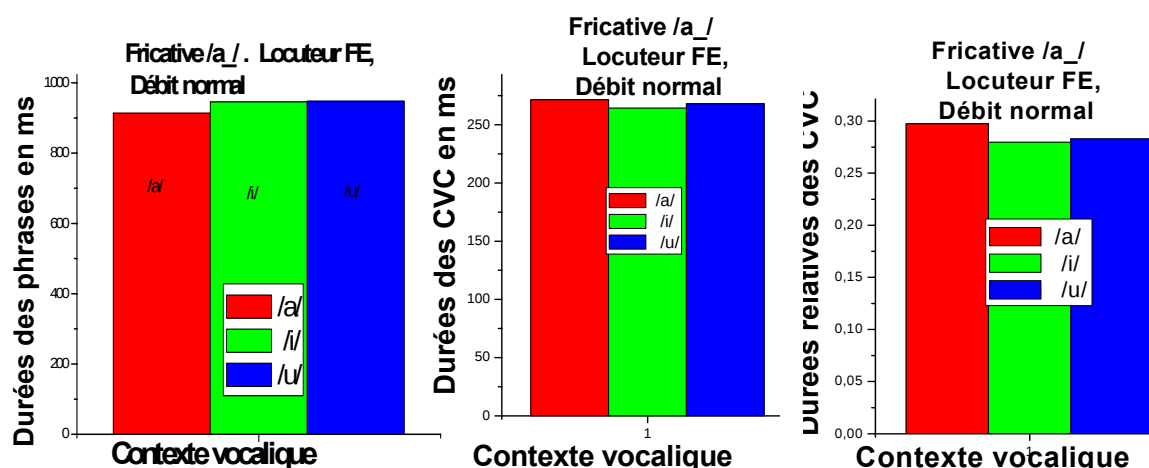
2.6 Analyse de l'influence du contexte vocalique par locuteur

a- Locuteur FE

a-1- fricative /a_ /

a-1-1- Analyses des durées des phrases, des CVC et des durées relatives, en contexte /a/, /i/ /u/

Les figures 2.5.a. 1 illustrent l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /a_ /



Figures 2.5.a. 2 : Influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /a_ /.

La comparaison des moyennes des durées à l'aide de l'ANOVA est affichée dans le tableau 2.6.a.1

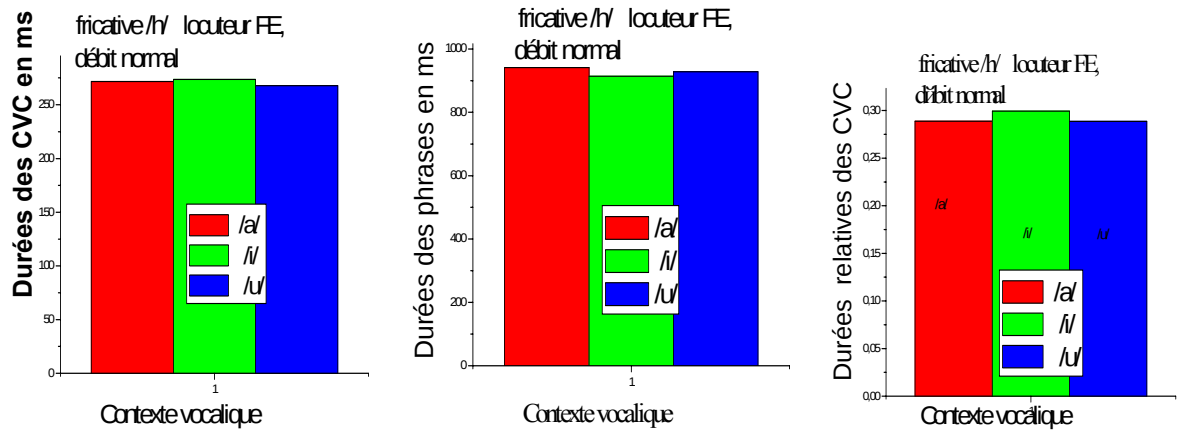
Fricative /a /			Locuteur FE	variation du contexte vocalique			
Test statistique	Résultat de l'ANOVA		Résultat de l'ANOVA avec test de Scheffé	Test statistique			
Durée de l'unité	F	p	Significativement différentes	Test Scheffé		Résultats avec test de Scheffé	
dph	4,03758	0,02924	non	Différence de durées entre 2 contextes consonantiques en ms		Significativement différentes	
				da-di	0		non
				da-du	0		non
				di-du	0		non
dCVC	0,86579	0,43208	non	da-di	0	non	
				da-du	0	non	
				di-du	0	non	
dr	5,4477	0.01029	oui	da-di	0,01769	oui	
				da-du	0	non	
				di-du	0	non	

Tableau 2.6.a. 1. Influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /a_ /

a-2- fricative /h/

a-2-1- Analyses des durées des phrases en contexte /a/, /i/, /u/

Les figures 2.5.a. 2 illustrent l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h/



Figures 2.5.a. 3: Influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h/

La comparaison des moyennes des durées à l'aide de l'ANOVA est affichée dans le tableau 2.6.a. 2

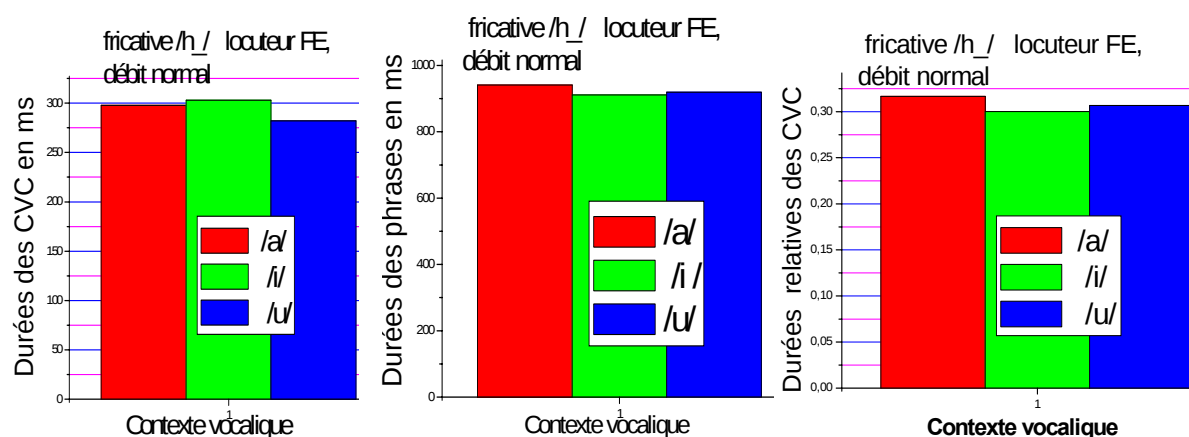
Durée de l'unité	Test statistique		Résultat de l'ANOVA	Résultat de l'ANOVA avec test de Scheffé		
	F	p	Significativement différentes	Test scheffé	Résultats avec test de Scheffé	
dph	2,41634	0,10830	non	Différence de durées entre 2 contextes consonantiques en ms		Significativement différentes
				da-di	0	non
				da-du	0	non
				di-du	0	non
dCVC	0,70999	0,500	non	da-di	0	non
				da-du	0	non
				di-du	0	non
				da-di	0	non
dr	3,10615	0,06108	non	da-du	0	non
				di-du	0	non
				da-di	0	non
				da-du	0	non

Tableau 2.6.a. 2. Résultats obtenus pour l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h/

a-3- fricative /h_/

a-3-1- Analyses des durées des phrases en contexte /a/, /i/, /u/

Les figures 2.5.a. 3 illustrent l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h_/



Figures 2.5.a. 4: Influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h_/

La comparaison des moyennes des durées à l'aide de l'ANOVA est affichée dans le tableau 2.6.a. 3

Fricative /h / Locuteur FE variation du contexte vocalique						
	Test statistique		Résultat de l'ANOVA	Résultat de l'ANOVA avec test de Scheffé		
Durée de l'unité	F	p	Significativement différentes	Test scheffé		
dph	3,54542	0,04293	oui	Différence de durées entre 2 contextes consonantiques en ms		Significativement différentes
				da-di	0	non
				da-du	0	non
				di-du	0	non
dCVC	6,83244	0,00397	oui	da-di	0	non
				da-du	15,9 ms	oui
				di-du	21ms	oui
dr	2,85365	0,0751	non	da-di	0	non
				da-du	0	non
				di-du	0	non

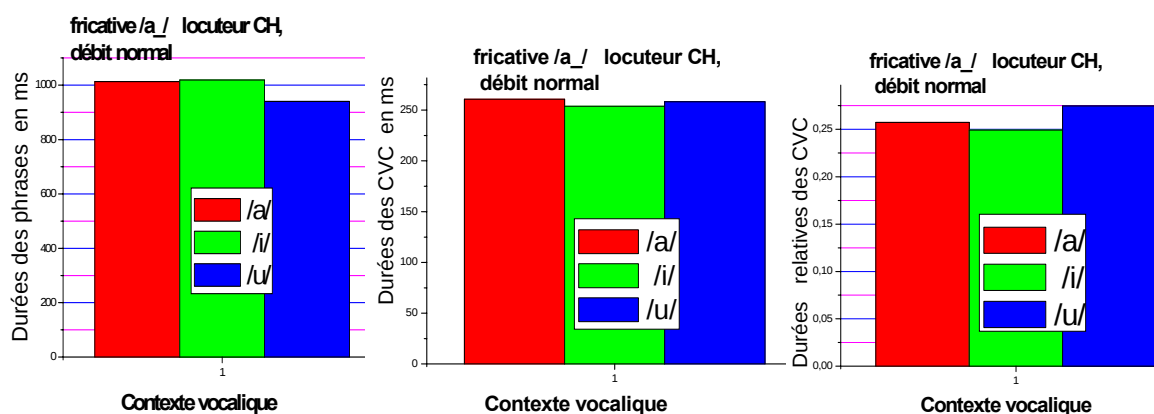
Tableau 2.6.a. 4. Résultats obtenus pour l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h/

b- Locuteur CH

b-1- fricative /a_/

b-1-1- Analyses des durées des phrases en contexte /a/, /i/, /u/

Les figures 2.5.b.1 illustrent l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /a_/



Figures 2.6.b. 1 : Influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /a_/

La comparaison des moyennes des durées à l'aide de l'ANOVA est affichée dans le tableau 2.6.a. 5

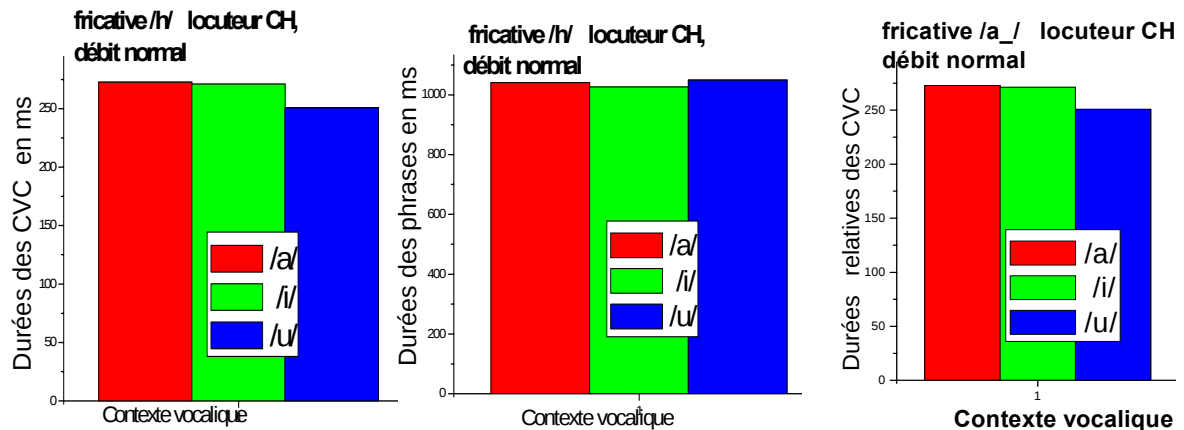
Fricative /a / Locuteur CH variation du contexte vocalique						
	Test statistique		Résultat de l'ANOVA	Résultat de l'ANOVA avec test de Scheffé		
Durée de l'unité	F	p	Significativement différentes	Test de scheffé		Résultats avec test de Scheffé
dph	24,09032	9,90786E-7	oui	Différence de durées entre 2 contextes consonantiques en ms		Significativement différentes
				da-di	0	non
				da-du	72,9ms	oui
				di-du	79,1ms	oui
dCVC	0,44474	0,64560	non	da-di	0	Non
				da-du	0	Non
				di-du	0	non
dr	6,21875	0,00601	oui	da-di	0	Non
				da-du	0	Non
				di-du	-0.0255	oui

Tableau 2.6.b. 1. Résultats obtenus pour l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h/

b-2- fricative /h/

b-2-1- Analyses des durées des phrases, des CVC et des durées relatives, en contexte /a/, /i/, /u/.

Les figures 2.5.b.2 illustrent l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h/



Figures 2.5.b. 2: Influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h/

La comparaison des moyennes des durées à l'aide de l'ANOVA est affichée dans le tableau 2.6.b.2

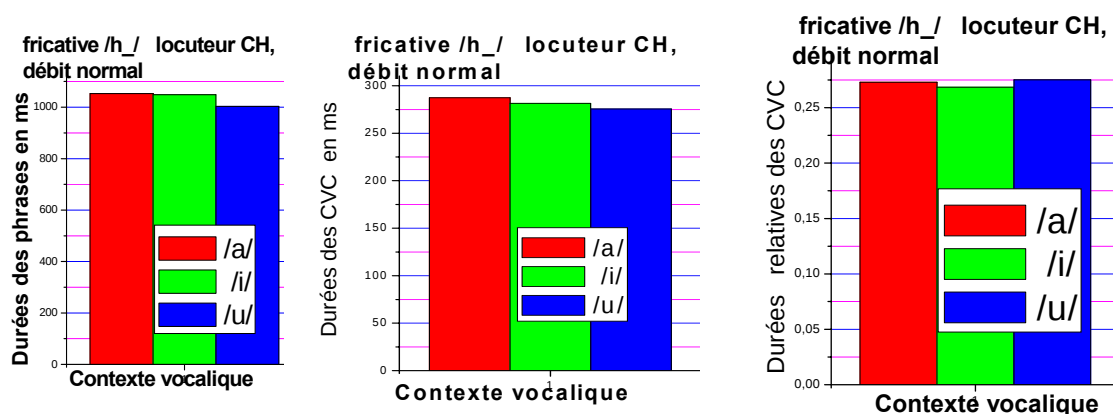
Fricative /h/ Locuteur CH variation du contexte vocalique					
Durée de l'unité	Test statistique		Résultat de l'ANOVA	Résultat de l'ANOVA avec test de Scheffé	
	F	p	Significativement différentes	Test de scheffé	Résultats avec test de Scheffé
dph	2,22762	0,12722	non	Différence de durées entre 2 contextes consonantiques en ms	
				da-di	0
				da-du	0
				di-du	0
dCVC	11,21824	2,842E-4	oui	da-di	0
				da-du	22ms
				di-du	20,3ms
dr	13,97323	6,826E-5	oui	da-di	0
				da-du	0,00952
				di-du	0,02531

Tableau 2.6.b. 2: Résultats obtenus pour l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h/

b-3- fricative /h_/

b-3-1- Analyses des durées des phrases, des CVC et des durées relatives, en contexte /a/, /i/ /u/

Les figures 2.5.b.3 illustrent l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h_/



Figures 2.5.b. 3: Influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h_/

La comparaison des moyennes des durées à l'aide de l'ANOVA est affichée dans le tableau 2.6.b.3

Fricative /h_ / Locuteur CH variation du contexte vocalique						
	Test statistique		Résultat de l'ANOVA	Résultat de l'ANOVA avec test de Scheffé		
Durée de l'unité	F	p	Significativement différentes	Test de scheffé		Résultats avec test de Scheffé
dph	2,59614	1,36687E-4	oui	Différence de durées entre 2 contextes consonantiques en ms		Significativement différentes
				da-di	0	Non
				da-du	50 ms	Oui
				di-du	46,3ms	oui
dCVC	2,25229	0,12456	non	da-di	0	Non
				da-du	0	Non
				di-du	0	non
dr	1,05028	0,36370	non	da-d	0	Non
				da-du	0	Non
				di-du	0	non

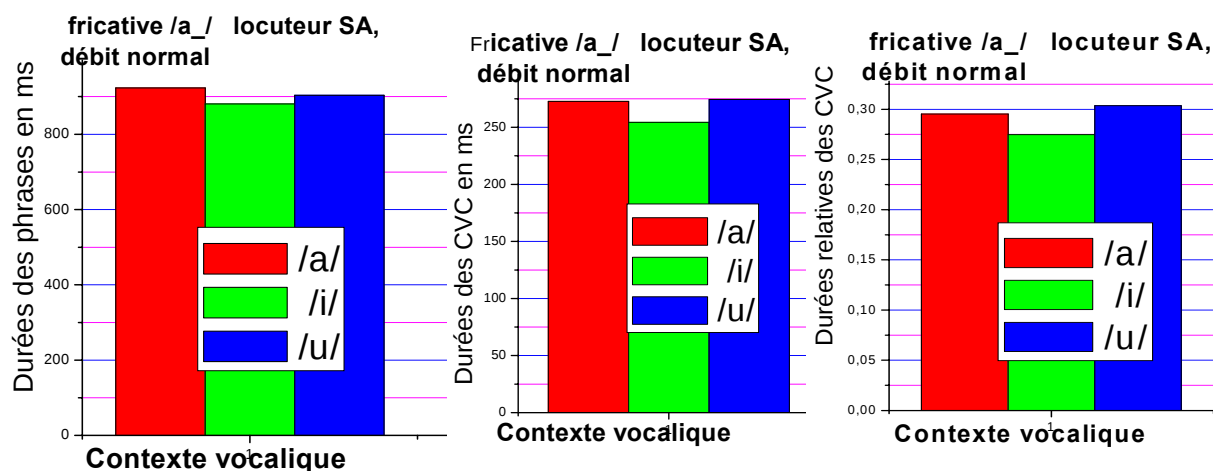
Tableau 2.6.b. 3 : Résultats obtenus pour l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h_/

c- Locuteur SA

c-1- fricative /a_ /

c-1-1- Analyses des durées des phrases en contexte /a/, /i/, /u/

Les figures 2.5.c.1 illustrent l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /a_ /



Figures 2.5.c. 1: Influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /a_ /

La comparaison des moyennes des durées à l'aide de l'ANOVA est affichée dans le tableau 2.6.c.1

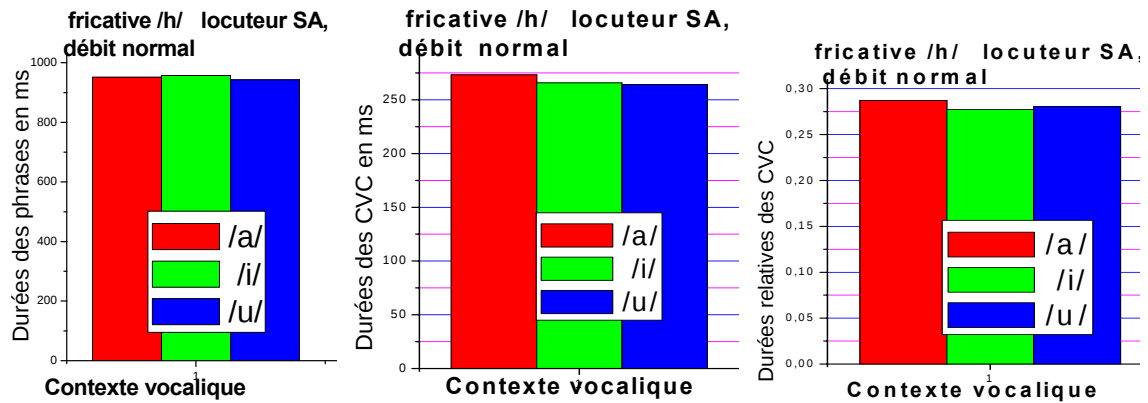
Fricative /a_ / Locuteur SA variation du contexte vocalique						
	Test statistique		Résultat de l'ANOVA	Résultat de l'ANOVA avec test de Scheffé		
Durée de l'unité	F	p	Significativ ement différentes	Test de scheffé		Résultats avec test de Scheffé
dph	5,28537	0,01156	oui	Différence de durées entre 2 contextes consonantiques en ms		Significativeme nt différentes
				da-di	42,8ms	oui
				da-du	0	Non
				di-du	0	non
dCVC	14,72549	4,74073E-5	oui	da-di	18,3ms	Oui
				da-du	0	Non
				di-du	-20 ms	oui
dr	2,83856	0,07605	non	da-di	0	non
				da-du	0	non
				di-du	0	non

Tableau 2.6.c. 1. Résultats obtenus pour l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /a_ /

c-2- fricative /h/

c-2-1- Analyses des durées des phrases en contexte /a/, /i/, /u/

Les figures 2.5.c.1 illustrent l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h/



Figures 2.5.c. 2 : Influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h/

La comparaison des moyennes des durées à l'aide de l'ANOVA est affichée dans le tableau 2.6.c.2

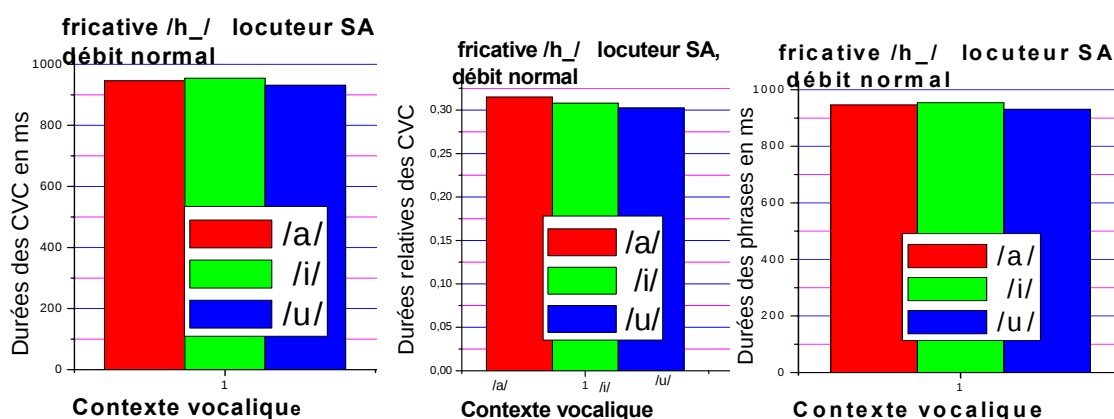
Fricative /h/ Locuteur SA variation du contexte vocalique						
	Test statistique		Résultat de l'ANOVA	Résultat de l'ANOVA avec test de Scheffé		
Durée de l'unité	F	p	Significativement différentes	Test de scheffé		Résultats avec test de Scheffé
dph	0,91735	0,41167	non	Différence de durées entre 2 contextes consonantiques en ms		Significativement différentes
				da-di	0	non
				da-du	0	non
				di-du	0	non
dCVC	1,43219	0,25635	non	da-di	0	non
				da-du	0	non
				di-du	0	non
dr	2,27093	0,12258	non	da-di	0	non
				da-du	0	non
				di-du	0	non

Tableau 2.6.c. 2. Résultats obtenus pour l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h/

c-3- fricative /h_/

c-3-1- Analyses des durées des phrases en contexte /a/, /i/, /u/.

Les figures 2.5.c3 illustrent l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h_/



Figures 2.5.c. 3: Influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h_/

La comparaison des moyennes des durées à l'aide de l'ANOVA est affichée dans le tableau 2.6.c.3

Fricative /h_/ Locuteur SA variation du contexte vocalique							
Durée de l'unité	F	p	Significativement différentes	Test scheffé		Résultats avec test de Scheffé	
dph	2,17960	0,13258	non	Différence de durées entre 2 contextes consonantiques en ms		Significativement différentes	
				da-di	0		non
				da-du	0		non
				di-du	0		non
dCVC	7,75061	0,00219	oui	da-di	0	non	
				da-du	16,5ms	Oui	
				di-du	12,2 ms	oui	
dr	2,64677	0,08920	non	da-di	0	non	
				da-du	0	non	
				di-du	0	non	

Tableau 2.6.c. 3. Résultats obtenus pour l'influence des contextes vocaliques /a/, /i/, /u/ en contexte consonantique /h_/

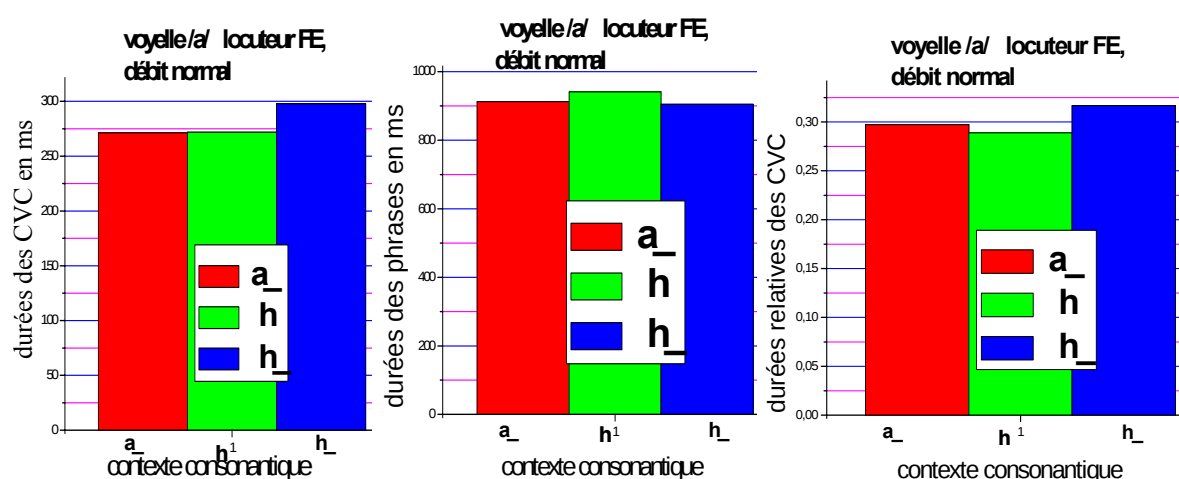
2.7 Etude de l'influence du contexte consonantique, pour un même contexte vocalique pour un même locuteur, sur les unités de parole analysées

a- Locuteur FE

a-1- contexte vocalique /a/

a-1-1- Analyses des durées des phrases, des CVC er des durées relatives en contexte consonantique /a_/, /h/ /h_ /

Les figures 2.6.a. 1 illustrent l'influence des contextes consonantique /a_/, /h/ /h_ / en contexte vocaliques /a/



Figures 2.6.a. 2: Influence des contextes consonantiques /a_/, /h/, /h_ / , en contexte vocalique /a/

La comparaison des moyennes des durées à l'aide de l'ANOVA est affichée dans le tableau 2.7.a. 1

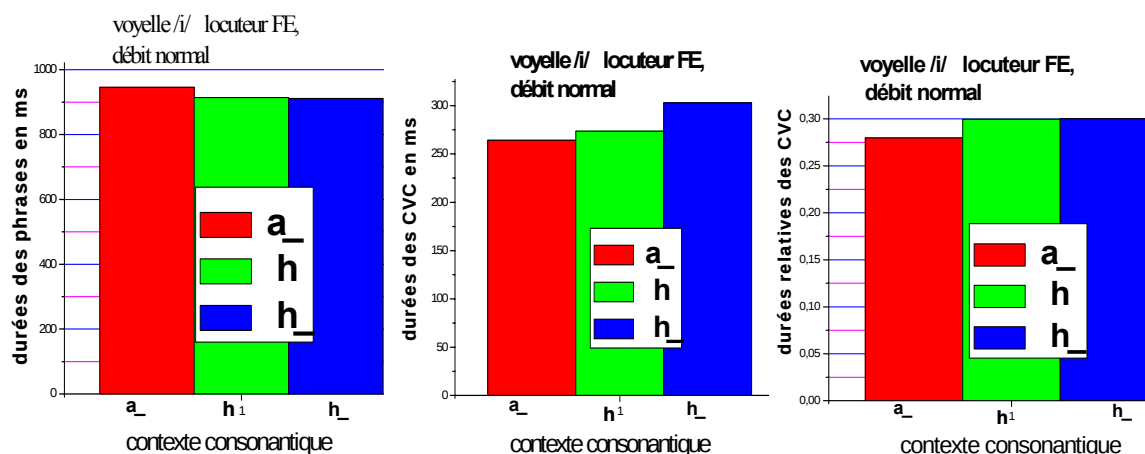
Voyelle /a/ Locuteur FE variation du contexte consonantique						
	Test statistique		Résultat de l'ANOVA	Résultat de l'ANOVA avec test de Scheffé		
Durée de l'unité	F	p	Significativement différentes	Test de scheffé		Résultats avec test de Scheffé
dph	4,79369	0,01654	oui	Différence de durées entre 2 contextes consonantiques en ms		Significativement différentes
				da _dh	0	non
				da _dh_	0	non
				dh-dh	36,5ms	oui
dCVC	11,00654	3,193E-4	oui	da _dh	0	Non
				da _dh_	-26,5	Oui
				dh-dh	-26,1	oui
dr	7,75919	0,00218	oui	da _dh	0	non
				da _dh_	-0,01935	oui
				dh-dh	-0.0279	oui

Tableau 2.7.a. 2 Résultats obtenus pour l'influence des contextes consonantiques /a_/, /h/, /h_ /, en contexte vocalique /a/

a-2- contexte vocalique /i/

a-2-1- Analyses des durées des phrases, des CVC et des durées relatives, en contexte consonantique /a_/, /h/, /h_/.

Les figures 2.6.a. 2 illustrent l'influence des contextes consonantique /a_/, /h/, /h_/ en contexte vocaliques /i/



Figures 2.6.a. 3: Influence des contextes consonantiques /a_/, /h/, /h_/, en contexte vocalique /i/

La comparaison des moyennes des durées à l'aide de l'ANOVA est affichée dans le tableau 2.7.a. 2

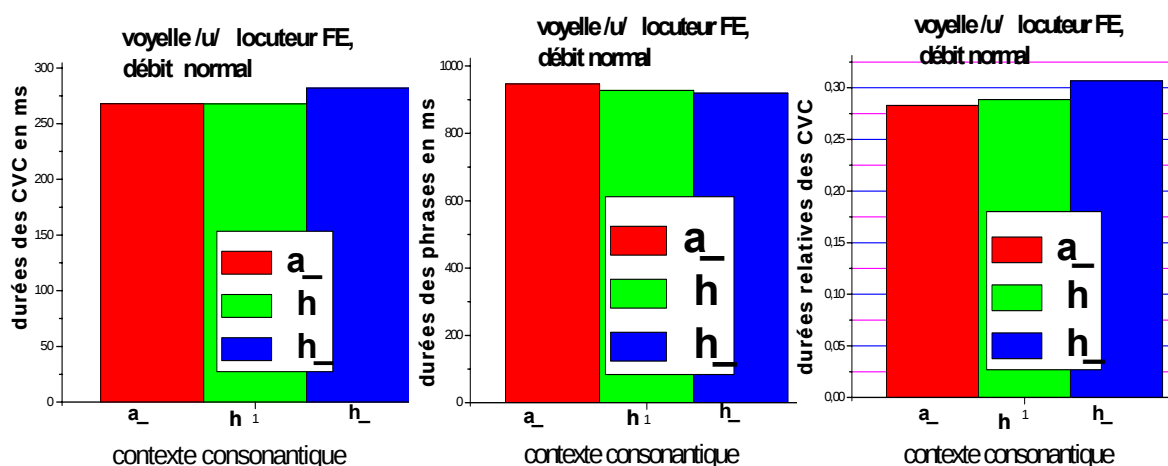
Voyelle /i/ Locuteur FE variation du contexte consonantique						
	Test statistique		Résultat de l'ANOVA	Résultat de l'ANOVA avec test de Scheffé		
Durée de l'unité	F	p	Significativement différentes	Test de scheffé		Résultats avec test de Scheffé
dph	3,72666	0,03722	oui	Différence de durées entre 2 contextes consonantiques en ms		Significativement différentes
				da _dh	31,8ms	oui
				da _dh_	34,2ms	oui
				dh-dh_	2,4ms	non
dCVC	45,93884	2,039E-9	oui	da _dh	0ms	non
				da _dh_	-38,7ms	oui
				dh-dh_	-29,4ms	oui
dr	14,06687	6,520E-5	oui	da _dh	-0,01968	oui
				da _dh_	-0,02048	oui
				dh-dh	0	non

Tableau 2.7.a. 3: Résultats obtenus pour l'influence des contextes consonantiques /a_/, /h/, /h_/, en contexte vocalique /i/

a-3- contexte vocalique /u/

a-3-1- Analyses des durées des phrases, des CVC et des durées relatives, en contexte consonantique /a_/, /h/ /h_ /

Les figures 2.6.a. 3 illustrent l'influence des contextes consonantique /a_/, /h/ /h_ / en contexte vocaliques /u/.



Figures 2.6.a. 4: Influence des contextes consonantiques /a_/, /h/, /h_ / , en contexte vocalique /u/

La comparaison des moyennes des durées à l'aide de l'ANOVA est affichée dans le tableau 2.7.a. 3.

Voyelle /u/ Locuteur FE variation du contexte consonantique						
	Test statistique		Résultat de l'ANOVA	Résultat de l'ANOVA avec test de Scheffé		
Durée de l'unité	F	p	Significativement différentes	Test de scheffé		Résultats avec test de Scheffé
dph	3,52786	0,04354	oui	Différence de durées entre 2 contextes consonantiques en ms		Significativement différentes
				da _dh	19,4	non
				da _dh	27,8	non
				dh-dh	8,4	non
dCVC	4,50816	0,02045	oui	da _dh	0,2	non
				da _dh	-14	non
				dh-dh	-14,2	oui
dr	9,21698	8,88737E-4	oui	da _dh	-0,00582	non
				da _dh	-0,0238	oui
				dh-dh	-0.01805	oui

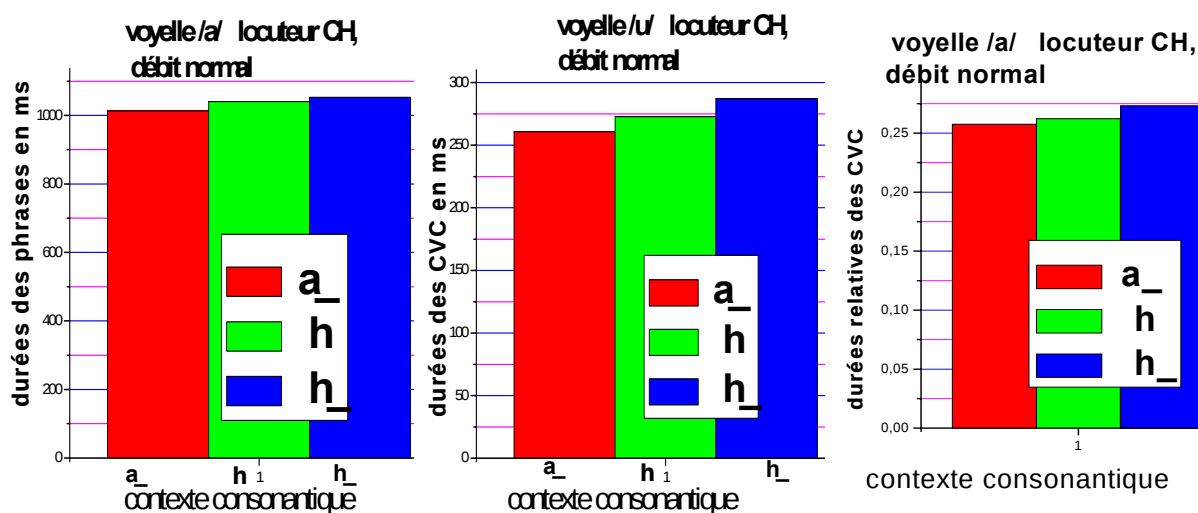
Tableau 2.7.a. 4: Résultats obtenus pour l'influence des contextes consonantiques /a_/, /h/, /h_ / , en contexte vocalique /u/

b- LocuteurCH

b-1- contexte vocalique /a/

b-1-1- Analyses des durées des phrases, des CVC et des durées relatives en contexte consonantique /a_/, /h/ /h_/

Les figures 2.6.b. 1 illustrent l'influence des contextes consonantique /a_/, /h/ /h_/ en contexte vocaliques /a/.



Figures 2.6.b. 1: Influence des contextes consonantiques /a_/, /h/, /h_/, en contexte vocalique /a/

La comparaison des moyennes des durées à l'aide de l'ANOVA est affichée dans le tableau 2.7.b. 1.

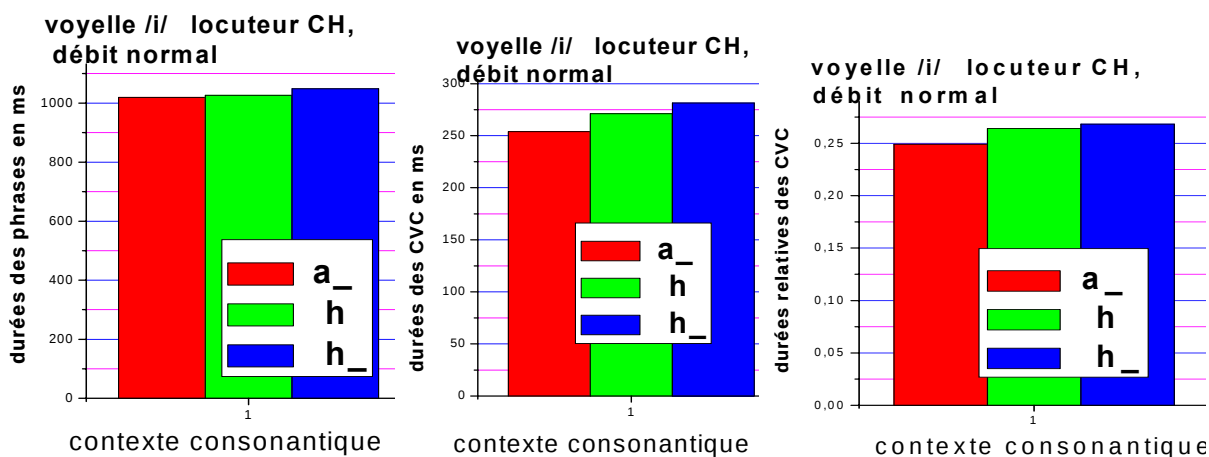
Voyelle /a/ Locuteur CH variation du contexte consonantique					
Test statistique		Résultat de l'ANOVA	Résultat de l'ANOVA avec test de Scheffé		
Durée de l'unité	F	p	Significativement différentes	Test de scheffé	Résultats avec test de Scheffé
dph	5,50327	0,00989	oui	Différence de durées entre 2 contextes consonantiques en ms	Significativement différentes
				da_-dh	-27,6 Non
				da_-dh_	-39,5 Oui
				dh-dh_	-11,9 non
dCVC	8,20080	0,00165	oui	da_-dh	-12 Non
				da_-dh_	-26,5 Oui
				dh-dh_	-14,5 non
dr	3,73493	0,03698	oui	da_-dh	-0,00483 Non
				da_-dh_	-0,01553 Oui
				dh-dh_	-0,0107 non

Tableau 2.7.b. 1: Tableau des résultats obtenus pour l'influence des contextes consonantiques /a_/, /h/, /h_/, en contexte vocalique /a/

b-2- contexte vocalique /i/

b-2-1- Analyses des durées des phrases, des CVC et des durées relatives, en contexte consonantique /a_/, /h/, /h_/

Les figures 2.6.b. 2 illustrent l'influence des contextes consonantique /a_/, /h/, /h_/ en contexte vocaliques /i/.



Figures 2.6.b. 2: Influence des contextes consonantiques /a_/, /h/, /h_/ , en contexte vocalique /i/

La comparaison des moyennes des durées à l'aide de l'ANOVA est affichée dans le tableau 2.7.b.2.

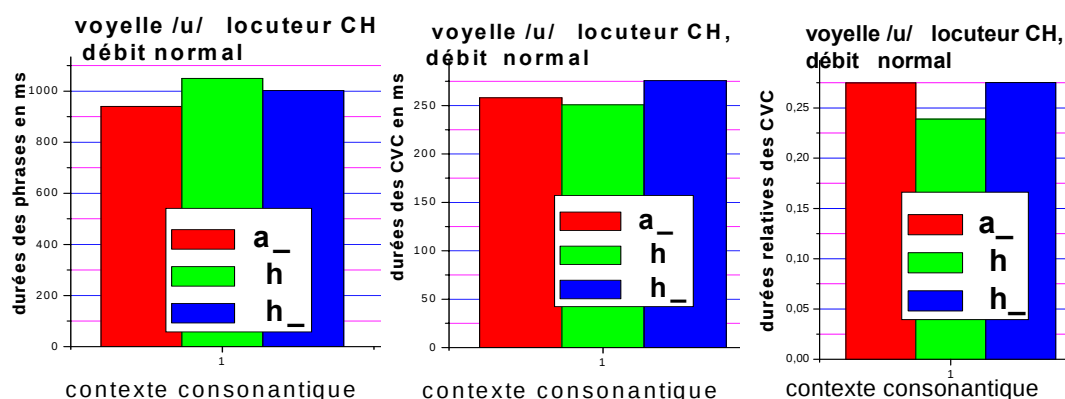
Voyelle /i/ Locuteur CH variation du contexte consonantique					
Test statistique		Résultat de l'ANOVA	Résultat de l'ANOVA avec test de Scheffé		
Durée de l'unité	F		Significativement différentes	Test scheffé	Résultats avec test de Scheffé
dph	3,14012	0,05941	non	Différence de durées entre 2 contextes consonantiques en ms	
				da -dh	-7,5
				da -dh_	-29,6
				dh-dh_	-22,1
dCVC	9,59232	7,123E-4	oui	da -dh	-17,3
				da -dh_	-27,7
				dh-dh_	-10,4
dr	4,87032	0,01563	oui	da -dh	-0,01506
				da -dh_	-0,0192
				dh-dh_	-0,00414

Tableau 2.7.b. 2: résultats obtenus pour l'influence des contextes consonantiques /a_/, /h/, /h_/, en contexte vocalique /i/

b-3- contexte vocalique /u/

b-3-1- Analyses des durées des phrases, des CVC et des durées relatives, en contexte consonantique /a_/, /h/ /h_/

Les figures 2.6.b. 3 illustrent l'influence des contextes consonantique /a_/, /h/ /h_/ en contexte vocaliques /u/.



Figures 2.6.b. 3: Influence des contextes consonantiques /a_/, /h/, /h_/_ , en contexte vocalique /u/

La comparaison des moyennes des durées à l'aide de l'ANOVA est affichée dans le tableau 2.7.b.3.

Voyelle /u/ Locuteur CH variation du contexte consonantique						
Durées	F	p	Significativement différentes	Test scheffé		Significativement différentes
dph	58,38737	1,56565E-10	oui	da_-dh	-109,5	oui
				da_-dh_	-62,4	oui
				dh-dh_	47,1	oui
dCVC	11,22934	2,82577E-4	oui	da_-dh	7,4	non
				da_-dh_	-17,5	oui
				dh-dh_	-24,9	oui
dr	28,880	1,96232E-7	oui	da_-dh	0,0358	Oui
				da_-dh_	-4,1E-4	Non
				dh-dh_	-0,03621	oui

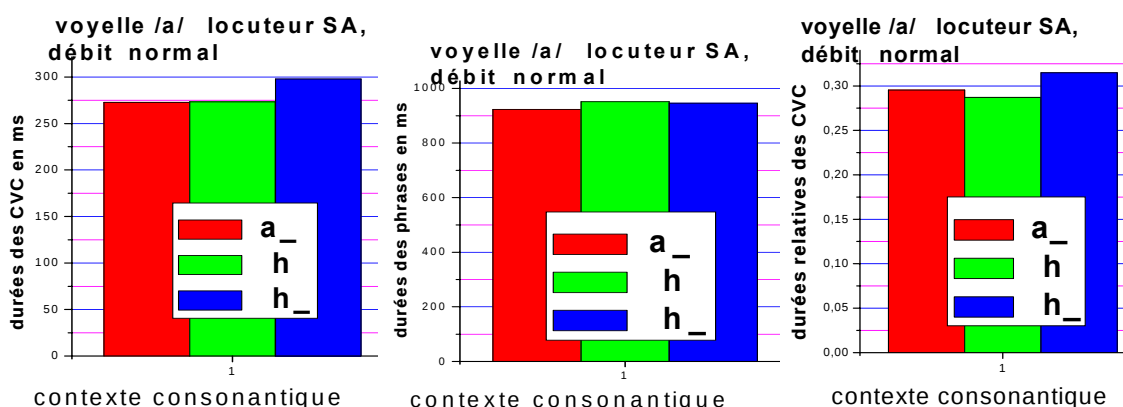
Tableau 2.7.b. 3:Tableau des résultats obtenus pour l'influence des contextes consonantiques /a_/, /h/, /h_/_ , en contexte vocalique /u/

c- Locuteur SA

c-1- contexte vocalique /a/

c-1-1- Analyses des durées des phrases, des CVC et des durées relatives, en contexte consonantique /a_/, /h/ /h_ /

Les figures 2.6.c. 1 illustrent l'influence des contextes consonantique /a_/, /h/ /h_ / en contexte vocaliques /a/.



Figures 2.6.c. 1: Influence des contextes consonantiques /a_/, /h/, /h_ / , en contexte vocalique /a/

La comparaison des moyennes des durées à l'aide de l'ANOVA est affichée dans le tableau 2.7.c.1.

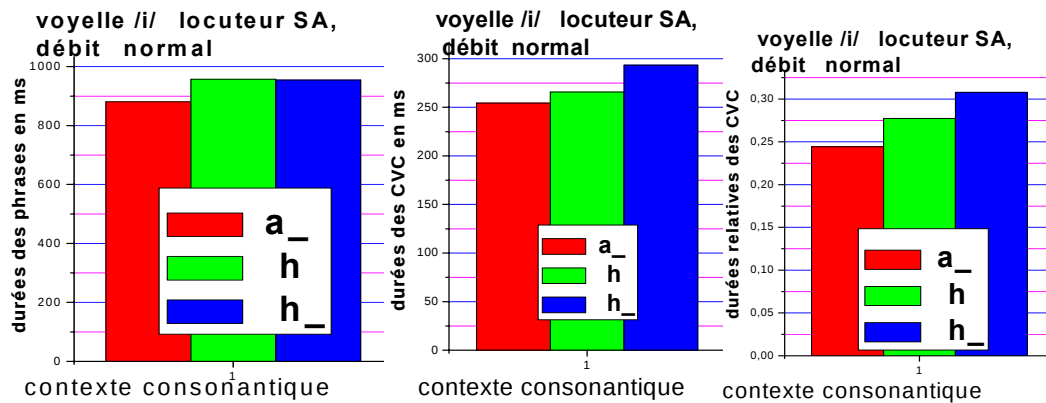
Voyelle /a/ Locuteur SA variation du contexte consonantique						
	Test statistique		Résultat de l'ANOVA	Résultat de l'ANOVA avec test de Scheffé		
Durée de l'unité	F	p	Significativement différentes	Test de scheffé		Résultats avec test de Scheffé
dph	3,90087	0,03249	oui	Différence de durées entre 2 contextes consonantiques en ms		Significativement différentes
				da _dh	-28,1	Oui
				da _dh	-22,8	Non
				dh-dh	5,3	non
dCVC	22,01656	2,1311E-6	oui	da _dh	-0,6	Non
				da _dh	-25,2	Oui
				dh-dh	-24,6	oui
dr	18,4485	8,90037E-6	oui	da _dh	0,00821	Non
				da _dh	-0,0195	Oui
				dh-dh	-0.0277	oui

Tableau 2.7.c. 1: résultats obtenus pour l'influence des contextes consonantiques /a_/, /h/, /h_ / , en contexte vocalique /a/

c-2- contexte vocalique /i/

c-2-1- Analyses des durées des phrases, des CVC et des durées relatives, en contexte consonantique /a_/, /h/ /h_ /

Les figures 2.6.c. 2 illustrent l'influence des contextes consonantique /a_/, /h/ /h_ / en contexte vocaliques /i/.



Figures 2.6.c. 2: Influence des contextes consonantiques /a_/, /h/, /h_ / , en contexte vocalique /i/

La comparaison des moyennes des durées à l'aide de l'ANOVA est affichée dans le tableau 2.7.c.2.

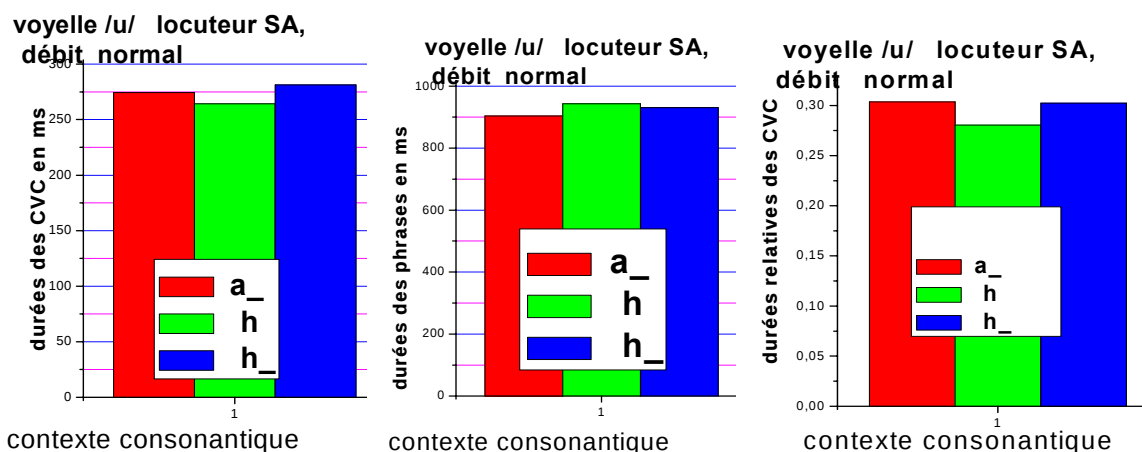
Voyelle /i/ Locuteur SA variation du contexte consonantique						
	Test statistique		Résultat de l'ANOVA	Résultat de l'ANOVA avec test de Scheffé		
Durée de l'unité	F	p	Significativement différentes	Test de scheffé		Significativement différentes
dph	18,23743	9,7339E-6	oui	Différence de durées entre 2 contextes consonantiques en ms		
				da _dh	-76,9	Oui
				da _dh_	-73,8	Oui
				dh-dh_	3,1	non
dCVC	3,21458	1,3622E-6	oui	da _dh	-11,4	Non
				da _dh_	-39,2	Oui
				dh-dh_	-27,8	oui
dr	3,11269	0,06075	non	da _dh	-0,0331	Non
				da _dh_	-0,06358	Non
				dh-dh	-0.03048	non

Tableau 2.7.c. 2: Résultats obtenus pour l'influence des contextes consonantiques /a_/, /h/, /h_ / , en contexte vocalique /i/

c-3- contexte vocalique /u/

c-3-1- Analyses des durées des phrases, des CVC et des durées relatives, en contexte consonantique /a_/, /h/ /h_ /

Les figures 2.6.c. 3 illustrent l'influence des contextes consonantique /a_/, /h/ /h_ / en contexte vocaliques /u/.



Figures 2.6.c. 3: Influence des contextes consonantiques /a_/, /h/, /h_ /, en contexte vocalique /u/

La comparaison des moyennes des durées à l'aide de l'ANOVA est affichée dans le tableau 2.7.c.3.

Voyelle /u/ Locuteur SA variation du contexte consonantique						
	Test statistique		Résultat de l'ANOVA	Résultat de l'ANOVA avec test de Scheffé		
Durée de l'unité	F	p	Significativement différentes	Test de scheffé		Résultats avec test de Scheffé
dph	8,29485	0,00155	oui	Différence de durées entre 2 contextes consonantiques en ms		Significativement différentes
				da _dh	-39,1	oui
				da _dh_	-27,1	oui
				dh-dh_	12	non
dCVC	10,5950	4,013 E-4	oui	da _dh	10	oui
				da _dh_	-7	Non
				dh-dh_	-17	oui
dr	17,06161	1,620E-5	oui	da _dh	0,0232	Oui
				da _dh_	0,00133	Non
				dh-dh_	-0.0218	oui

Tableau 2.7.c. 3: résultats obtenus pour l'influence des contextes consonantiques /a_/, /h/, /h_ /, en contexte vocalique /u/

2.8 Discussion des résultats

2.8.1 Récapitulatif des résultats pour l'étude de l'influence du contexte vocalique sur les durées

Le « non » correspond à : pas significativement différent

Le « oui » correspond à : significativement différent

L'influence du contexte vocalique en contexte consonantique /a_/ est représentée par le tableau 2.8.1.1

/a_/	Durées des unités considérées	Résultats de l'ANOVA	da-di en ms	da-du en ms	di-du en ms
FE	Dph	Pas d'influence	-32,8 non	-34,4 non	-1,6 non
	dCVC	Pas d'influence	7,1 non	3,4 non	-3,7 non
	dr	augmentation en contexte /u/	0,01769 oui	0,01448 non	-0,00321 oui
CH	Dph	augmentation en contexte /a/ et /i/, diminution en contexte /u/	-6,2 non	72,9 oui	79,1 oui
	dCVC	Pas d'influence	7 non	2,6 non	-4,4 non
	dr	Pas d'influence	0,00823 non	-0,01732 non	-0,02555 non
SA	Dph	Augmentation en contexte /a/	42,8 oui	19,5 non	-23,3 non
	dCVC	diminution en contexte /i/	18,3 oui	-1,7 non	-20 oui
	dr	Pas d'influence	0,02065 non	-0,00823 non	-0,02888 non

Tableau 2.8.1. 1:Récapitulatif des résultats de l'influence du contexte vocalique sur les durées en présence de la fricative /a_/

L'influence du contexte vocalique en contexte consonantique /h/ est représentée par le tableau 2.8.1.2

/h/	Durées des unités considérées	Résultats	da-di=	da-du=	di-du=
FE	Dph	Pas d'influence	27,3 non	13,3 non	-14 non
	dCVC	Pas d'influence	-1,8 non	4 non	5,8 non
	dr	Pas d'influence	-0,01054 non	1,1E-4 non	0,01065 non
CH	Dph	Pas d'influence	13,9 non	-9 non	-22,9 non
	dCVC	diminution en contexte /u/	1,7 non	22 yes	20,3 yes
	dr	diminution en contexte /u/	-0,002 non	0,0233 oui	0,02531 oui
SA	Dph	Pas d'influence	-6 non	8,5 non	14,5 non
	dCVC	Pas d'influence	7,5 non	8,9 non	1,4 non
	dr	Pas d'influence	0,00983 non	0,00676 non	-0,00307 non

Tableau 2.8.1. 2:Récapitulatif des résultats de l'influence du contexte vocalique sur les durées en présence de la fricative /h/

L'influence du contexte vocalique en contexte consonantique /h_/ est représentée par le tableau 2.8.1.3

/h_/	Durées des unités considérées	Résultats	da-di=	da-du=	di-du=
FE	Dph	Pas d'influence	29,7 non	21,7 non	-8 non
	dCVC	diminution en contexte /u/	-5,1 non	15,9 oui	21 oui
	dr	Pas d'influence	0,01656 non	0,00996 non	-0,0066 non
CH	Dph	diminution en contexte /u/	3,7 non	50 oui	46,3 oui
	dCVC	Pas d'influence	5,8 non	11,6 non	5,8 non
	dr	Pas d'influence	0,00456 non	-0,0022 non	-0,00676 non
SA	Dph	Pas d'influence	-8,2 non	15,2 non	23,4 non
	dCVC	diminution en contexte /u/	4,3 non	16,5 oui	12,2 oui
	dr	Pas d'influence	0,00709 non	0,01263 non	0,00554 non

Tableau 2.8.1. 3:Récapitulatif des résultats de l'influence du contexte vocalique sur les durées en présence de la fricative /h_/

Nous dirons donc, qu'en présence de ces fricatives, les durées des unités analysées ne semblent pas être influencées par le contexte vocalique

2.8.2 Récapitulatif des résultats pour l'étude de l'influence du contexte consonantique sur les durées

Le « non » correspond à : pas significativement différent

Le « oui » correspond à : significativement différent

L'influence du contexte consonantique en contexte vocalique /a/ est représentée par le tableau 2.8.2.1

/a/	Durées des unités considérées	Résultats	da_-dh	da_-dh_	dh-dh_
FE	Dph	augmentation en contexte /h/	-28,3 oui	8,2 non	36,5 oui
	dCVC	augmentation en contexte /a /	-0,4 non	26.5ms oui	-26,1 oui
	dr	augmentation en contexte /h /	0,00855 no	-0,01935 oui	-0,0279 oui
CH	Dph	augmentation en contexte /h_ /	-27,6 non	-39,5 oui	-11,9 non
	dCVC	augmentation en contexte /h_ /	-12 non	-26,5 oui	-14,5 non
	dr	augmentation en contexte /h_ /	-0,00483 Non	-0,01553 oui	-0,0107 non
SA	Dph	augmentation en contexte /h/	-28,1 oui	-22,8 non	5,3 non
	dCVC	augmentation en contexte /h_ /	-0,6 non	-25,2 oui	-24,6 oui
	dr	augmentation en contexte /h_ /	0,00821 non	-0,01953 oui	-0,02774 oui

Tableau 2.8.2. 1:Récapitulatif des résultats de l'influence du contexte consonantique sur les durées en présence de la voyelle /a/

Conclusion du tableau : Allongement des durées relatives en contexte vocalique /a/ pour la fricative /h_/. Pour le /h/ et /a_ / aucun allongement en contexte vocalique /a/

L'influence du contexte consonantique en contexte vocalique /i/ est représentée par le tableau 2.8.2.2.

/i/	Durées des unités considérées	Résultats	da_-dh	da_-dh_	dh-dh_
FE	Dph	Pas d'influence	31,8 non	34,2 non	2,4 non
	dCVC	augmentation en contexte /h_/	-9,3 non	-38,7 oui	-29,4 oui
	dr	augmentation en contexte /h_/ et /h/, /a_/ est le plus petit	-0,01968 oui	-0,02048 oui	-8E-4 non
CH	Dph	Pas d'influence	-7,5 non	-29,6 non	-22,1 non
	dCVC	augmentation en contexte /h_/ et /h/, /a_/ est le plus petit	-17,3 oui	-27,7 oui	-10,4 non
	dr	augmentation en contexte /h_/	-0,01506 non	-0,0192 oui	-0,00414 non
SA	Dph	augmentation en contexte /h/ et /h_/, /a_/ est le plus petit	-76,9 oui	-73,8 oui	3,1 non
	dCVC	augmentation en contexte /h_/	-11,4 non	-39,2 oui	-27,8 oui
	dr	Pas d'influence	-0,0331 non	-0,06358 non	-0,03048 non

Tableau 2.8.2. 2:Récapitulatif des résultats de l'influence du contexte consonantique sur les durées en présence de la voyelle /a/

Conclusion : Généralement nous avons un allongement des durées en présence de la fricative /h_/ quelque soit le contexte vocalique. Quelques allongements sont observés aussi en présence du /h/.

L'influence du contexte consonantique en contexte vocalique /u/ est représentée par le tableau 2.8.2.1

/u/	Durées des unités considérées	Résultats	da_-dh	da_-dh_	dh-dh_
FE	Dph	Pas d'influence	19,4 non	27,8 non	8,4 non
	dCVC	augmentation en contexte /h_/	0,2 non	-14 non	-14,2 oui
	dr	augmentation en contexte /h_/	-0,00582 non	-0,02387 oui	-0,01805 oui
CH	Dph	augmentation en contexte /h/ et /h_/_	-109,5 oui	-62,4 oui	47,1 oui
	dCVC	augmentation en contexte /h_/_	7,4 non	-17,5 oui	-24,9 oui
	dr	augmentation en contexte /h_/_ et /a_/_	0,0358 oui	-4,1E-4 non	-0,0362 oui
SA	Dph	augmentation en contexte /h/ et /h_/_	-39,1 oui	-27,1 oui	12 non
	dCVC	augmentation en contexte /h_/_ et du /h/	10 oui	-7 non	-17 oui
	dr	augmentation en contexte /h_/_ et du /h/	0,0232 oui	0,00133 non	-0,02187 oui

Tableau 2.8.2. 3:Récapitulatif des résultats de l'influence du contexte consonantique sur les durées en présence de la voyelle /a/

Conclusion : Même remarque, en général, allongement des durées en présence du /h_/ et du /h/ (pour le /h/ en contexte /i/ et /u/) dans tous les contextes vocaliques.

2.8.3. Conclusion

a. Influence du contexte vocalique

D'après les résultats obtenus réunis dans les tableaux 2.4 :

Nous ne notons globalement aucune influence du contexte vocalique sur les durées analysées et ce quelque soit la fricative considérée.

b. Influence du contexte consonantique

D'après les résultats obtenus réunis dans les tableaux 4.5 :

En présence de la fricative /h_/, nous avons un allongement des durées pour tous les contextes vocaliques et tous les locuteurs.

En contexte /i/ et /u/ nous constatons aussi un allongement des durées du à la présence du /h/. Cet allongement est moins fréquent que pour la présence du /h_/.

Par contre, la présence du /a_/ ne provoque aucun allongement quelque soit le contexte vocalique et le locuteur.

Nous pensons que puisque la pharyngale neutre /h/ est aussi concernée par cet allongement et que d'après A.Djeradi [1], le pharynx est complètement ouvert durant la production du /h/, seul le rétrécissement laryngale qui est d'ailleurs similaire aux autres configurations, reste comme cause plausible de la phonation de cette consonne, et plusieurs auteurs affirment ne voir aucun mouvement musculaire se produire au cours de l'émission de cette consonne. Donc, la variation de ces durées ne peut être due aux muscles pharyngaux mais au rétrécissement laryngale.

En effet, cet allongement diminue avec le voisement des consonnes fricatives considérées. Le caractère voisement semble nécessiter moins d'effort. Cela pourrait s'expliquer par le fait que la production des consonnes sourdes n'exige pas que les cordes vocales soient accolées. La dépense d'air est nettement plus importante que lors de la production des sons voisés. Donc, plus d'apport d'air des poumons est nécessaire pour produire le /h_/ que le /h/ et le /a_/.

Là encore, comme pour l'influence du contexte vocalique, nous voyons que le lieu d'articulation n'a aucun impact sur le timing pour ces trois fricatives.

En somme nous dirons que les lieux d'articulation des pharyngales utilisées, en contextes vocaliques, semblent n'avoir aucun effet sur le timing de la parole. Nous pensons que cela est du au fait que les lieux d'articulation de ces trois fricatives sont très proches. Ce qui doit provoquer des effets inappréciables au plan temporel.

2.9 Bibliographie du chapitre 2

- [1] A.Djeradi "Analyse et caractérisation des fricatives d'arrière de l'arabe .Acoustique, articulatoire et timing. Thèse de doctorat d'état en électronique. USTHB . Alger. 1992
- [2] Stevens K.N & Klatt "Role of formant transitions in the voiced-voiceless distinctions for stops".J.Acoust.Soc.Am.55, 653-659. 1974
- [3] Butcher A. Ahmad K. " Some Acoustic and Aerodynamis Characteristics of Pharyngeal Consonants in Iraqi Arabic". *Phonetica* 44, 156-172. 1987
- [4] C. Zeroual. A fiberscopic and acoustic study of « guttural » and emphatic consonants of Moroccan Arabic. *Proceedings of the XIVth International Congress of Phonetic Sciences*, San Francisco: 997- 1000, 1999.
- [5] C. Zeroual. "*Propos controversés sur la phonétique et la phonologie de l'arabe marocain*". Thèse de Doctorat, Université de Paris 8, 2000.
- [6] C. Zeroual C. and L. Crevier-Buchman." L'arabe marocain possède des consonnes epiglottales et non pharyngales"XXIVèmes *Journées d'Etude sur laParole*, 24-27 Juin 2002, Nancy, France : 237-240, 2002.
- [7] C. Zeroual. "Aerodynamic study of Moroccan Arabic guttural consonants". *Proceedings of the XVth International Congress of Phonetic Sciences, Barcelona*: 1859-1862, 2003.
- [8] Chakir Zeroual, Lise Crevier-Buchman & John H. Esling. "Ajustements laryngaux durant les consonnes gutturales et l'emphatique [S]"simples et géminées de l'arabe marocain. JEP2004.
- [9] Ramousse R., Le Berre M. & Le Guelte L. « Introduction aux statistiques » - © 1996,

Chapitre 3

La variation de débits et le contrôle des gestes de la parole.

3.1 Introduction

Une préoccupation centrale serait de comprendre comment se fait le passage du niveau symbolique de la description phonétique aux niveaux articulatoire et acoustique. A partir de ceci, on peut dire qu'une exigence fondamentale pour les modèles de production serait de permettre le passage entre les différents niveaux de la chaîne de production de la parole : depuis le niveau symbolique jusqu'aux réalisations articulatoire et acoustique en passant par le processus d'encodage et de contrôle moteur. L'un des points que nous voulons traiter est le contrôle vers les cibles. On doit alors définir le comportement dynamique des articulateurs et le mode de gestion de la coordination des articulateurs. Il est bien admis (cf. Théorie quantique) qu'en parole il y a des cibles, donc les gestes de la parole sont des gestes optimisés, résultant d'un apprentissage, destinés à opérer le passage de cibles à cibles. Le contrôle moteur introduit l'idée que les gestes de la parole seraient comparables à des mouvements de pointage de cibles, objet d'étude spécifique dans le champ du contrôle moteur. Nous sommes dans un champ de recherche lieu de controverses :

- Quelles spécifications pour les cibles? Cible équilibre (Bizzi) ou spatiale (Mac Neilage)?
- Quelles sont les paramètres contrôlés? Force, vitesse, ou longueur ? (Stein 82)
- Quel est le statut du temps ? variable intrinsèque (Fowler) ou commande extrinsèque ? (Schmidt)

C'est cette question que nous voulons examiner ici. Cette étude porte sur les gestes articulatoires occasionnés lors d'une variation de débit dans la production de phrases courtes et interrogatives en arabe : la variation de la vitesse d'élocution. Le but est de montrer comment chaque locuteur agit sur ses patrons temporels et acoustiques pour produire une parole encore avec une certaine intelligibilité.

3.2. Méthodologie

Nous utilisons le débit comme paradigme de recherche, appliqué à un corpus de phrases arabes.

Les phrases du corpus sont une expression de l'arabe standard c'est la langue arabe apprise à l'école pour les six locuteurs. Les six locuteurs sont des étudiants.

- CH et LI (LI: locutrice), libanais, de langue maternelle libanaise, vivant au Liban.

- FE et NA : sont des algériens, vivant à Alger

- SA et MA sont des algériens, vivant à Tizi Ouzou, région berbérophone.

Le matériau linguistique utilisé est le corpus de phrases défini est réalisé au chapitre 2.

3.2.1. Les unités de parole analysées:

Comme souligné dans le chapitre 2, mais nous le rappelons encore dans ce chapitre, il s'agit des unités /CjVil/, Vi étant les trois voyelles de l'arabe (i=/a/, /i/, /u/), en contexte consonantique /Cj/: j=/a_/, /h/, /h_/, pour chaque locuteur, soit :

/a_al/, /a_il/, /a_ul/, /hal/, /hil/, /hul/, /h_al/, /h_il/, /h_ul/

Pour obtenir ces différentes unités, nous avons segmenté manuellement les phrases du corpus en se basant sur le signal acoustique et sur le spectrogramme, à l'aide du logiciel Praat. L'analyse a été faite dans le domaine temporel et dans le domaine fréquentiel.

Dans le domaine temporel nous avons analysé les durées relatives d(CjVil)/d(phrase), (la phrase correspondante à l'unité /CjVil/.

Dans le domaine fréquentiel, nous avons analysé les formants F1 et F2 de la voyelle Vi (au cœur de la voyelle, prélevés manuellement à l'aide du logiciel Pratt).

3.3. Analyse des durées relatives d(CjViL)/ d(phrase)

Le but est d'analyser les durées temporelles mises par chaque locuteur pour produire les fricatives en fonction du changement de débit. On pourra obtenir les patrons temporels de chaque locuteur.

Nous avons calculé les valeurs moyennes des durées relatives d(CjVil)/d(phrase), en fonction du débit et du contexte consonantique, pour chaque locuteur.

3.3.1 Résultats obtenus pour les moyennes des durées relatives

Ils sont illustrés par les tableaux 3.1. Nous avons obtenus trois tableaux. Chaque tableau contient les durées relatives moyennes d(CjVil)/d(phrase, par contexte consonantique (/a_/, /h/, /h_/). A l'intérieur de ce tableau nous trouvons les durées relatives moyennes par locuteur, par débit et par contexte vocalique (/a/, /i. /u/)

		Contexte consonantique /a_/		
locuteur	débit	a_al/ph	a_il/ph	a_ul/ph
FE	N	0,29738	0,27969	0,2829
FE	R	0,34181	0,33866	0,33107
FE	L	0,25439	0,25021	0,24646
loc	débit			
CH	N	0,27278	0,26423	0,26434
CH	R	0,26856	0,27956	0,27322
CH	L	0,27047	0,26557	0,27008
loc	débit			
SA	N	0,2954	0,24426	0,30363
SA	R	0,36495	0,38015	0,36766
SA	L	0,26369	0,26247	0,25374
loc	débit			
MA	N	0,28132	0,27309	0,2795
MA	R	0,36962	0,35033	0,31978
MA	L	0,29857	0,27326	0,28682
loc	débit			
LI	N	0,26252	0,26281	0,25937
LI	R	0,27637	0,27243	0,28034
LI	L	0,26196	0,2685	0,25823
Loc	débit			
NA	N	0,25738	0,2524	0,26816
NA	R	0,36343	0,35224	0,33244
NA	L	0,30756	0,30195	0,30933

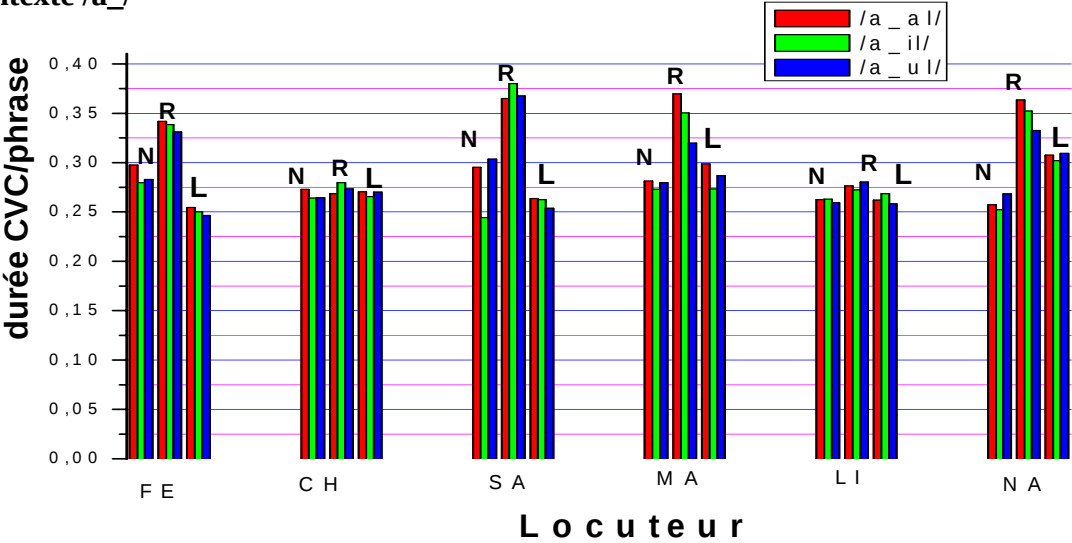
		Contexte consonantique /h/		
locuteur	débit	hal/ph	hil/ph	hul/ph
FE	N	0,29081	0,29158	0,29453
FE	R	0,33578	0,33183	0,3439
FE	L	0,27862	0,2771	0,27966
loc	débit			
CH	N	0,25927	0,25627	0,24894
CH	R	0,25441	0,2514	0,25104
CH	L	0,24589	0,26053	0,25121
loc	débit			
SA	N	0,28606	0,2783	0,2806
SA	R	0,40538	0,41012	0,4119
SA	L	0,32256	0,26165	0,25551
loc	débit			
MA	N	0,28454	0,26435	0,27322
MA	R	0,3124	0,36091	0,37574
MA	L	0,27514	0,28711	0,2922
loc	débit			
LI	N	0,27815	0,27709	0,27411
LI	R	0,28	0,28095	0,27309
LI	L	0,27982	0,27427	0,27759
Loc	débit			
NA	N	0,25992	0,26121	0,25479
NA	R	0,35088	0,36376	0,3522
NA	L	0,25315	0,30011	0,26254

		Contexte consonantique /h_/		
locuteur	débit	h_al/ph	h_il/ph	hul/ph
FE	N	0,30202	0,30851	0,31314
FE	R	0,38508	0,37363	0,37277
FE	L	0,30275	0,29672	0,30118
loc	débit			
CH	N	0,27299	0,27495	0,27352
CH	R	0,2861	0,2897	0,28625
CH	L	0,27822	0,27727	0,27225
loc	débit			
SA	N	0,3117	0,30994	0,30343
SA	R	0,35182	0,34073	0,35638
SA	L	0,28067	0,27774	0,28091
loc	débit			
MA	N	0,26381	0,24412	0,25953
MA	R	0,34728	0,33445	0,3471
MA	L	0,26463	0,27334	0,26047
loc	débit			
LI	N	0,29122	0,26723	0,28736
LI	R	0,29178	0,29006	0,29943
LI	L	0,29105	0,28441	0,29362
Loc	débit			
NA	N	0,28431	0,31775	0,30337
NA	R	0,36282	0,34405	0,33235
NA	L	0,2913	0,3095	0,3129

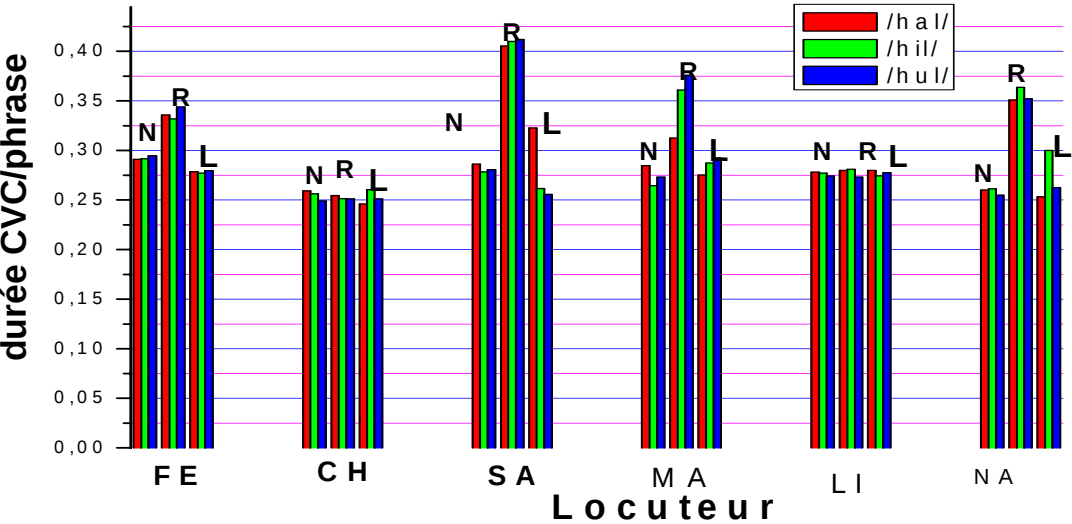
Tableau 3. 1. Variations moyennes des durées CjVil/phrased

Les figures 3.1 obtenues à partir des tableaux 3.1, illustrent la distribution du temps dans la phrase par fricative, en fonction du locuteur et du débit.

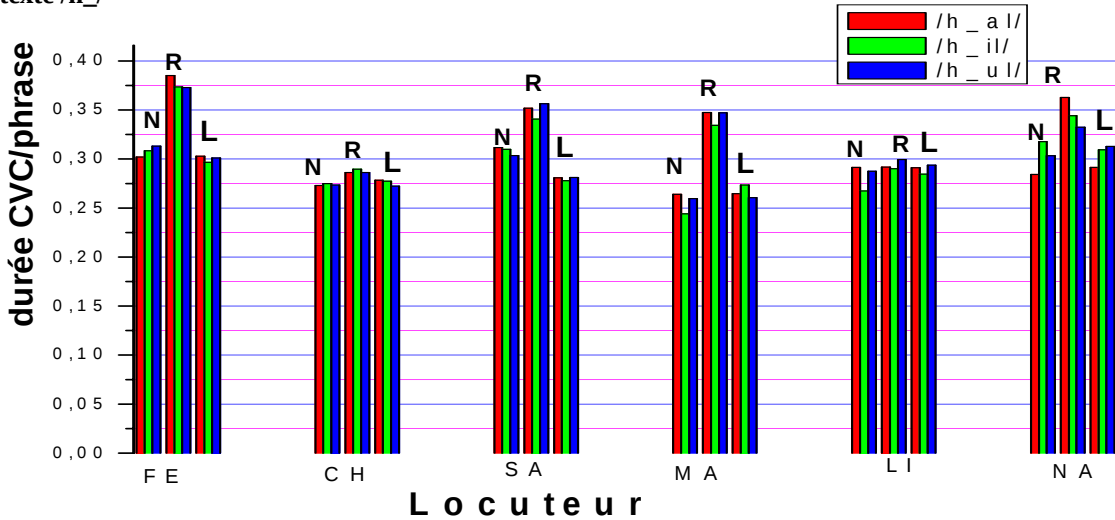
Contexte /a_/



Contexte /h/



Contexte /h_/



Figures 3. 1. Représentation des variations des durées relatives des CjVil/phrase

3.3.2 Interprétation des résultats

Si l'on considère la majorité de ces bâtonnets, on voit que la durée rapide relative reste quand même plus importante que les autres, alors les débits lent et normal ne sont pas très différents. Par ailleurs les locuteurs CH et LI montrent des durées relatives constantes quelque soit le débit. Pour plus de rigueur dans ces résultats, nous allons appliquer la méthode de l'analyse de la variance "ANOVA" aux 540 stimuli de chaque fricative. Cette méthode comme décrite dans le chapitre 2, va permettre de comparer les moyennes des différentes durées relatives représentées dans les figures 3.2 entre elles au sens statistique, afin de voir si les résultats que nous venons de trouver restent inchangés et asseoir ainsi une conclusion quant à la distribution temporelle du CVC dans la phrase.

L'ANOVA a deux facteurs intra : contexte vocalique (/a/, /i/ /u/) et débit (normal, rapide et lent) pour une même fricative.

Nous avons mélangé les données des trois contextes vocaliques aux trois différents débits, avec dix répétitions par contexte vocalique, nous avons obtenu 540 stimuli par fricative et pour les 6 locuteurs.

- L'hypothèse nulle H0 : les moyennes des durées relatives du groupe ne sont pas significativement différentes
- L'hypothèse H1 : les moyennes des durées relatives du groupe sont significativement différentes

Les résultats obtenus pour l'ANOVA sont regroupés dans le tableau 3.2.

Pour les locuteurs CH et LI : la valeur de la probabilité p étant largement supérieure au seuil de signification 0.05, on conserve donc l'hypothèse nulle, ce qui rejoint notre remarque première, lisible sur le graphe à bâtonnets.

Pour les locuteurs FE, SA, NA et MA : la valeur de p étant très en deçà du seuil de signification 0.05, on conserve donc l'hypothèse H1

Par ailleurs, d'après le test de Scheffé qui permet de comparer les moyennes deux à deux, les moyennes obtenues pour les durées relatives intra débit et inter contextes vocaliques, ne sont pas significativement différentes pour les six locuteurs (globalement).

Nous avons aussi représenté en figure 3.2, les valeurs des F (rapport entre variance inter groupe/ variance intra groupe) du test statistique

D'après le tableau 3.2 et la figure 3.2, les résultats des F et des p du test statistique, obtenus pour chacun des locuteurs, FE et MA présentent les différences les plus significatives (en particulier en contexte consonantique /a_ / et /h/).

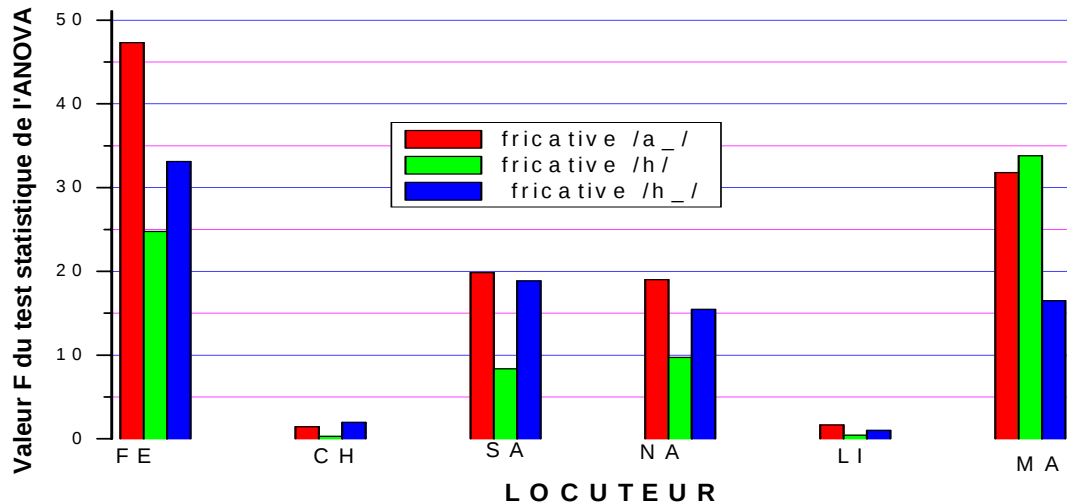
Ce qui nous laisse penser que

- **les locuteurs CH et LI** maîtrisent bien le geste pharyngale car le temps ne varie pas avec le débit d'une part, avec le lieu articulaire d'autre part, le contrôle du pharynx est plus stable

- **les locuteurs FE, MA, NA et SA** agissent différemment sur le muscle pharyngale, la réalisation de la constriction n'est pas précise d'où un temps plus lent pour atteindre cette réalisation en débit rapide.

Locuteur	Fricative /a_/	F	p	ANOVA des moyennes des durées relatives pour les trois débits, pour les 6 locuteurs à 0.05	Test de scheffé
FE		47,326	0	significativement différentes	Pas significativement différentes intra débit mais significativement différente inter débit
CH		1,437	0,193	pas significativement différentes	Pas significativement différentes intra débit et inter débit
SA		19,870	3,33067E-16	significativement différentes	Pas significativement différentes intra débit mais significativement différente inter débit
NA		19,018	8,88178E-16	significativement différentes	Pas significativement différentes intra débit mais significativement différente inter débit
LI		1,645	0,1248	pas significativement différentes	Pas significativement différentes intra débit et inter débit
MA		31,794	0	significativement différentes	Pas significativement différentes intra débit mais significativement différente inter débit
	/h/				
FE		24,755	0	significativement différentes	Pas significativement différentes intra débit mais significativement différente inter débit
CH		0,2975	0,9649	pas significativement différentes	Pas significativement différentes intra débit et inter débit
SA		8,355	6846E-8	significativement différentes	Pas significativement différentes intra débit mais significativement différente inter débit
NA		9,70237	2,44338E-9	significativement différentes	Pas significativement différentes intra débit mais significativement différente inter débit
LI		0,4079	0,9129	pas significativement différentes	Pas significativement différentes intra débit et inter débit
MA		33,817	0	significativement différentes	Pas significativement différentes intra débit mais significativement différente inter débit
	/h_/				
FE		33,134	0	significativement différentes	Pas significativement différentes intra débit mais significativement différente inter débit
CH		1,9593	0,0622	pas significativement différentes	Pas significativement différentes intra débit et inter débit
SA		18,86	1,11022E-15	significativement différentes	Pas significativement différentes intra débit mais significativement différente inter débit
NA		15,470	1,4444E-13	significativement différentes	Pas significativement différentes intra débit mais significativement différente inter débit
LI		0,980	0,4576	pas significativement différentes	Pas significativement différentes intra débit et inter débit
MA		16,485	3,20854E-14	significativement différentes	Pas significativement différentes intra débit mais significativement différente inter débit

Tableau 3. 2 : Résultats de l'analyse de la variance (ANOVA) des durées relatives par locuteur



Figures 3. 2. Valeurs des F du test statistique en fonction du contexte consonantique pour chaque locuteur

On pourrait en conclure que FE, MA, NA et SA ont approximativement la même stratégie de contrôle du pharynx

3.3.3. Discussions :

D'après ces résultats :

* En débits normal et lent :

Nous pouvons dire que les locuteurs adoptent approximativement tous la même stratégie sur le plan temporel, puisqu'ils ont tous approximativement la même durée relative pour les unités /CjVil/.

* En débit rapide :

- Ch et LI adoptent la même stratégie temporelle qu'en débits normal et lent, puisque leurs durées relatives demeurent inchangées et du même ordre de grandeur.

- FE et SA, NA et MA augmentent leurs durées relatives par rapport aux débits normal et lent (Les F du test statistique sont plus significatifs pour les locuteurs FE et NA par rapport à SA et MA) pour reproduire leurs contrastes phonémiques des unités considérées.

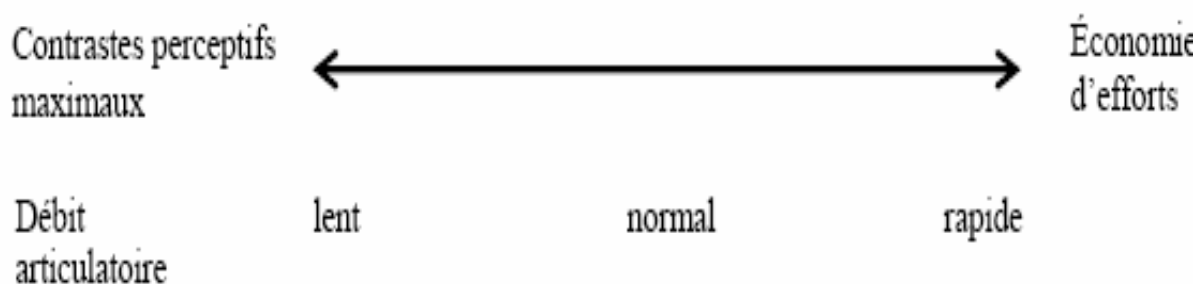
Ils allouent donc plus de temps dans la phrase pour pouvoir reproduire les mêmes unités produites en débit normal. Cela veut dire qu'il leur sera difficile de reproduire les unités considérées si le temps qui leur est alloué est relativement court par rapport au débit normal. Donc au plan temporel, CH et LI ont plus de facilité à reproduire les unités considérées même en contrainte temporelle. Ce qui n'est pas le cas des locuteurs SA, NA, MA et FE (FE et MA plus que SA et NA : d'après la figure 3.2)

D'après Perkell *et al.* (2000) [16], les mécanismes pour la production des contrastes phonémiques sont basés sur des buts auditifs. Ainsi, la production de la parole est programmée dans le système nerveux central pour atteindre des séquences de buts acoustiques.

Comme ces buts acoustiques sont en général construits dès la plus tendre enfance par le biais de la langue maternelle, donc le locuteur produira ses cibles acoustiques en fonction de modèles de patrons acoustiques pré enregistrés dans son cerveau, à l'âge adulte. C'est pour cela que nous obtiendrons des durées relatives propres à chaque locuteur. C'est ce qui lui permet d'atteindre ses cibles acoustiques. Voyons alors ce qu'il en est dans le domaine acoustique.

3.4. Analyses acoustiques des voyelles arabes Vi (i=/a/, /i/, /u/) des unités /CjVil/, en contrainte de débit.

Perkell *et al.* (2000) [16] rajoutent dans leur théorie de production de la parole sur la base de buts auditifs, que la dimension articulatoire est régie inconsciemment par le locuteur: le degré d'articulation d'un locuteur donné dans une condition de parole donnée, varie entre son désir d'obtenir les meilleurs contrastes acoustiques possibles, et donc les meilleurs contrastes articulatoires, et sa volonté de minimiser ses patrons articulatoires dans une perspective d'économie d'efforts (figure 3.3 perkell 2000 [16]). C'est sur cet axe que le débit joue un rôle très important.



Figures 3. 3. Axe de production de la parole selon la contrainte du débit

Par exemple, si on demande à un locuteur adulte ayant un contrôle moteur fin, d'une part de parler lentement, il augmentera l'amplitude de ses contrastes acoustiques, (donc la distance formantique entre les voyelles sera plus grande) ce qui résultera par ailleurs en un effort plus marqué au plan articulatoire.

D'autre part, s'il parle plus rapidement, il diminuera son activité articulatoire, ce qui aura pour effet de minimiser les efforts du locuteur et se traduira par une baisse des contrastes acoustiques (diminution des distances entre des paires de voyelles dans l'espace acoustique).

Cela veut dire qu'en situation de contrainte, l'espace vocalique varie par rapport à l'espace vocalique occupé en débit normal.

Toutefois, en débit rapide, la baisse de ces contrastes acoustiques ne signifie pas que le locuteur devient inintelligible.

Il serait donc intéressant pour nous de savoir si ces locuteurs, malgré le fait qu'ils n'aient pas la même stratégie, ont des espaces vocaliques qui varient avec le débit ? Est-ce que cette variation de l'espace vocalique est un signe de bon contrôle moteur. Comment en être sûr ?

Gay (1978) [9] propose deux stratégies de contrôle moteur en production de la parole :

- La stratégie d'amplitude (*amplitude strategy*)
- ou la stratégie de vitesse (*velocity strategy*).

Dans le premier cas, un locuteur A garde à peu près le même débit articulatoire (la distribution du temps est la même) qu'en condition de parole normale, mais il réduit l'amplitude de ses mouvements (réduction vocalique). Il centralise davantage ses mouvements articulatoires. Pour cette stratégie, on retrouve beaucoup plus de coarticulation [5]. La coarticulation se produit lorsqu'une voyelle ou une consonne voit ses caractéristiques changer à cause de son environnement phonétique. La configuration des articulateurs pour une voyelle donnée peut alors être influencée par les phonèmes précédents ou suivants.

Pour la seconde stratégie, un locuteur B garde la même amplitude de mouvement (réduction vocalique minime) et maximise ses patrons articulatoires, Il compense en accélérant le mouvement de ses articulateurs (augmente la distribution du temps).

Pour cette stratégie, les cibles acoustiques atteintes sont sensiblement les mêmes que celles en condition normale. Il y a moins de centralisation.

Cette deuxième stratégie témoigne selon nous d'un contrôle moteur moindre, moins fin.

En effet, pour un but acoustique semblable produit en condition identique, le locuteur fournit plus d'efforts.

Nous pensons que CH et LI (libanais) utilisent la stratégie d'amplitude car ils gardent leurs durées relatives constantes, comme en débit normal, reste à vérifier s'ils réduisent l'amplitude de leurs mouvements, c'est-à-dire s'ils réduisent leurs espaces vocaliques en augmentant le débit, donc est ce qu'ils centralisent ?

Nous classerons par contre les 4 autres locuteurs SA, MA (kabyles), NA et FE (algerois) dans le 2ème cas (stratégie de vitesse (+*velocity strategy*)), car en effet, en débit rapide, les durées relatives sont plus élevées, donc différentes des durées en débit normal, donc maximisent leurs patrons articulatoires. Reste à vérifier si l'espace acoustique est presque le même qu'en débit normal. Il faudra vérifier le degré de centralisation.

Pour vérifier cela, nous commencerons par une analyse de la variation de l'espace vocalique.

3.4.1 Méthodologie

Pour calculer l'espace vocalique occupé par les voyelles V_i , nous avons calculé les surfaces des triangles vocaliques des voyelles arabes /a/, /i/ et /u/, pour chaque débit et chaque fricative, dans le plan F1, F2.

Nous avons tracé les ellipses de dispersions et les sommets des triangles vocaliques sont les centres des ellipses de dispersion de chaque voyelle (/a/, /i/, /u/).

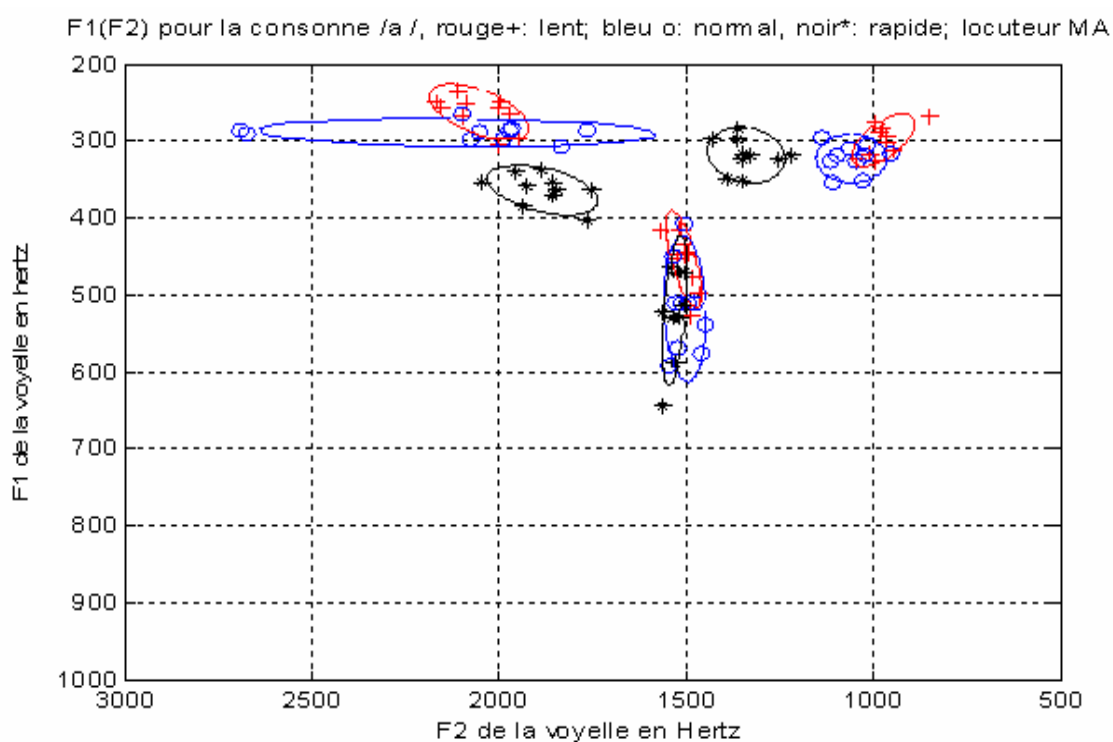
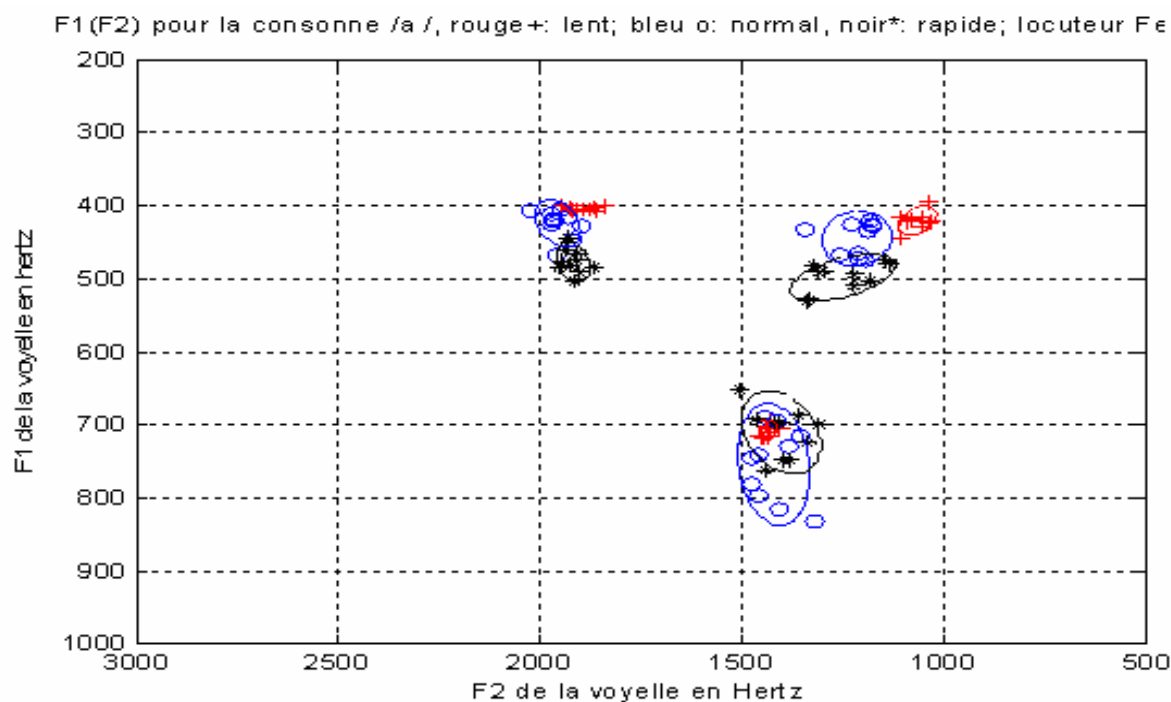
3.4.1.1 Tracé des ellipses de dispersion

Nous avons dans un premier temps tracé les ellipses de dispersion dans le plan F1, F2 pour les voyelles /a/, /i/, /u/, prélevées dans les unités de parole considérées, selon le contexte consonantique /a_/, /h/, /h_/, ce qui donne : /a_al/, /a_il/, /a_ul/ ; /hal/, /hil/, /hul/ ; /h_al/, /h_il/, /h_ul/, pour les trois débits d'élocution, pour chaque locuteur.

Les résultats obtenus sont illustrés par les figures 3.4. Pour chaque voyelle, nous avons trois ellipses correspondant aux trois débits : en bleu : débit normal ; en rouge : débit lent et en noir : débit rapide.

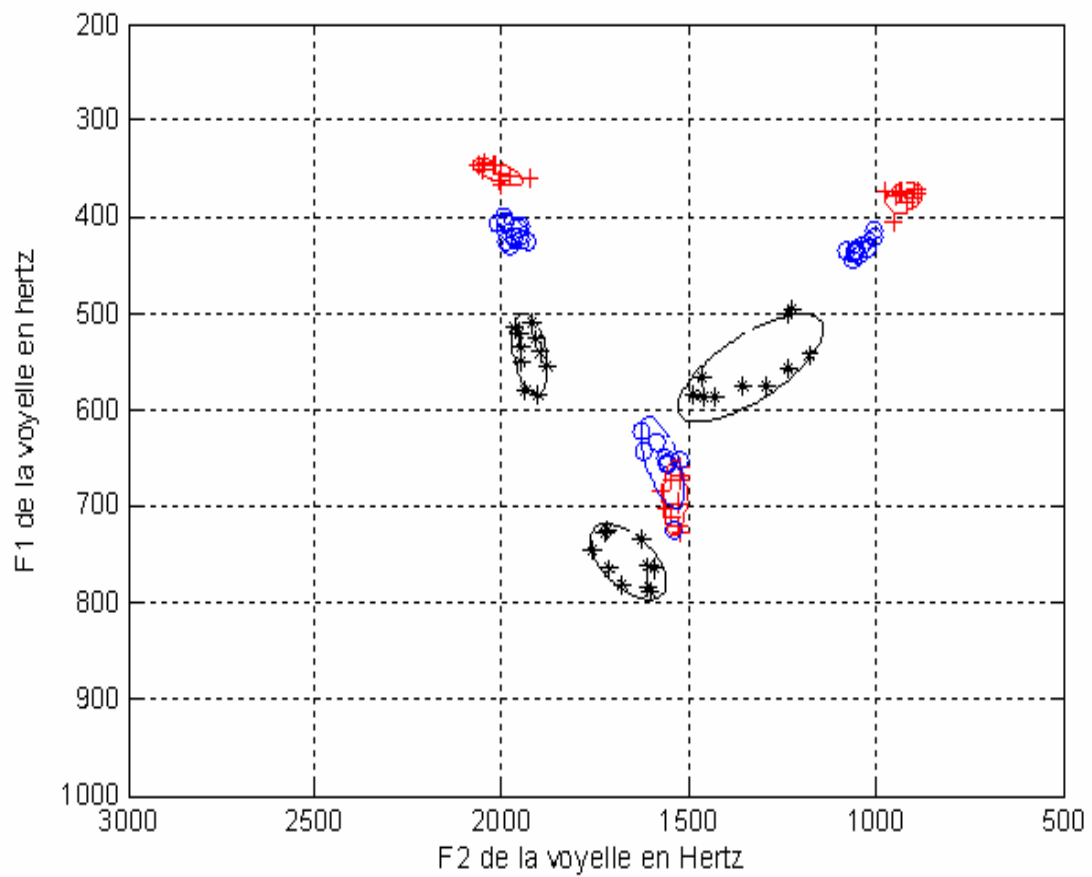
Contexte consonatique /a_/

Locuteurs Fe et MA (algerois)

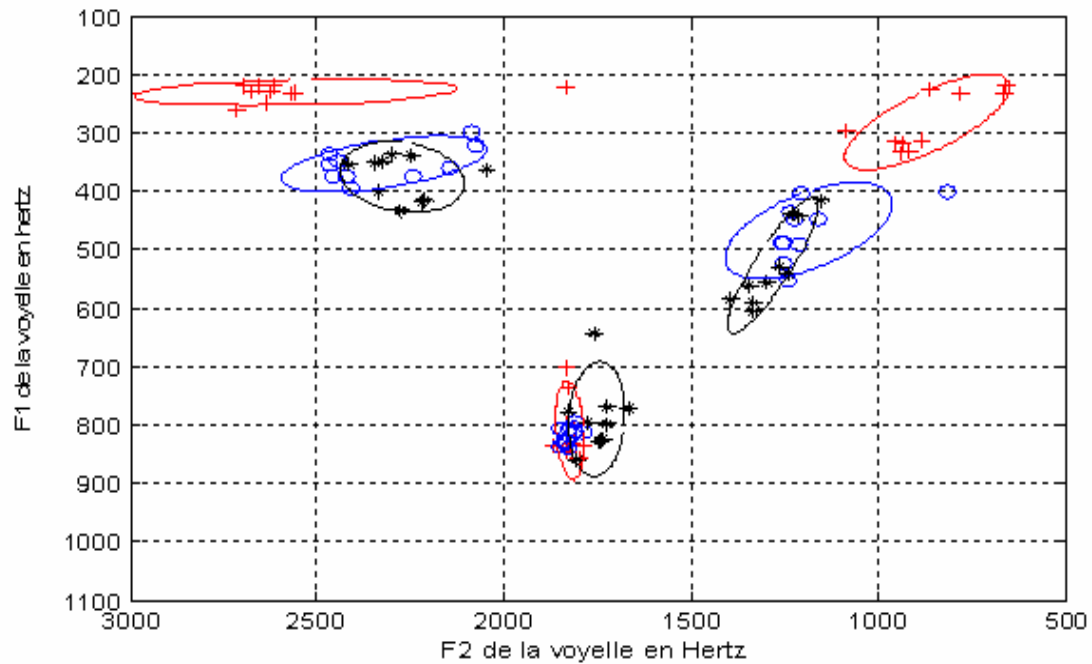


Locuteurs SA et NA (kabyles)

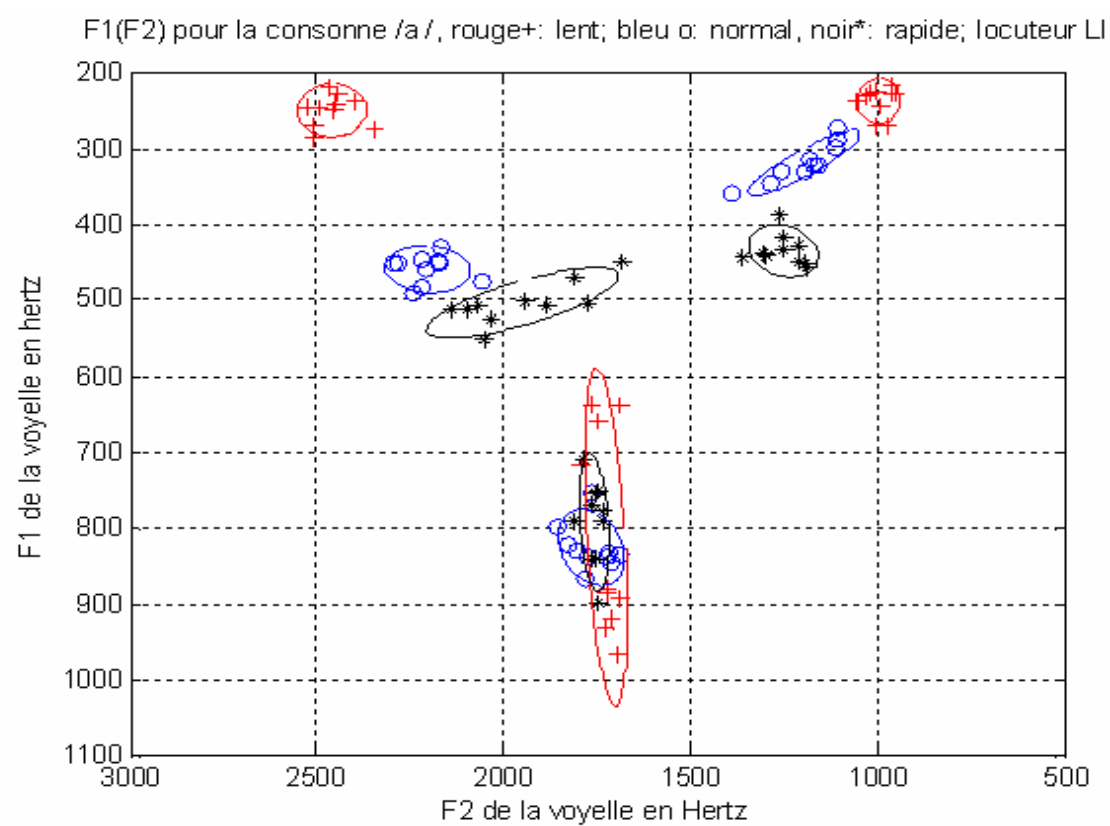
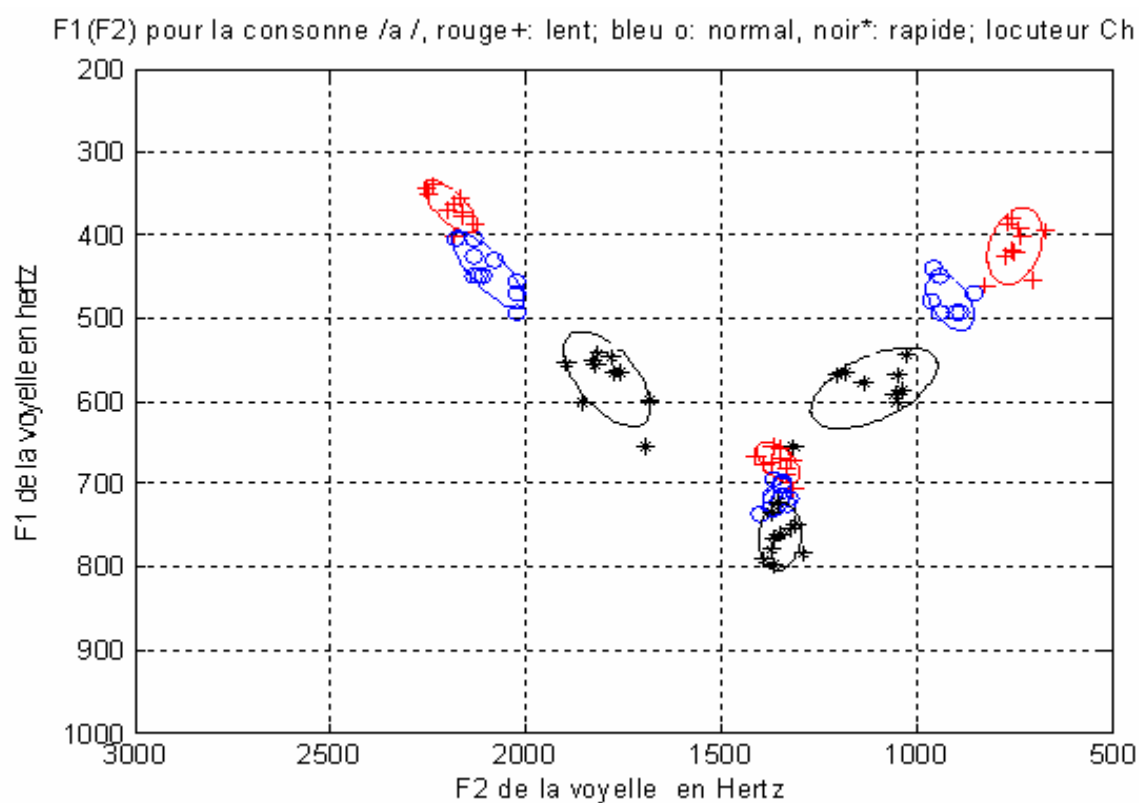
F1(F2) pour la consonne /a /, rouge+: lent; bleu o: normal, noir*: rapide; locuteur Sa



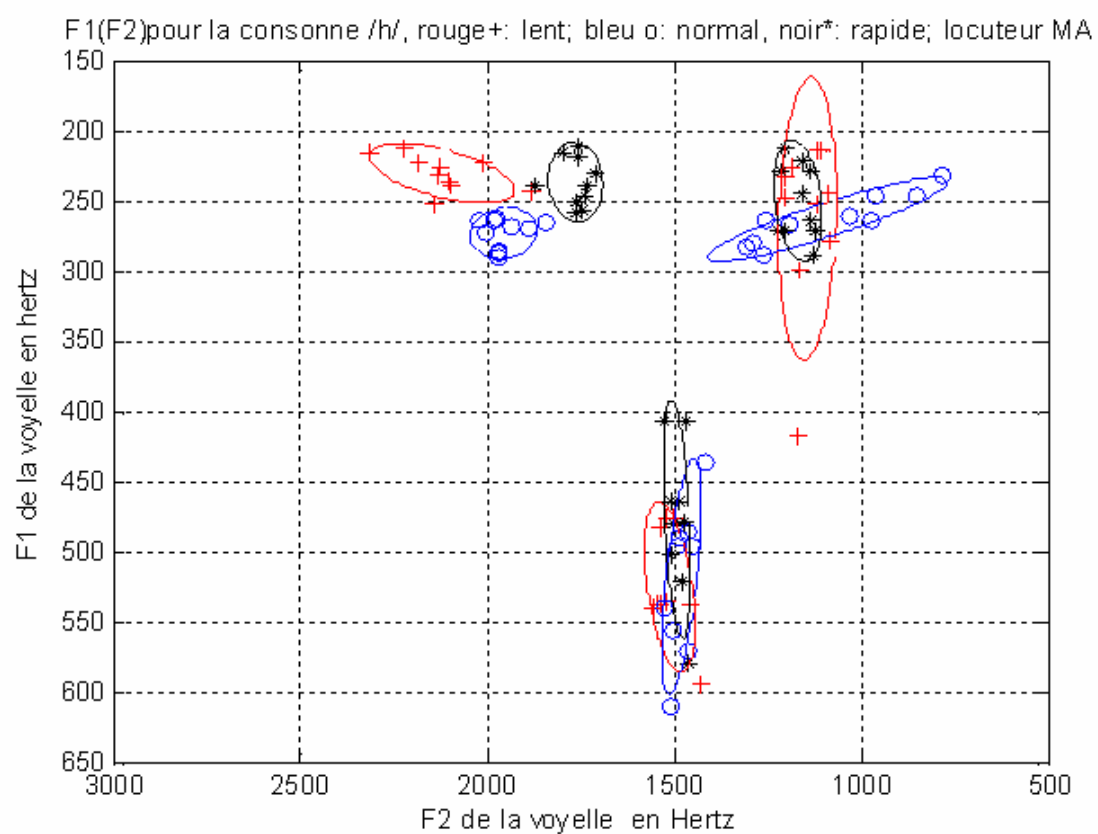
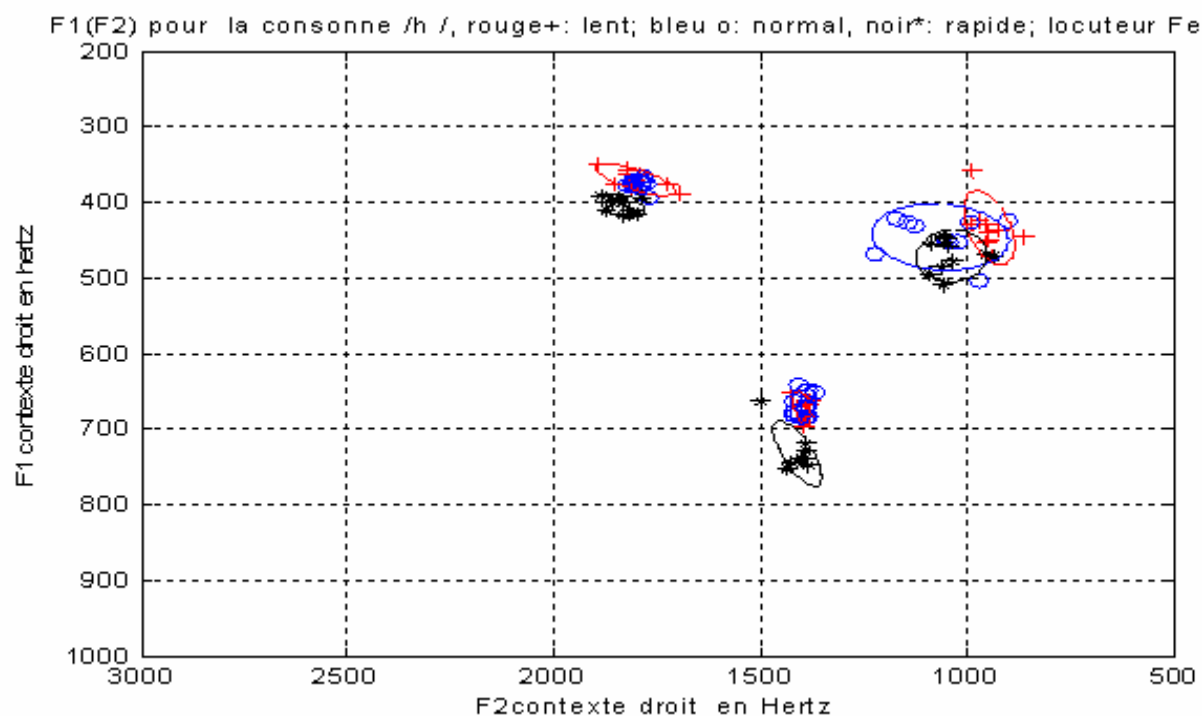
F1(F2) pour la consonne /a /, rouge+: lent; bleu o: normal, noir*: rapide; locuteur NA

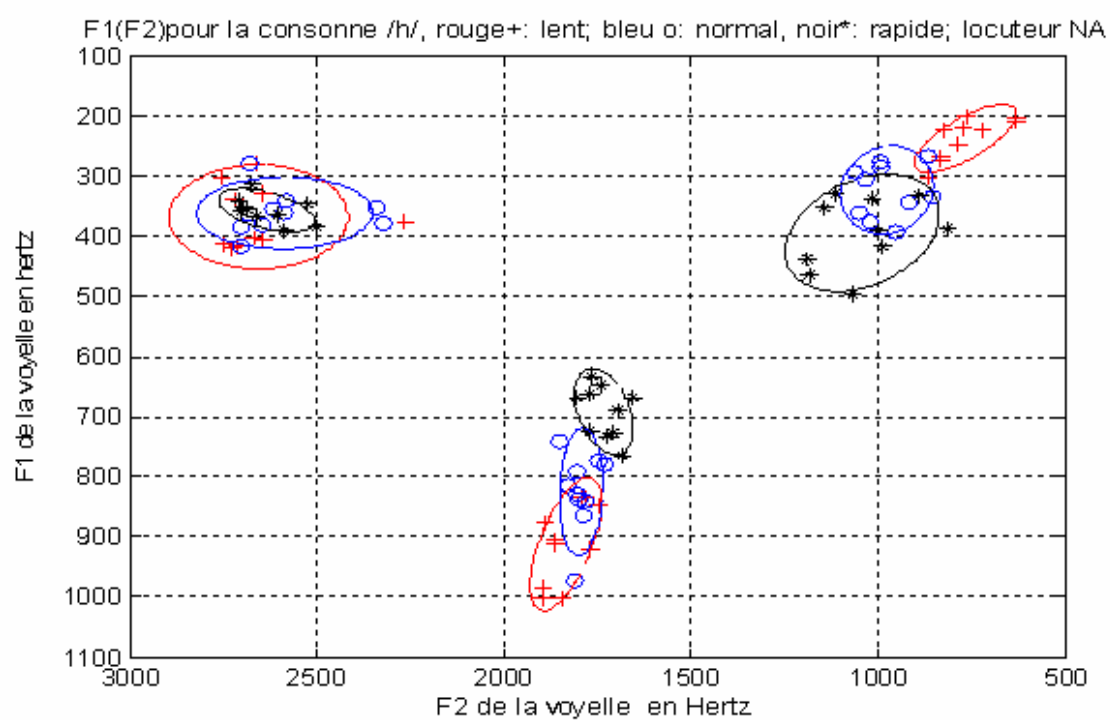
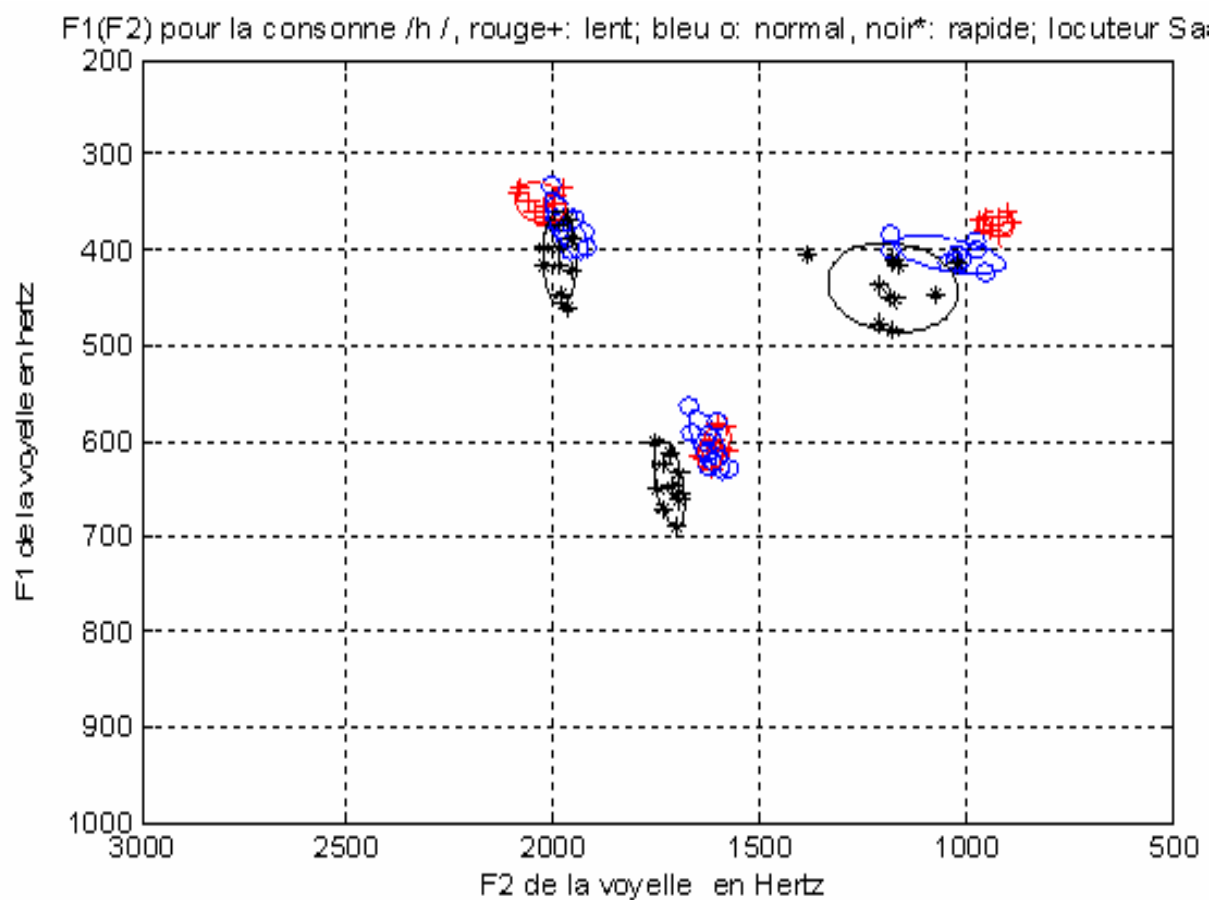


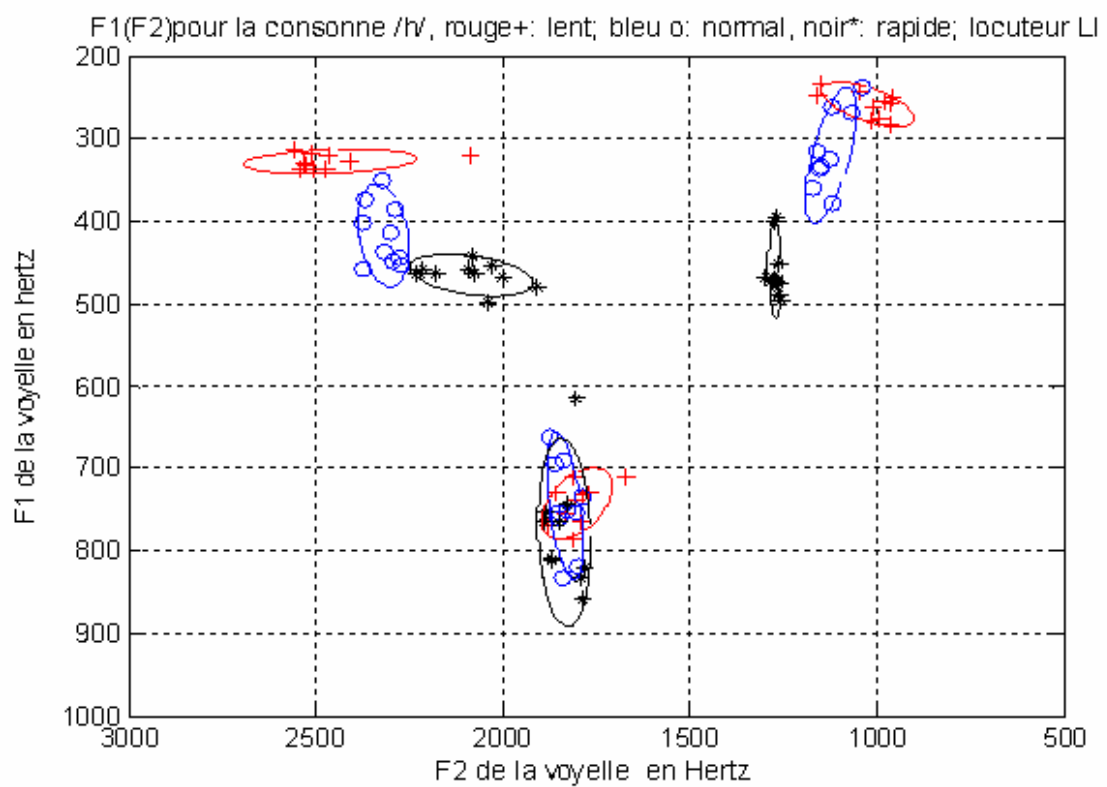
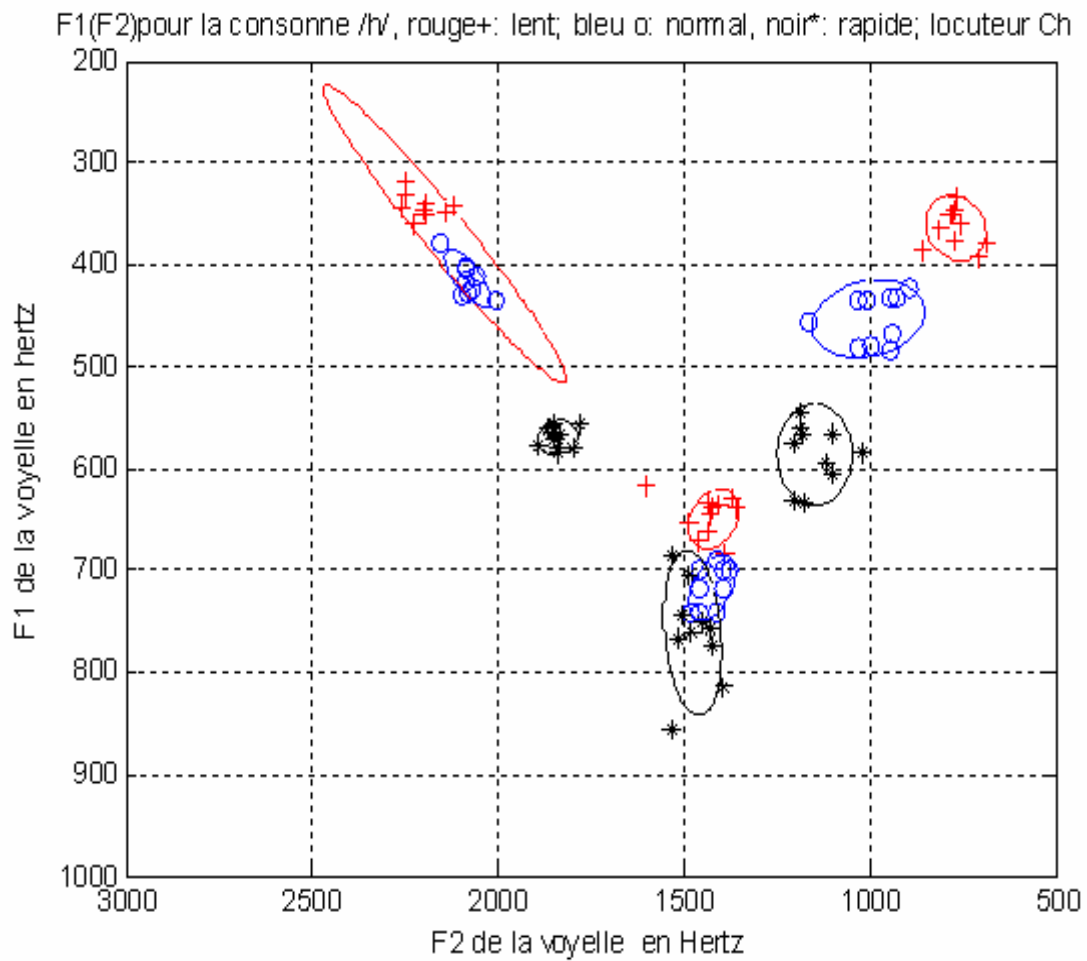
Locuteurs CH et LI (libanais)



Contexte consonantique /h/

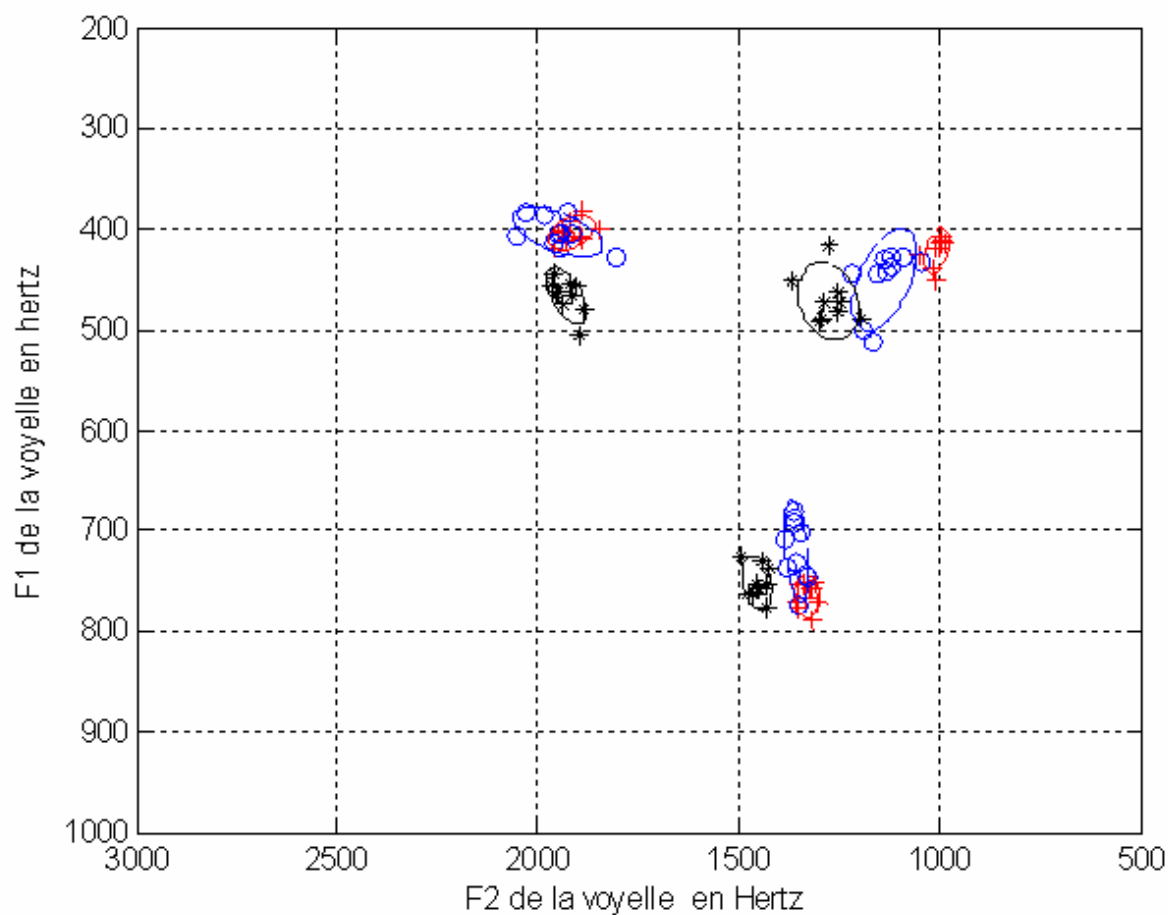




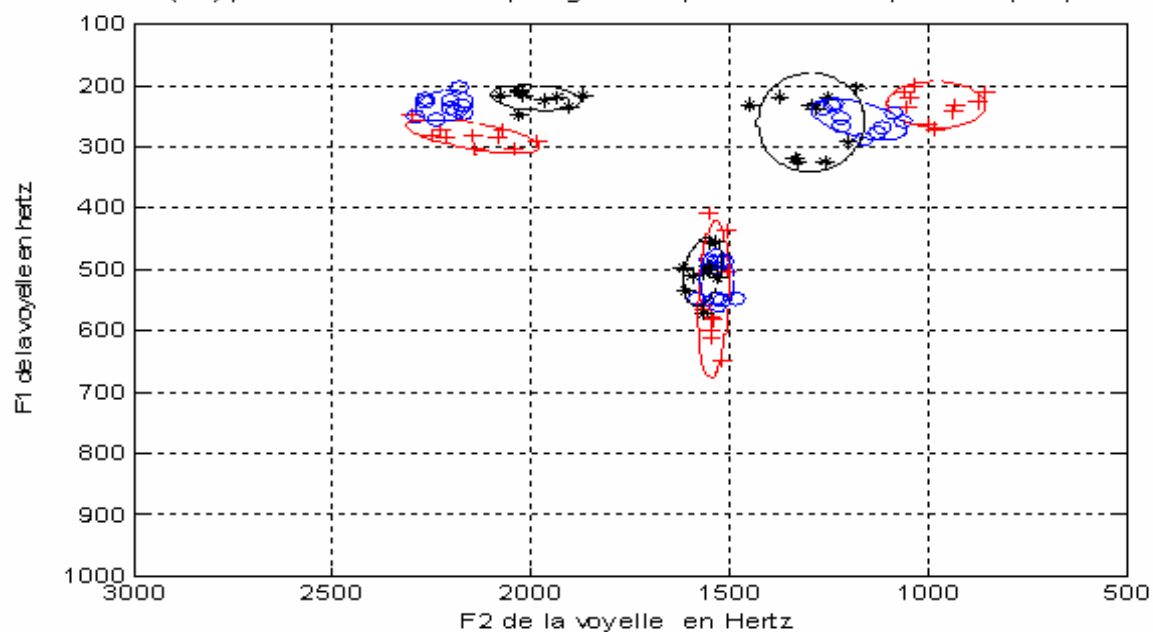


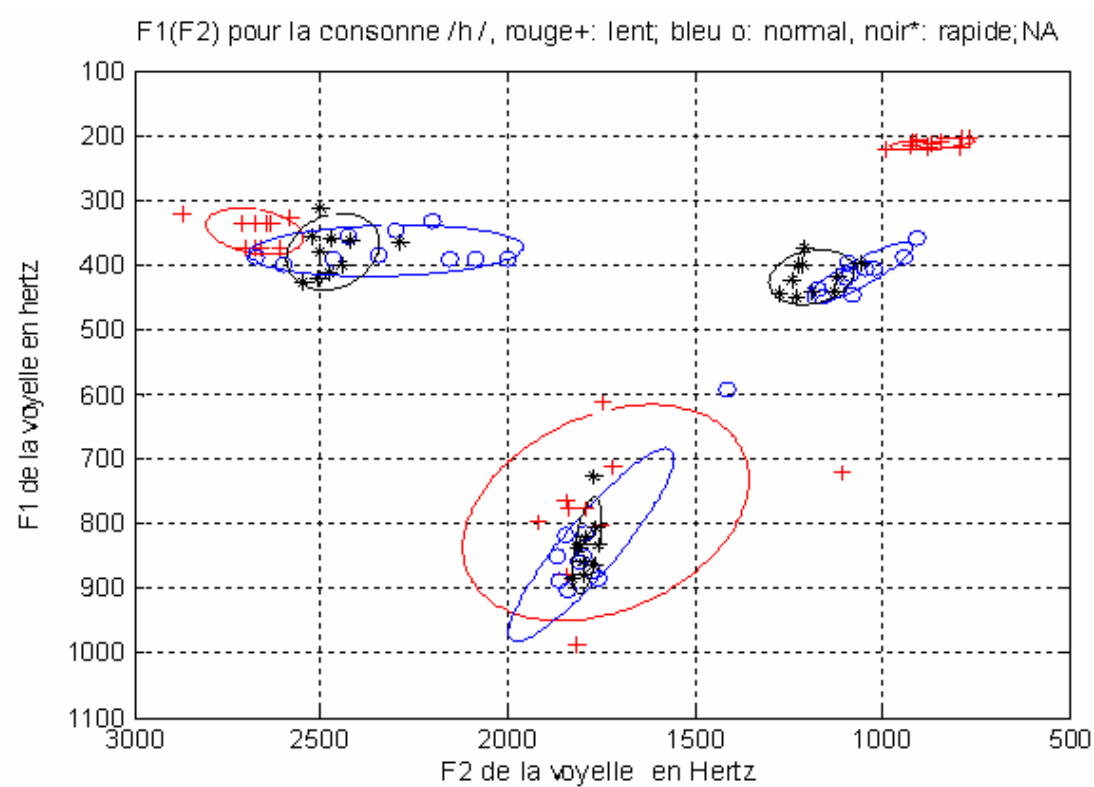
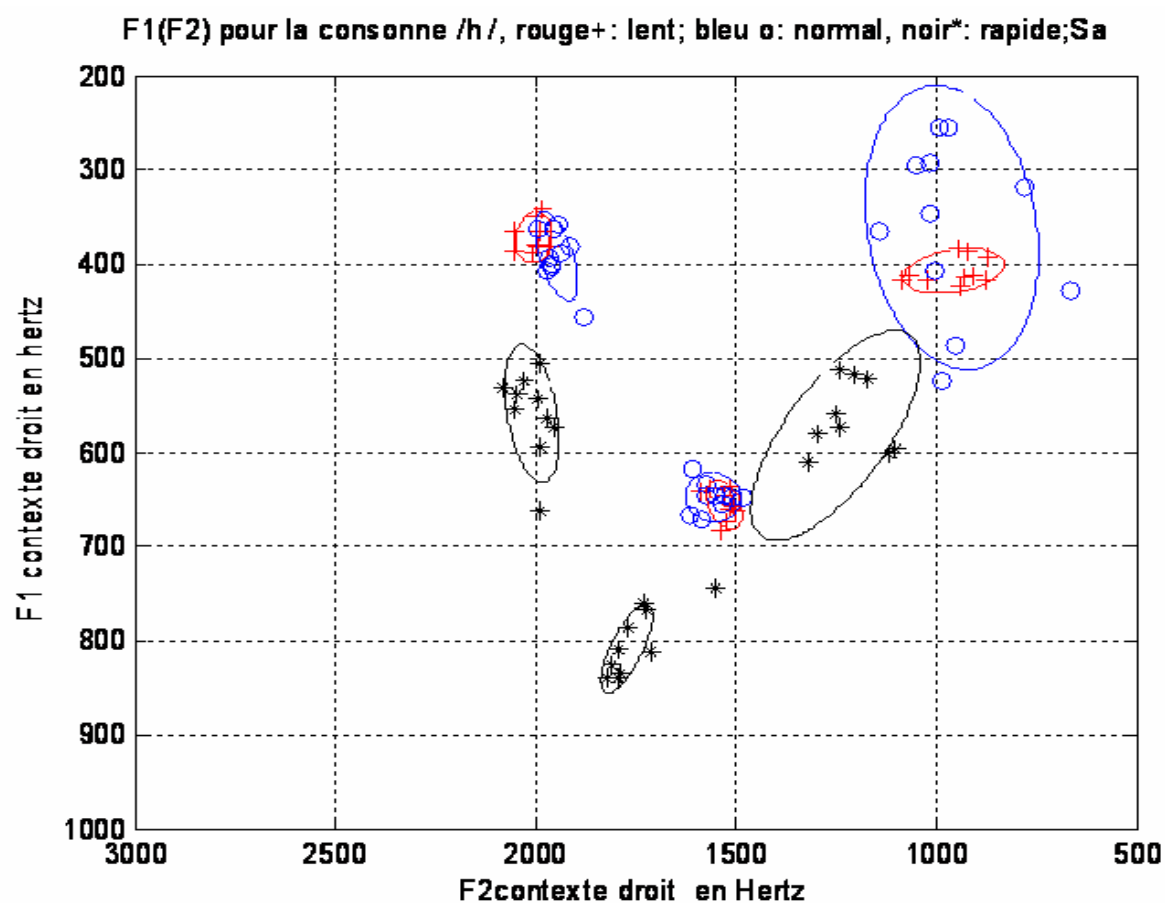
Contexte consonantique /h_/

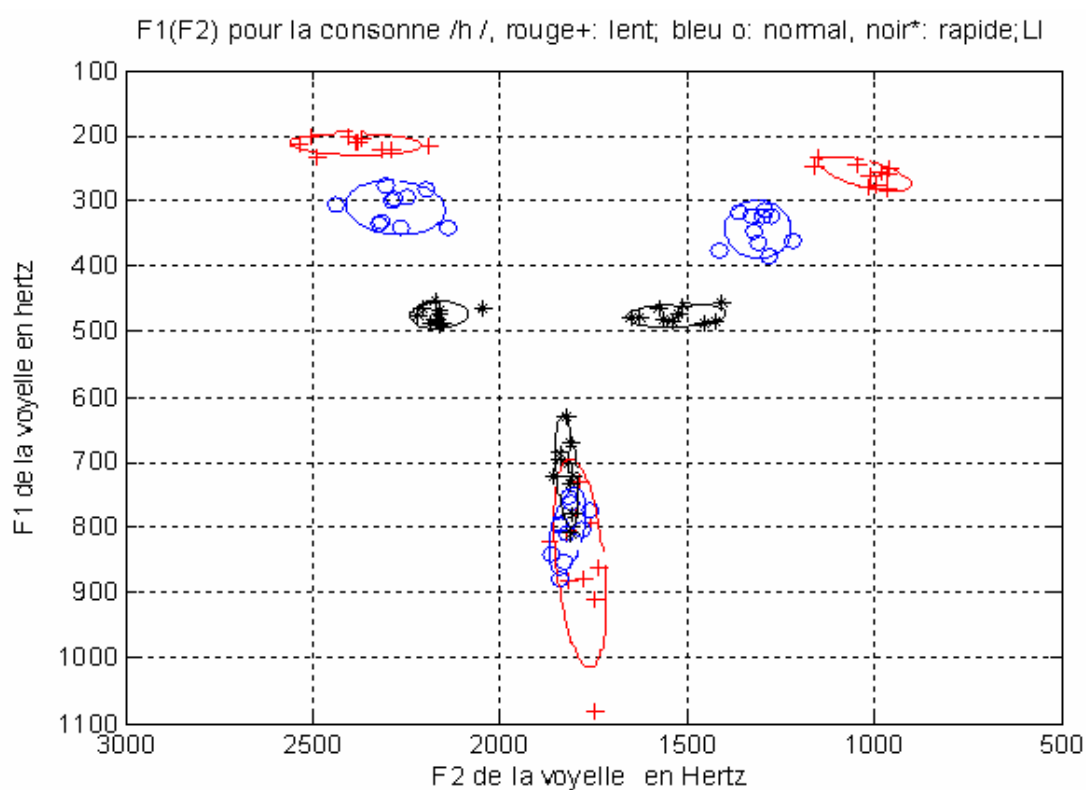
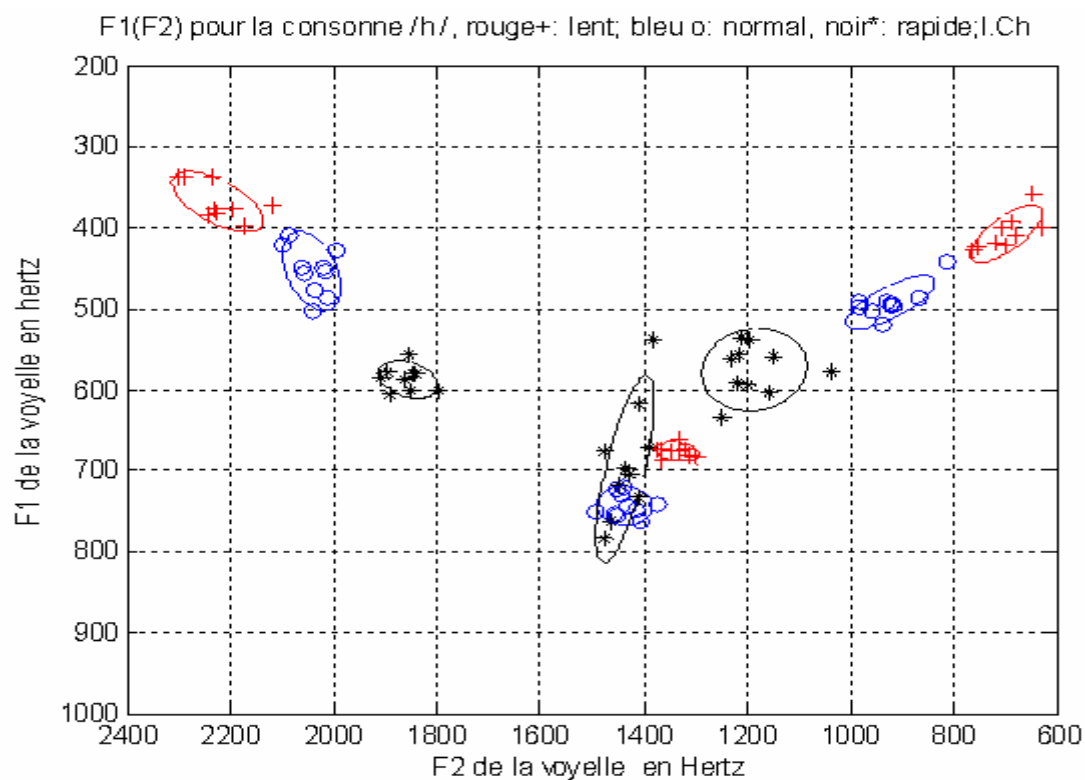
F1(F2) pour la consonne /h/, rouge+: lent; bleu o: normal, noir*: rapide; Fe



F1(F2) pour la consonne /h/, rouge+: lent; bleu o: normal, noir*: rapide; MA







Figures 3. 4. Ellipses de dispersion relatives à chaque locuteur, chaque contexte consonantique et chaque débit
En bleu : débit normal ; en rouge : débit lent et en noir : débit rapide

- En regardant les ellipses de dispersion de FE et MA (algerois) en figures 3.4, nous remarquons en effet qu'elles sont presque toutes superposées pour la même voyelle en fonction du débit, et ce quelque soit le contexte consonantique, donc son espace vocalique varie très peu avec le débit. Ce qui rejoint la théorie de Gay

- Les ellipses de CH et LI (libanais) sont par contre presque toutes disjointes et se dirigent vers le centre de l'espace F1, F2. Cela veut dire que l'espace vocalique change avec le débit. Ce qui vérifie la théorie de Gay prise en hypothèse.

- SA et NA (kabyles) par contre se trouvent entre les deux locuteurs. Il y a des ellipses disjointes et d'autres non. Donc pour le moment nous ne pouvons pas lui attribuer une stratégie.

Pour arriver à des résultats plus quantitatifs donc plus concrets, il serait intéressant de comparer les surfaces des triangles vocaliques en fonction du débit afin de voir s'il y a réduction de l'espace vocalique avec l'augmentation du débit.

3.4.1.2. Tracé des triangles vocaliques

A partir des ellipses de dispersion précédentes, nous avons tracé pour chaque contexte consonantique et pour chaque débit, les droites joignant les centres des ellipses de dispersion (correspondants aux moyennes des formants F1 et F2). Nous avons obtenu ainsi trois triangles vocaliques correspondant à chaque débit d'élocution. Les résultats obtenus sont illustrés par les figures 3.5:

3.4.1.3. Calcul de la surface des triangles vocaliques

La méthode utilisée est inspirée de l'article de Fabrice Hirsch [11], où l'aire des triangles vocalique est calculée de la façon suivante :

L'aire des triangles vocaliques est obtenue par la relation suivante :

$$\text{Aire} = \sqrt{P (P-a) (P-b) (P-c)} \text{ en kHz}^2$$

où a, b et c représentent la distance entre les coordonnées de deux voyelles qui est quantifiée par :

$$\sqrt{(x_b - x_a)^2 + (y_b - y_a)^2} \quad (1)$$

(x correspond à F1 d'une des trois voyelles et y à F2) et où P est le résultat de :

$$(a + b + c) / 2 \quad (2)$$

Le coté d'un triangle est calculé par

le coté « a » du triangle

$$a = \sqrt{(x_b - x_a)^2 + (y_b - y_a)^2} \quad (3)$$

le coté « b » du triangle:

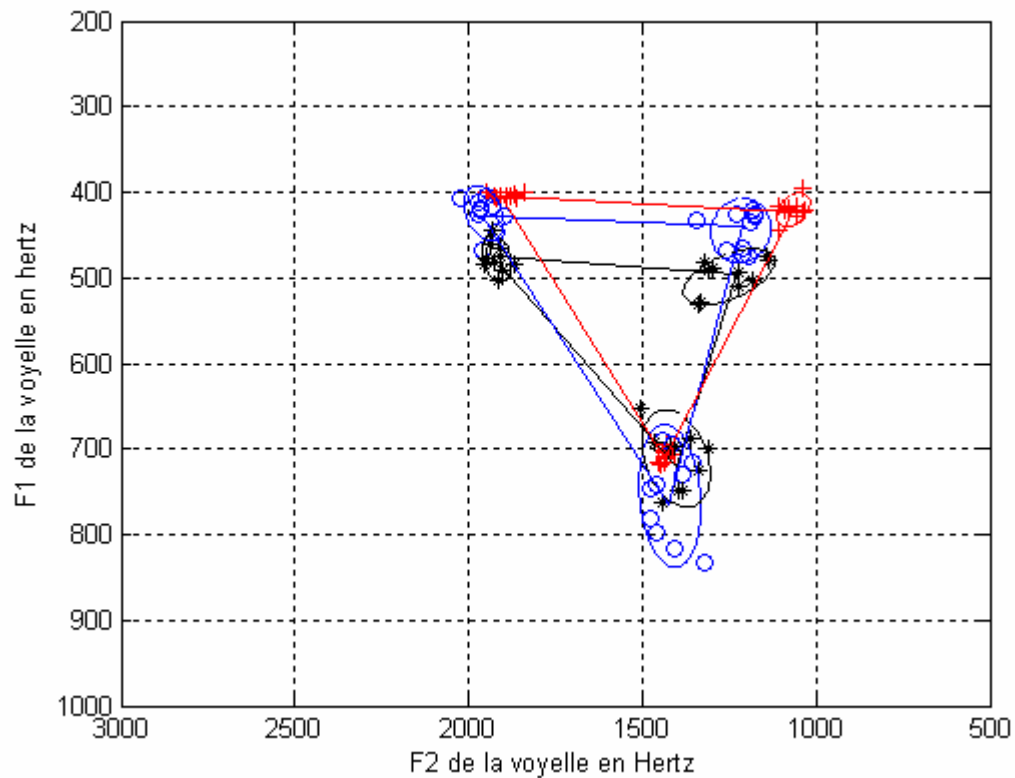
$$b = \sqrt{(x_c - x_b)^2 + (y_c - y_b)^2} \quad (4)$$

le coté « c » du triangle.

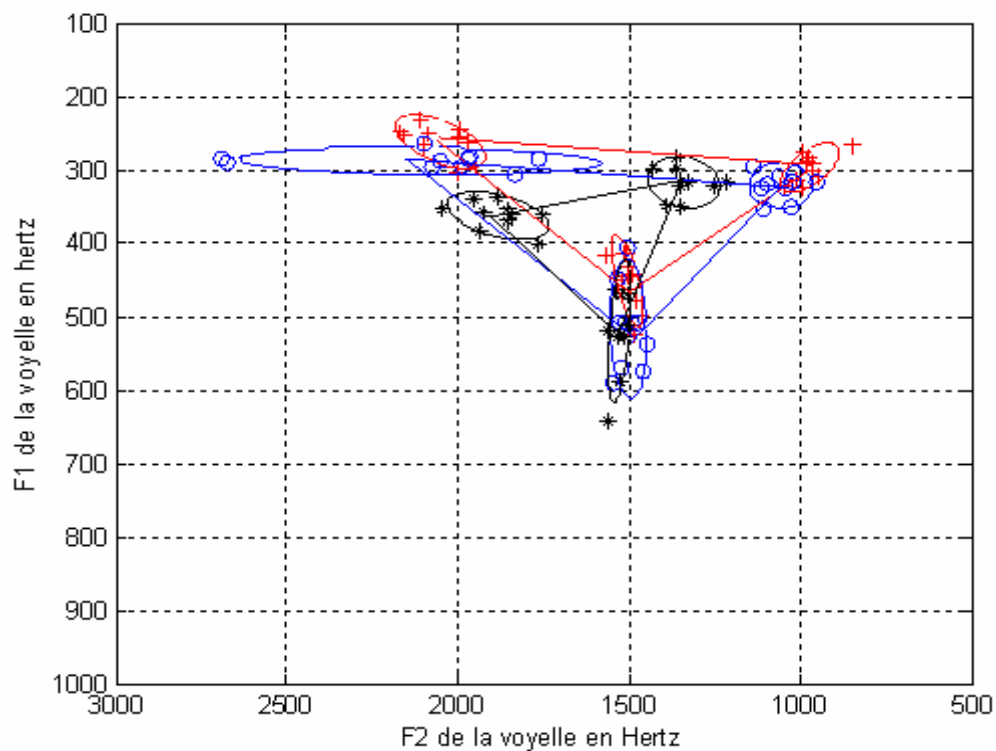
$$c = \sqrt{(x_a - x_c)^2 + (y_a - y_c)^2} \quad (5)$$

Pour la pharyngale /a_/

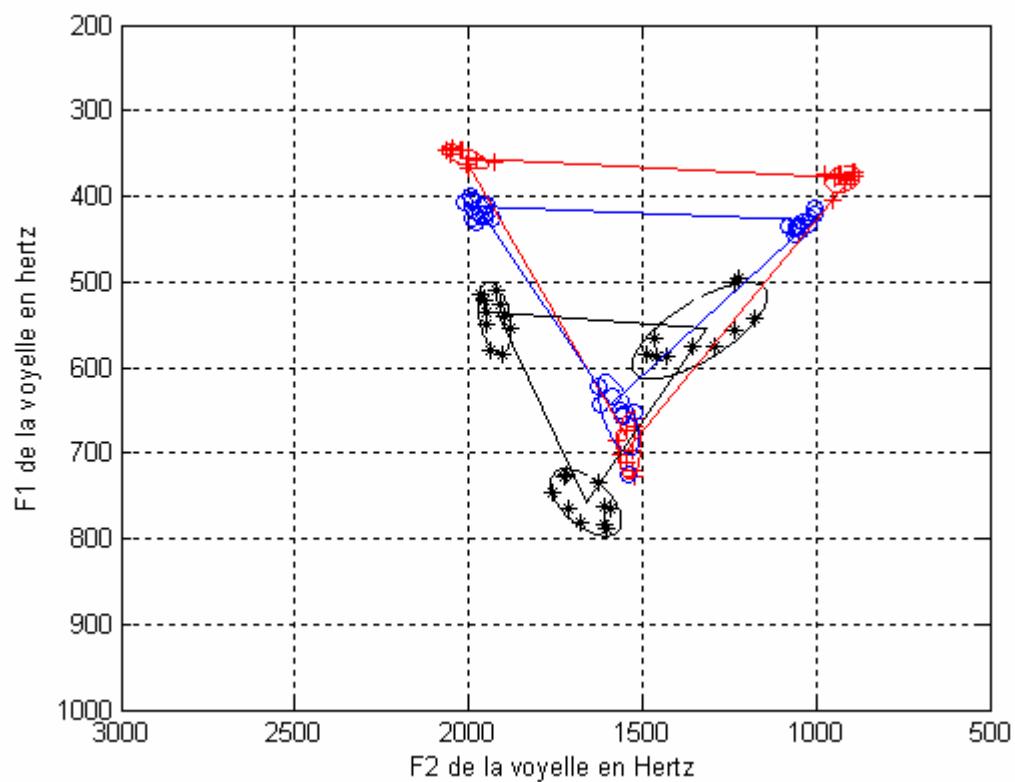
F1(F2) pour la consonne /a /, rouge+: lent; bleu o: normal, noir*: rapide; locuteur Fe



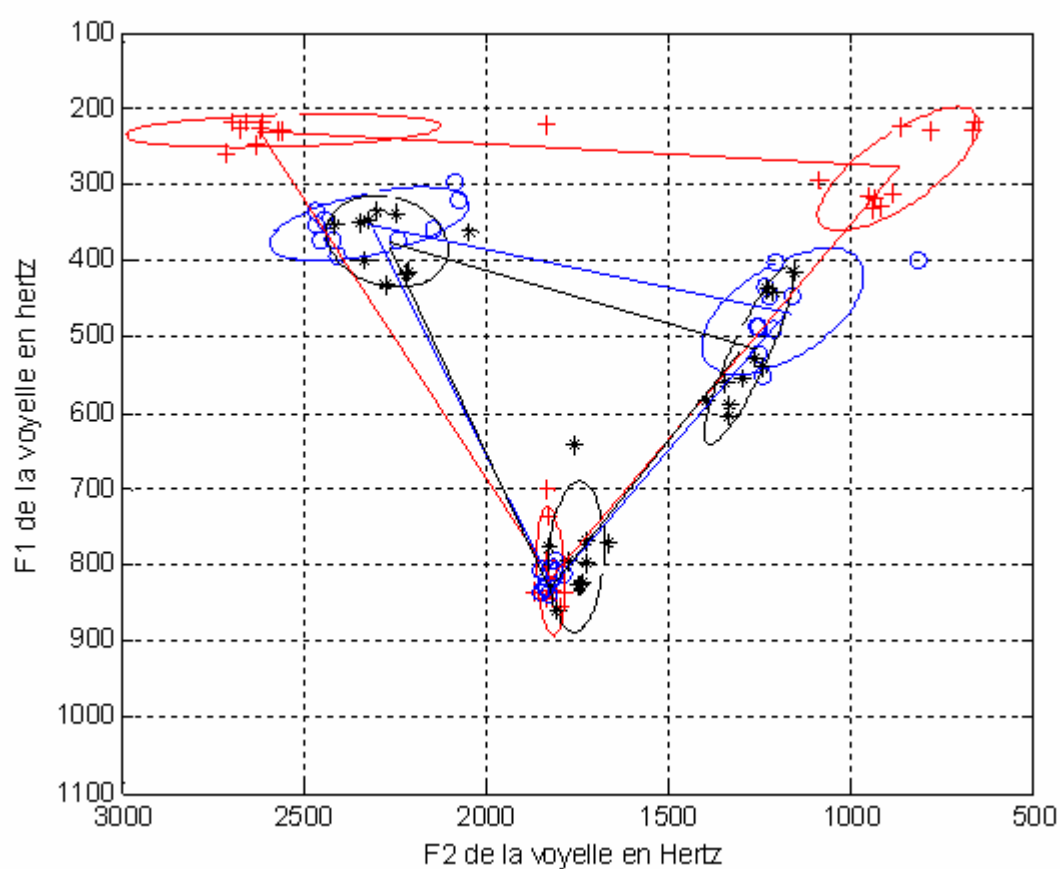
F1(F2) pour la consonne /a /, rouge+: lent; bleu o: normal, noir*: rapide; locuteur MA



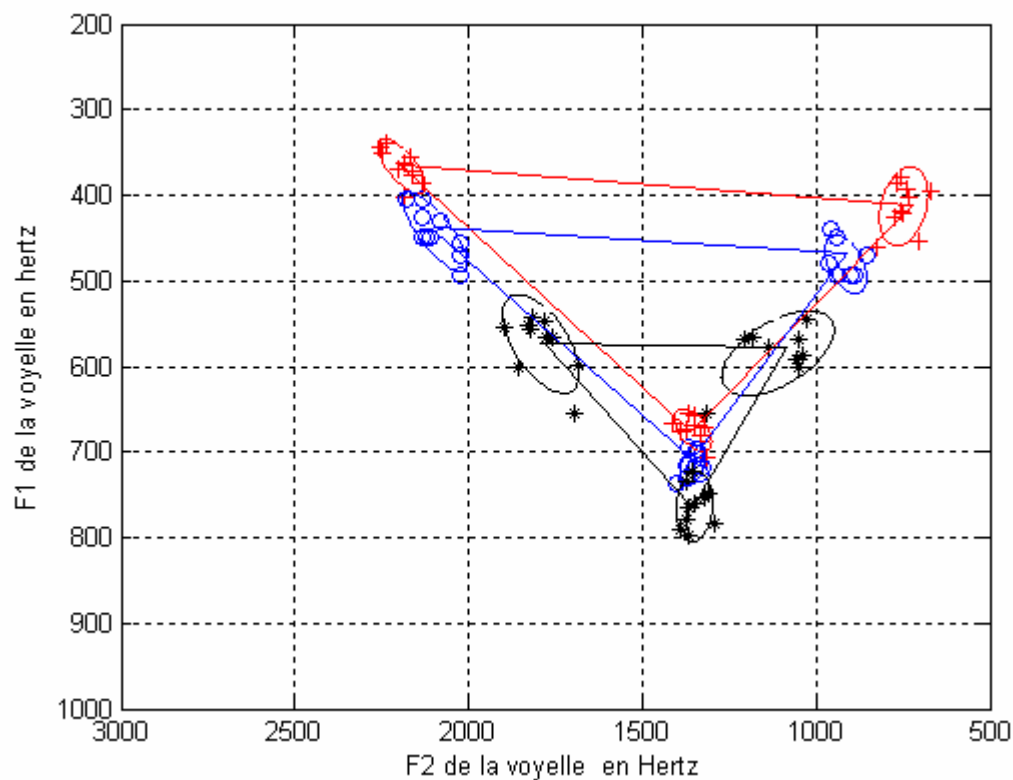
F1(F2) pour la consonne /a/, rouge+: lent; bleu o: normal, noir*: rapide; locuteur Sa



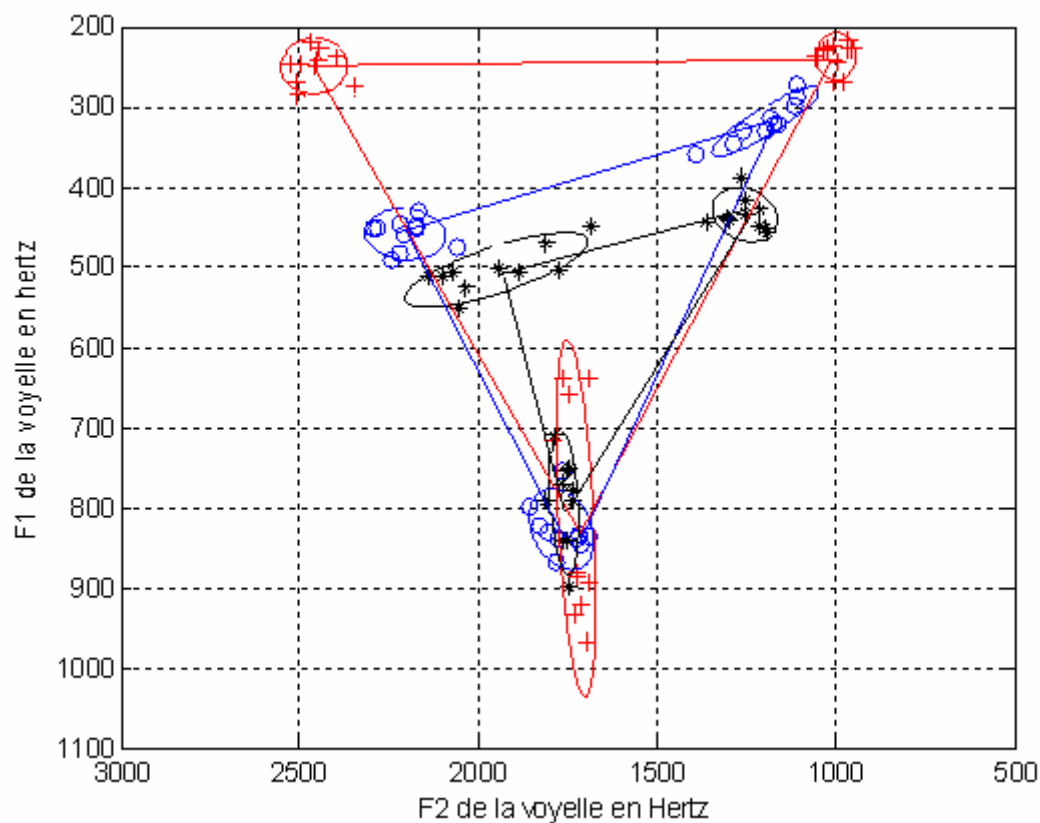
F1(F2) pour la consonne /a/, rouge+: lent; bleu o: normal, noir*: rapide; locuteur NA



F1(F2) pour la consonne /a /, rouge+: lent; bleu o: normal, noir*: rapide; locuteur Ch

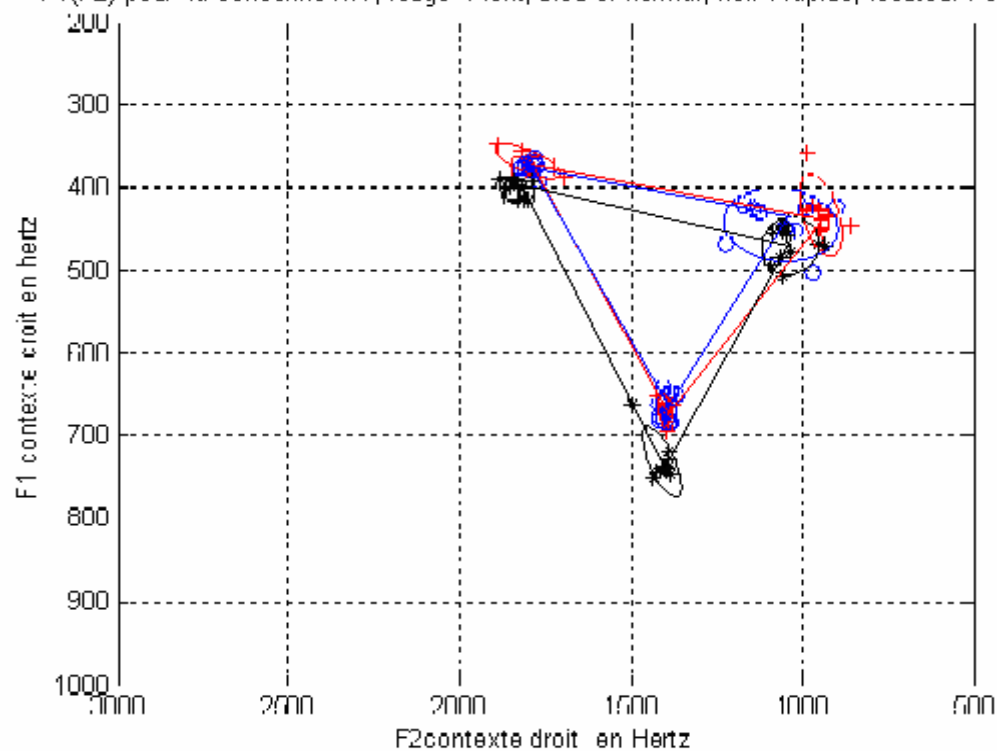


F1(F2) pour la consonne /a /, rouge+: lent; bleu o: normal, noir*: rapide; locuteur LI

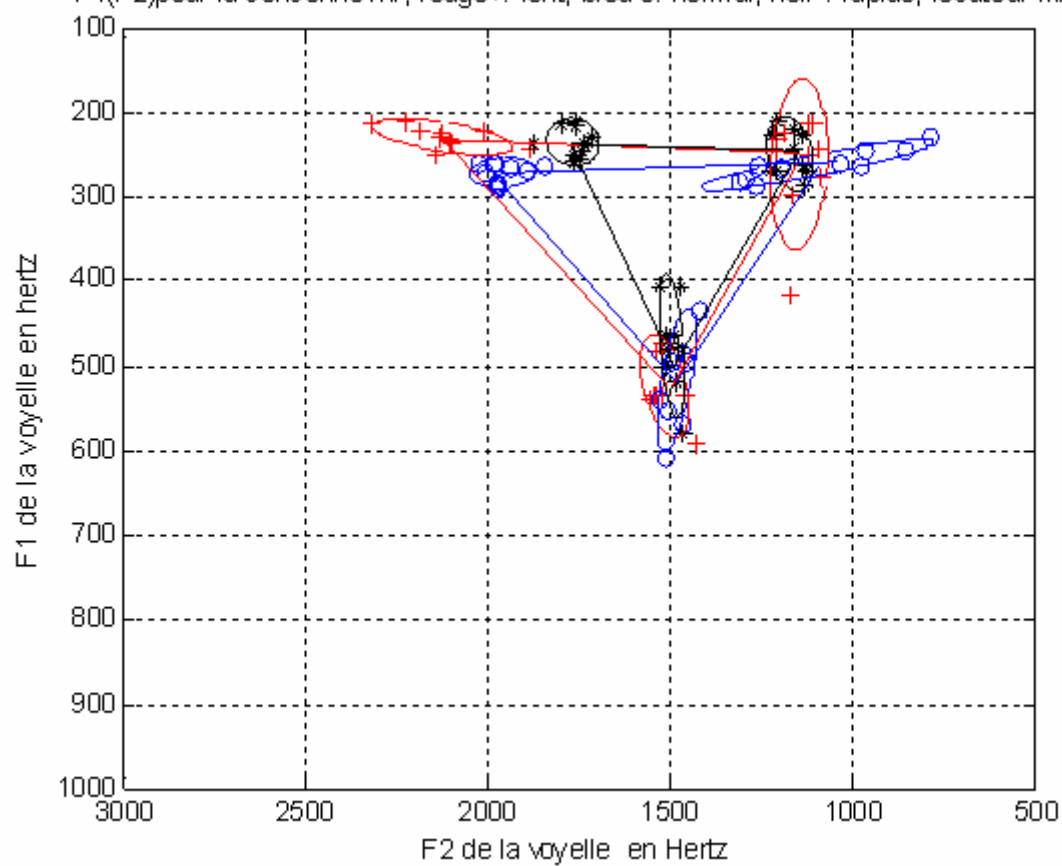


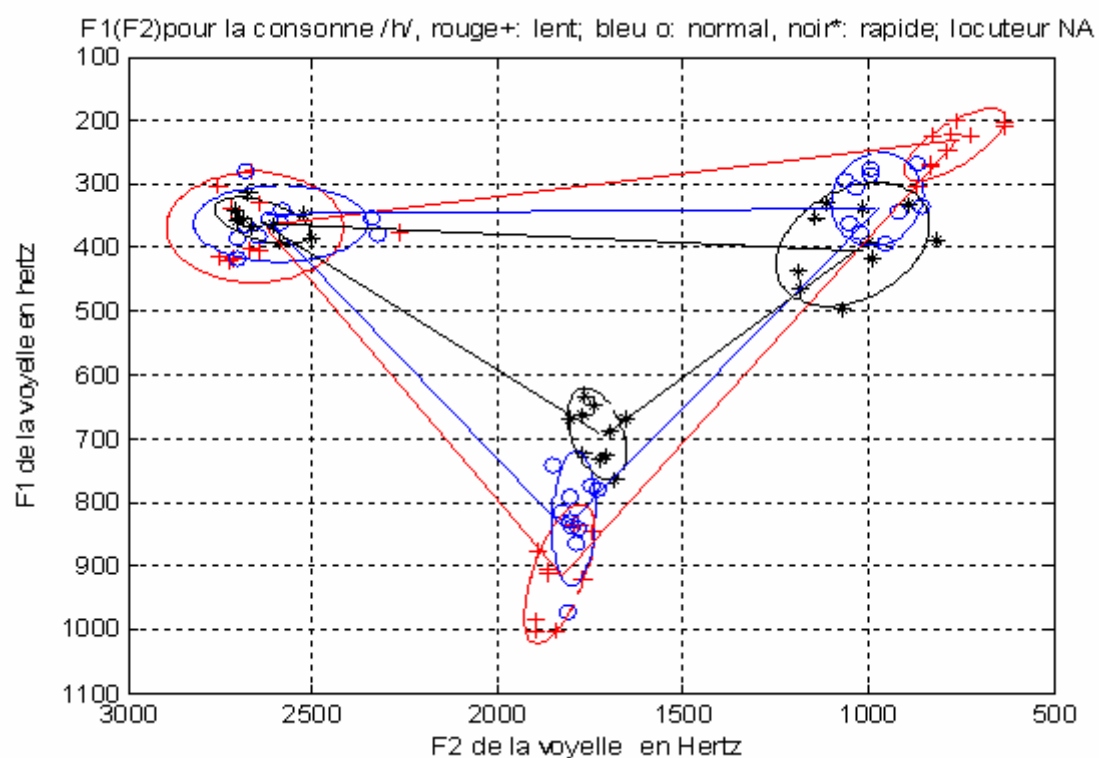
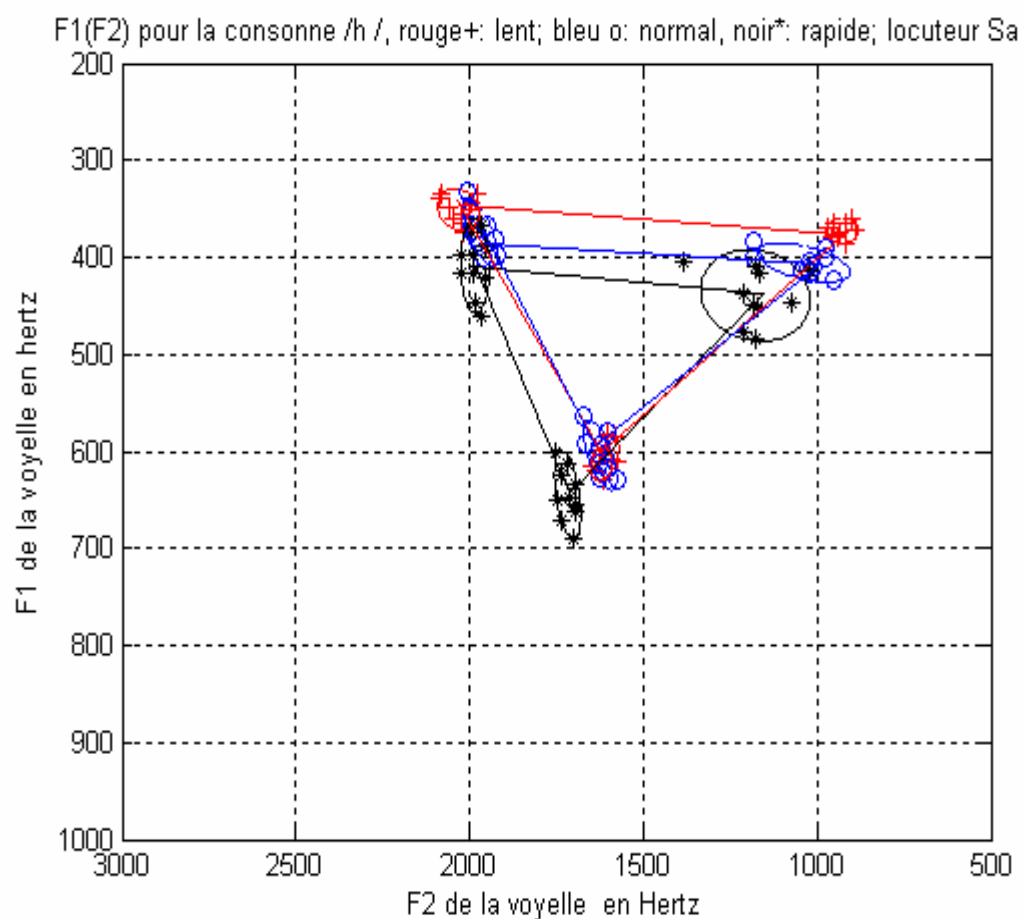
Pour la pharyngale /h/

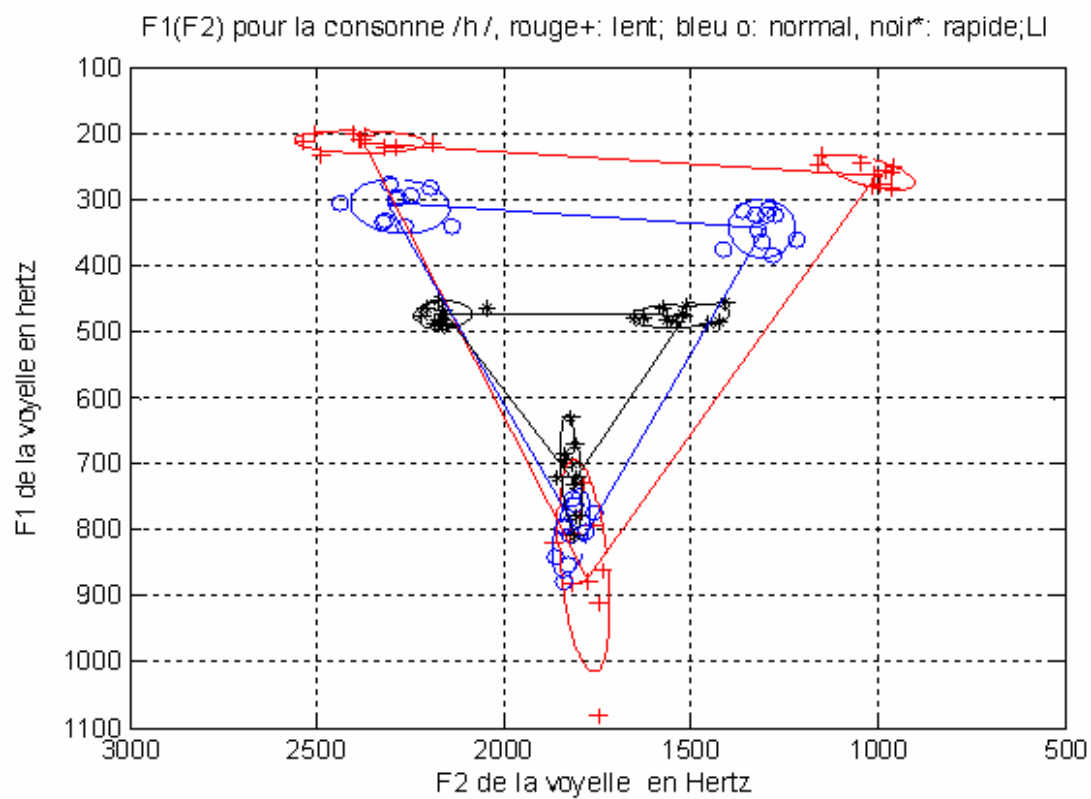
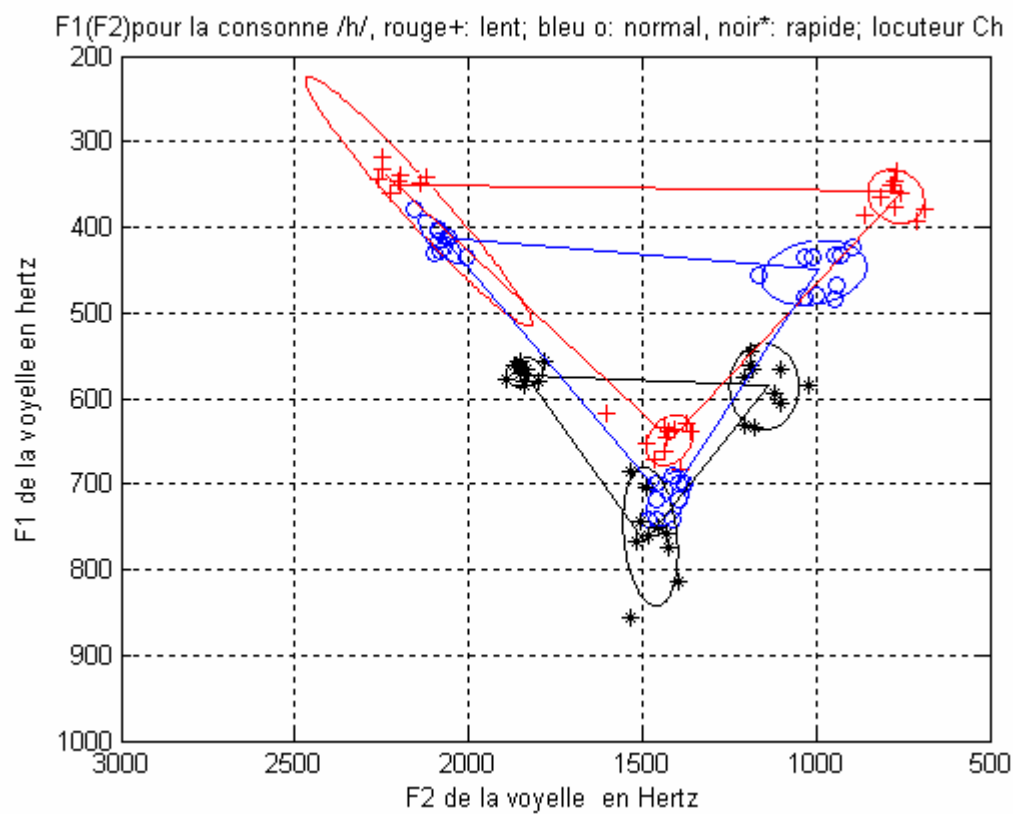
F1(F2) pour la consonne /h/, rouge+: lent; bleu o: normal, noir*: rapide; locuteur Fo



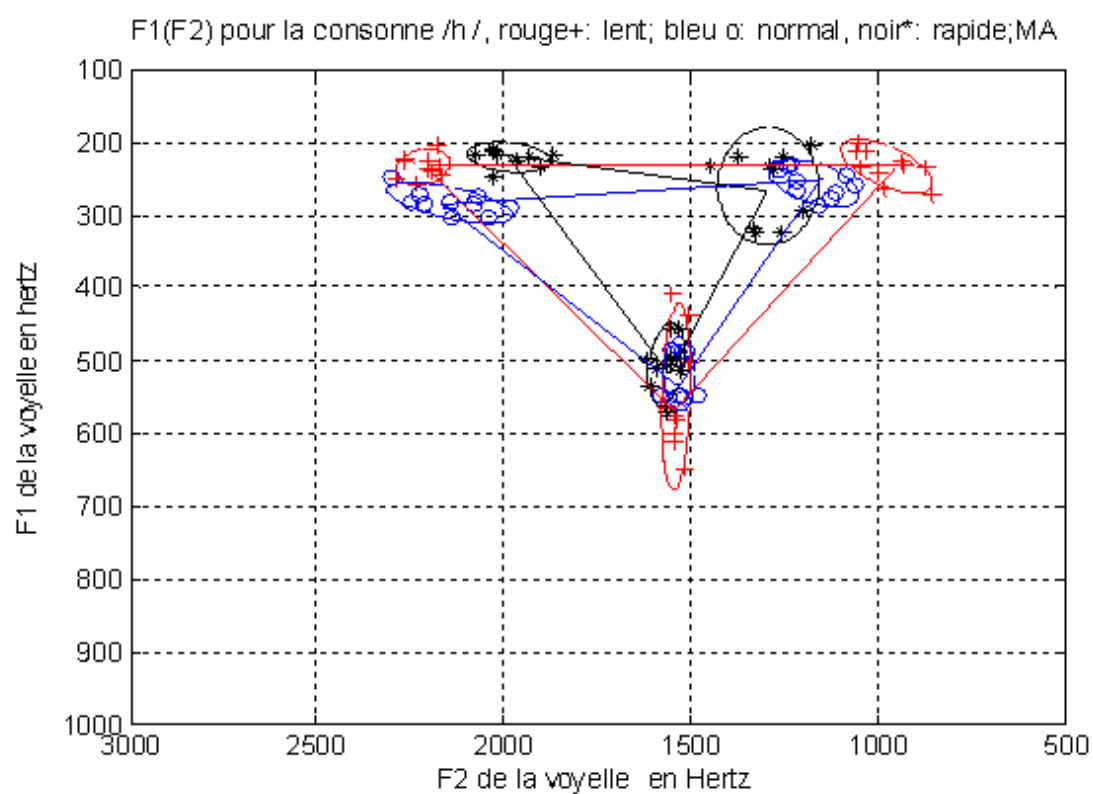
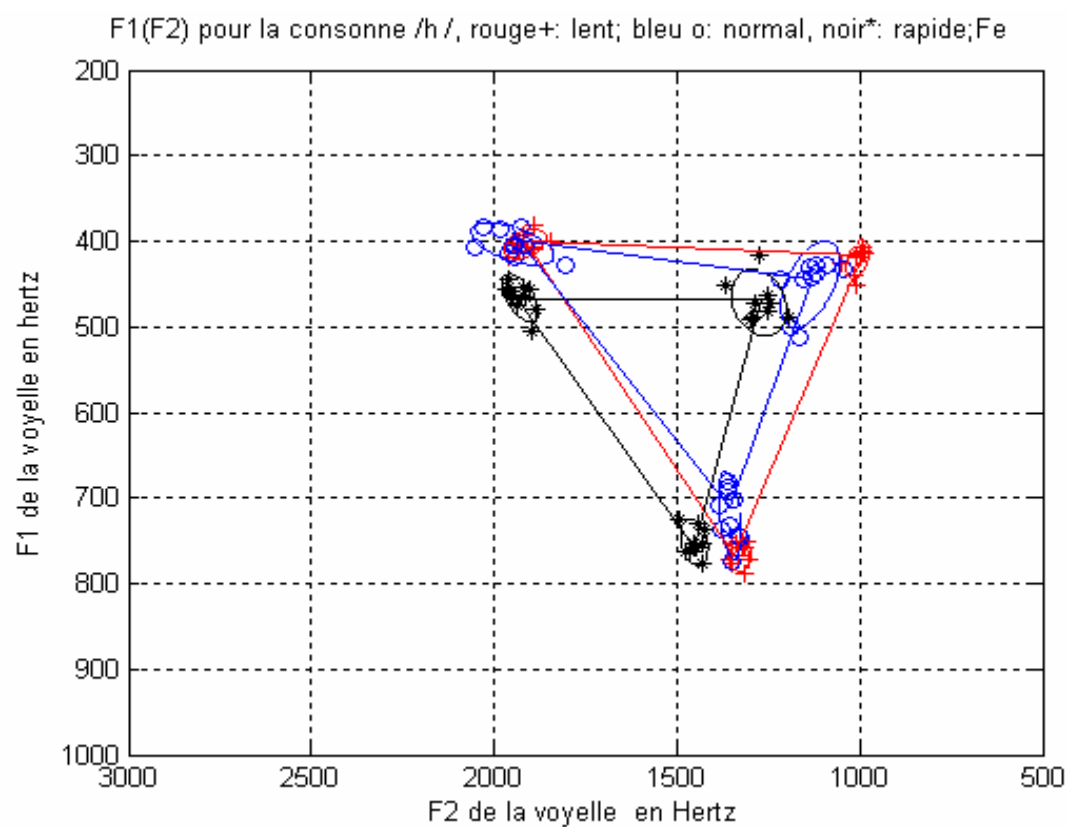
F1(F2) pour la consonne /h/, rouge+: lent; bleu o: normal, noir*: rapide; locuteur MA



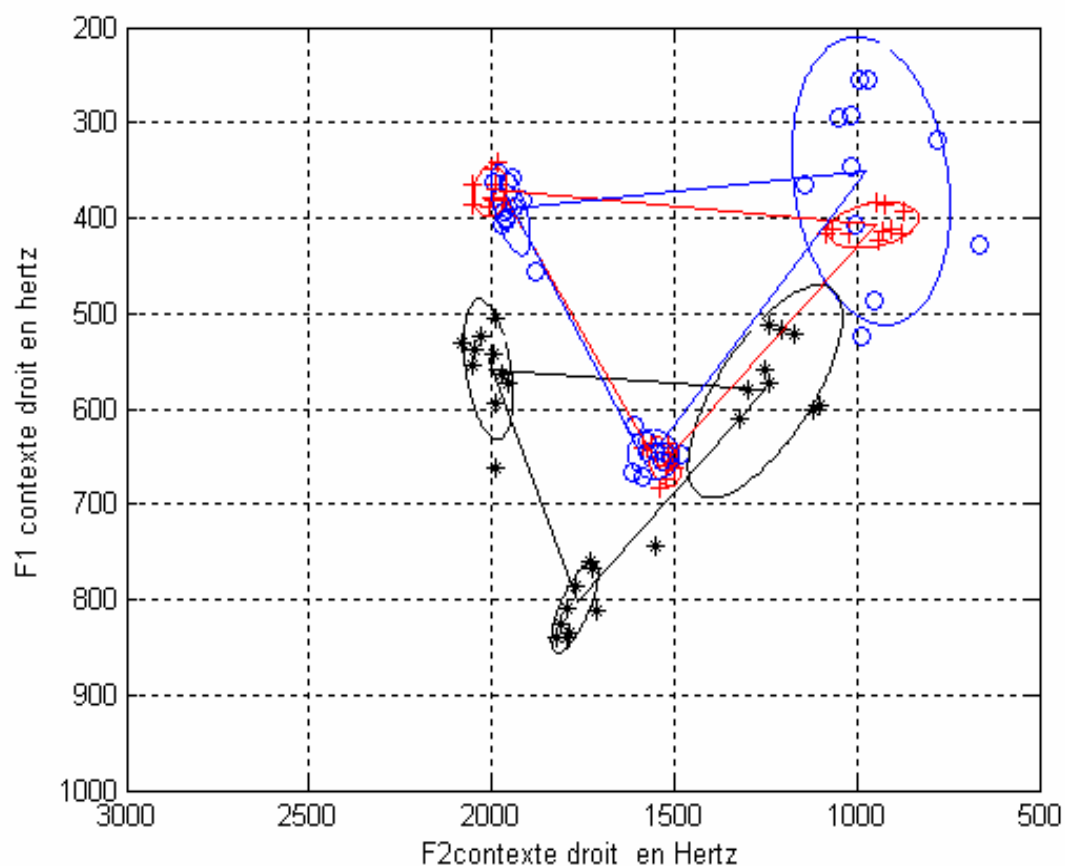




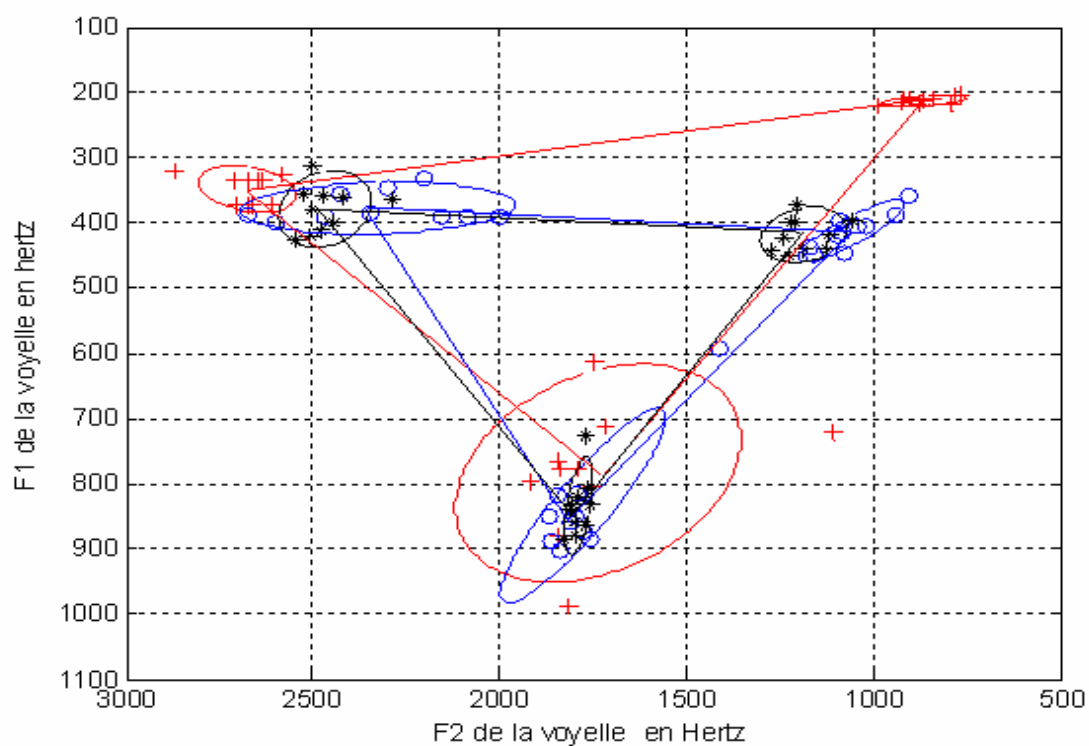
Pour la pharyngale /h_/

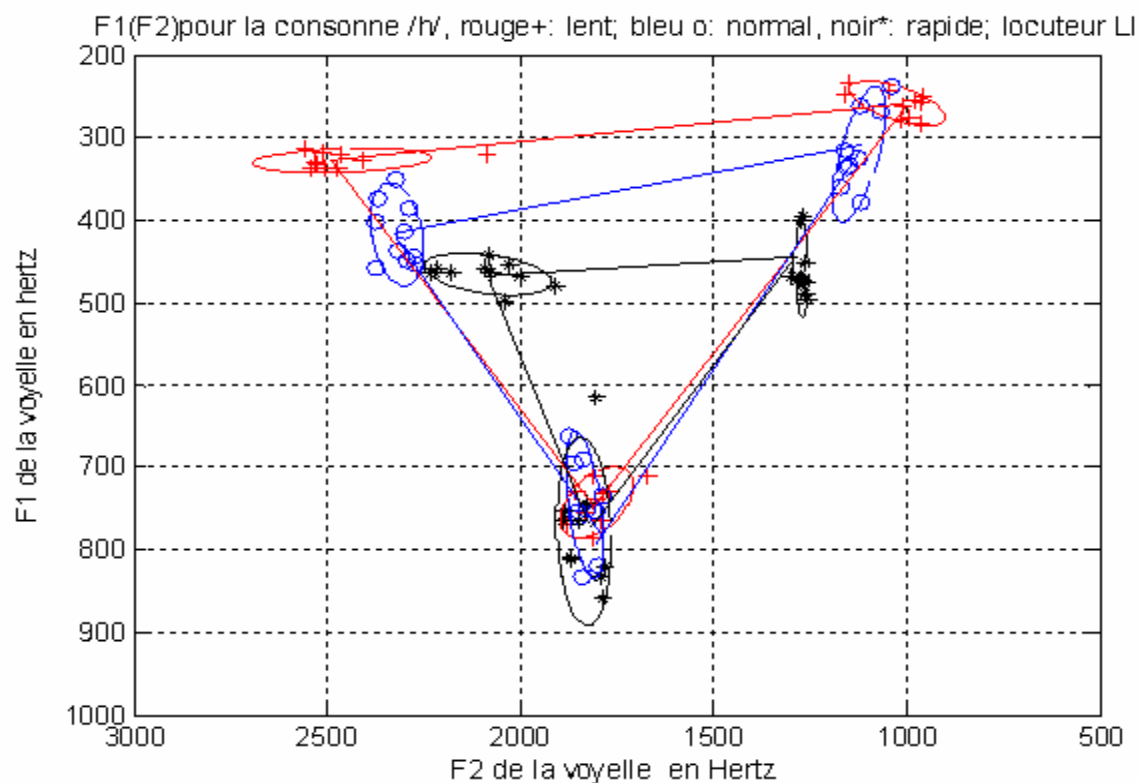
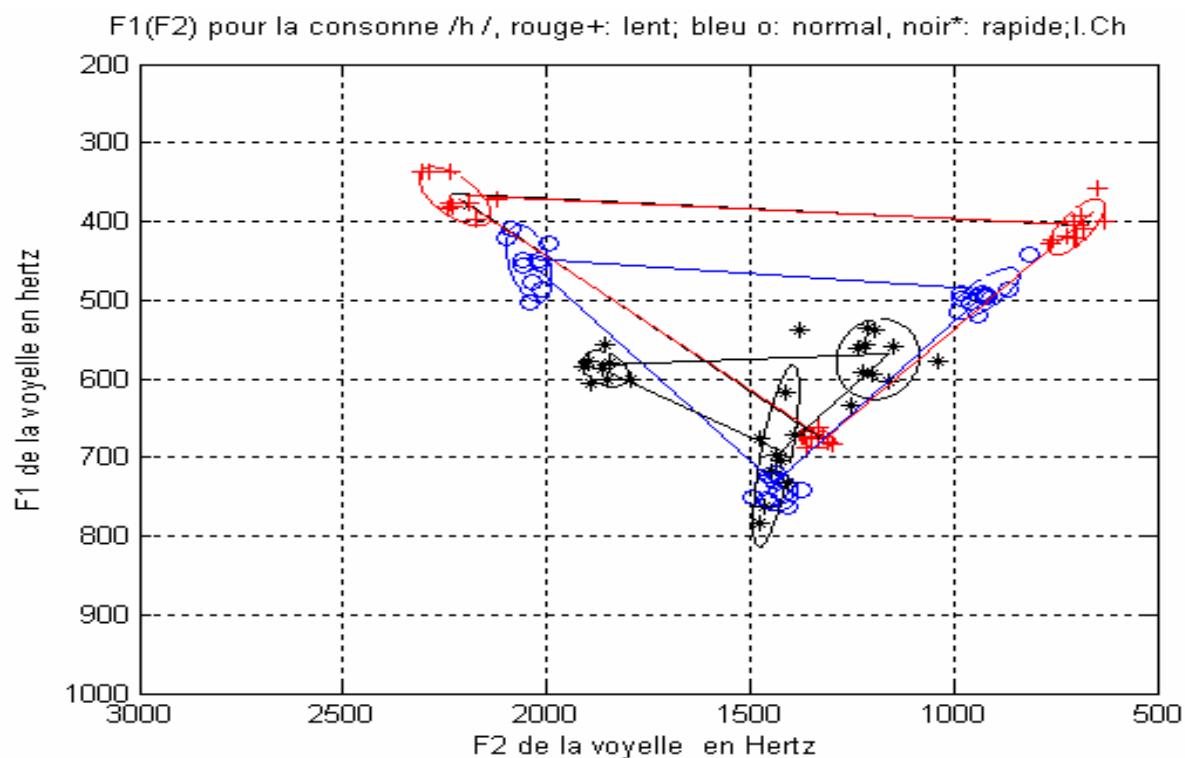


F1(F2) pour la consonne /h/, rouge+: lent; bleu o: normal, noir*: rapide;Saadi

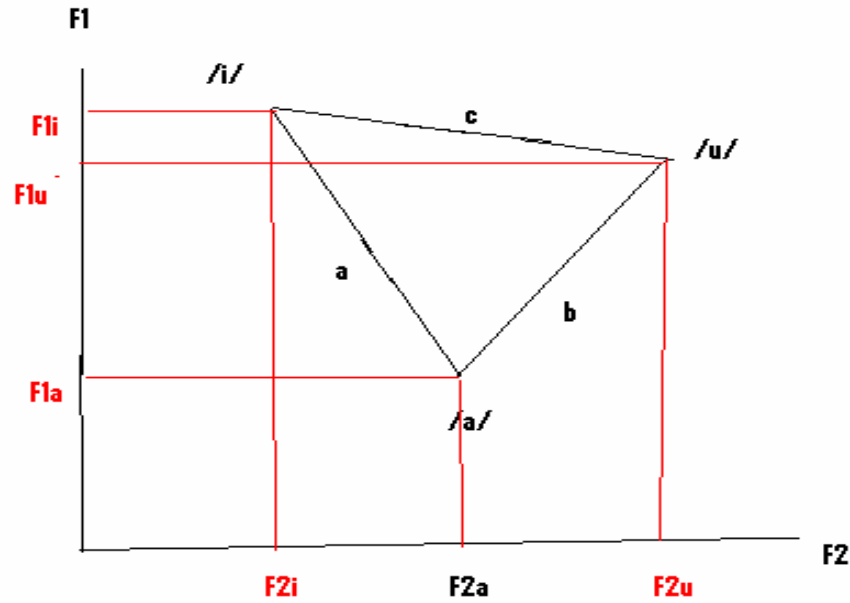


F1(F2) pour la consonne /h/, rouge+: lent; bleu o: normal, noir*: rapide;NA





Figures 3. 5. Triangles vocaliques tracé à partir des ellipses de dispersion, en fonction du débit pour chaque locuteur et chaque contexte consonantique
En bleu : débit normal ; en rouge : débit lent et en noir : débit rapide



Figures 3. 6. Représentation du triangle vocalique pour le calcul de sa surface

On notera ici que les sommets du triangle (voir figure 3.6) correspondent aux centres des ellipses de dispersion : /a/, /i/, /u/

Dans notre cas :

Le côté «a » est égal à la distance entre la voyelle /a/et la voyelle /i/

Le côté «b » est égal à la distance entre la voyelle /a/ et la voyelle /u/

Le côté «c » est égal à la distance entre la voyelle /i/ et la voyelle /u/

Les coordonnées des trois voyelles seront alors dans le plan (F2, F1) :

$$\begin{array}{ll} /a/ : x_a=F2a & y_a=F1a \\ /i/ : x_i=F2i & y_i=F1i \\ /u/ : x_u=F2u & y_u=F1u \end{array}$$

$$\begin{aligned} a &= \sqrt{(x_i - x_a)^2 + (y_i - y_a)^2} = \\ a &= \sqrt{(F2i - F2a)^2 + (F1i - F1a)^2} \end{aligned} \quad (3.6)$$

$$\begin{aligned} b &= \sqrt{(x_u - x_a)^2 + (y_u - y_a)^2} = \\ b &= \sqrt{(F2u - F2a)^2 + (F1u - F1a)^2} \end{aligned} \quad (3.7)$$

$$\begin{aligned} c &= \sqrt{(x_u - x_i)^2 + (y_u - y_i)^2} = \\ &= \sqrt{(F2u - F2i)^2 + (F1u - F1i)^2} \end{aligned} \quad (3.8)$$

$$p = (a + b + c) / 2 \quad (3.9)$$

3.4.1.4. Résultats des surfaces des triangles vocaliques

Les valeurs des surfaces des triangles vocaliques sont données par les tableaux 3.3, dans lesquels nous retrouvons : les valeurs des premier et deuxième formants par voyelle, les locuteurs et les aires en kHz^2 .

Globalement, les figures 3.7 montrent une réduction de l'espace vocalique avec l'augmentation du débit d'élocution pour les six locuteurs. (Notons cependant que les locutrices NA et LI présentent les surfaces vocaliques les plus élevées). Donc, le débit a un effet significatif sur l'espace vocalique pour les trois locuteurs.

Que pouvons-nous en conclure ?

Lane et al. [14] et bien d'autres auteurs ont remarqué que la taille des espaces vocaliques était réduite en cas de perturbation de la parole. Cependant, une étude concernant l'influence du débit sur le triangle vocalique en parole neutre [10] (c'est-à-dire sans expressivité comme la colère ou la joie), montre que les formants tendent vers une voyelle centrale pour les segments de courte durée. Une différence majeure existe entre le cas neutre et les autres expressivités: Le degré d'articulation n'est plus uniquement dépendant de la variable débit.

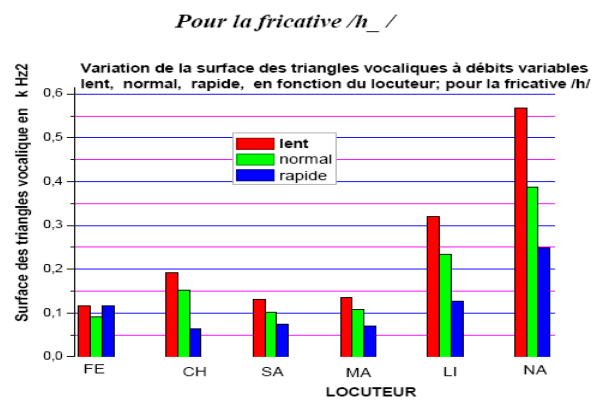
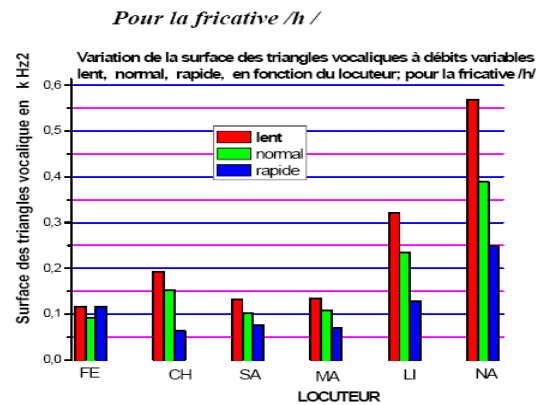
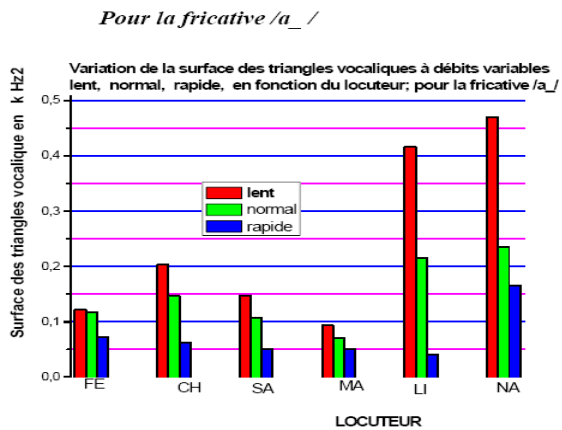
Dans le cas neutre, une accélération et une décélération correspondent respectivement à une réduction et à une expansion du triangle vocalique [10]. Il semble que cette tendance naturelle ne soit pas préservée dans le cas de certaines expressivités. Voir figure 3.8.

DEBIT NORMAL							
Formants	F1a En khz	F2a En khz	F1i En khz	F2i En khz	F1u khz	F2u En khz	Aire en kHz2
Locuteur	Consonne /a_/						
FE	0.754	1.420	0.424	1.956	0.444	1.216	0.1165
CH	0.715	1.354	0.443	2.096	0.480	0.915	0.1469
SA	0.652	1.571	0.415	1.971	0.432	1.040	0.1070
MA	0.516	1.5	0.3626	1.8831	0.3215	1.057	0.0712
LI	0.8261	1.7657	0.4587	2.2041	0.3188	1.1968	0.2157
NA	0.817	1.825	0.352	2.320	0.4668	1.1854	0.2354
	Consonne /h/						
FE	0.666	1.396	0.375	1.792	0.445	1.063	0.09232
CH	0.717	1.425	0.413	2.082	0.452	0.989	0.15291
SA	0.603	1.624	0.373	1.969	0.405	1.039	0.10156
MA	0.5174	1.4825	0.2723	1.9578	0.2628	1.0928	0.1083
LI	0.7451	1.8336	0.4170	2.3193	0.3203	1.1224	0.2338
NA	0.8251	1.7946	0.3620	2.5874	0.324	0.975	0.3884
	Consonne /h_/						
FE	0.720	1.355	0.403	1.949	0.450	1.135	0.11512
CH	0.743	1.436	0.453	2.043	0.491	0.922	0.15064
SA	0.647	1.557	0.391	1.948	0.3619	0.9600	0.13251
MA	0.5193	1.5293	0.2343	2.2166	0.2562	1.1717	0.14140
LI	0.8060	1.8193	0.3103	2.2797	0.3439	1.3098	0.23270
NA	0.8333	1.7769	0.3775	2.3264	0.4124	1.0592	0.27920

	DEBIT LENT						
Formants	F1a En khz	F2a En khz	F1i En khz	F2i En khz	F1u khz	F2u En khz	Aire en kHz2
Locuteur	Consonne /a_/						
FE	0.707	1.433	0.400	1.889	0.420	1.066	0.12160
CH	0.674	1.353	0.365	2.193	0.413	0.748	0.20300
SA	691,9	1.537	0.353	2.010	0.379	0.923	0.14680
MA	0.452	1.5	0.2618	2.0531	0.297	0.972	0.09310
LI	0.8129	1.7271	0.2497	2.4596	0.237	0.997	0.41650
NA	0.808	1.823	0.229	2.558	0.280	0.867	0.47080
	Consonne /h/						
FE	0.671	1.397	0.369	1.801	0.433	0.945	0.11620
CH	0.647	1.430	0.369	2.143	0.364	0.770	0.19260
SA	0.606	1.610	0.350	2.031	0.372	0.927	0.13180
MA	0.5245	1.5140	0.2298	2.1232	0.262	1.1458	0.13420
LI	0.7427	1.8021	0.3276	2.4615	0.258	1.0243	0.32100
NA	0.9120	1.8345	0.3686	2.6561	0.237	0.764	0.56810
	Consonne /h_/						
FE	0.765	1.328	0.403	1.915	0.421	1.005	0.1592
CH	0.676	1.335	0.368	2.225	0.407	0.699	0.21791
SA	0.655	1.528	0.371	2.003	0.407	0.957	0.14016
MA	0.550	0.1539	0.2211	1.9871	0.233	0.976	0.15540
LI	0.8546	1.7855	0.2128	2.3869	0.258	1.0243	0.42350
NA	0.7837	1.7379	0.3478	2.6767	0.212	0.866	0.45840

	DEBIT RAPIDE						
Formants	F1a	F2a	F1i	F2i	F1u	F2u	Aire en kHz2
Locuteur	Consonne /a_/						
FE	0.710	1.402	0.476	1.919	0.498	1.250	0.07258
CH	0.763	1.350	0.573	1.791	0.584	1.110	0.06229
SA	0.736	1.600	0.540	1.839	0.554	1.304	0.05091
MA	0.5195	1.529	0.3626	1.883	0.317	1.338	0.05080
LI	0.7923	1.7599	0.237	0.996	0.4346	1.2531	0.04100
NA	0.7887	1.751	0.374	2.271	0.5257	1.2802	0.16600
	Consonne /h/						
FE	0.731	1.410	0.404	1.835	0.470	1.035	0.11666
CH	0.761	1.476	0.569	1.840	0.586	1.149	0.06329
SA	0.645	1.718	0.408	1.984	0.439	1.177	0.07542
MA	0.4764	1.4944	0.2361	1.7650	0.2495	1.1696	0.06970
LI	0.7778	1.8358	0.4651	2.0850	0.4592	1.2720	0.12780
NA	0.6920	1.7319	0.3578	2.6354	0.3950	1.0403	0.24970
	Consonne /h_/						
FE	0.751	1.452	0.466	1.928	0.471	1.276	0.09178
CH	0.689	1.432	0.585	1.861	0.575	1.186	0.03740
SA	0.808	1.772	0.558	2.009	0.581	1.251	0.07599
MA	0.5049	1.5657	0.2833	2.1446	0.2608	1.2935	0.10080
LI	0.7151	1.8214	0.4744	2.1639	0.4751	1.5275	0.07650
NA	0.8341	1.7888	0.3789	2.4681	0.4183	1.1874	0.27810

Tableau 3. 3. Calcul des surfaces vocalique en fonction du contexte consonantique, du locuteur et du débit



Figures 3. 7. Surface des triangles vocaliques en fonction du contexte consonantique, du locuteur et du débit

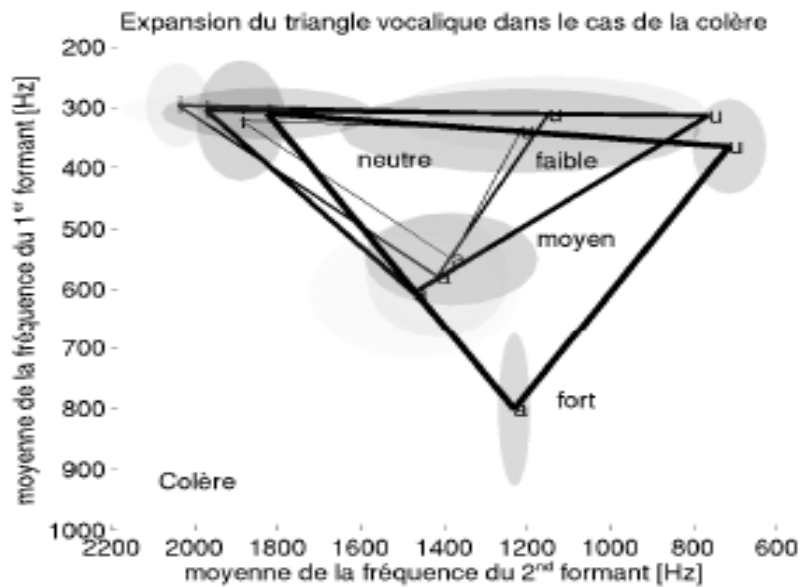


Figure 3. 8. Triangle vocalique neutre et selon trois niveaux d'intensité de la colère extravertie

D'après Gendrot et al. [10], en cas de parole neutre, comme dans le cas de variation de débit, la réduction de l'espace vocalique est réalisée suivant une centralisation des voyelles. Nos résultats montrent en effet une réduction de l'espace vocalique avec l'augmentation du débit, globalement pour les 6 locuteurs. Il serait bon de savoir si cette réduction de l'espace vocalique a aussi l'effet d'une centralisation des voyelles, comme dans le cas d'une parole neutre.

La figure 3.9 [4] montre un exemple de centralisation des voyelles /a/ et /i/ dans l'espace vocalique

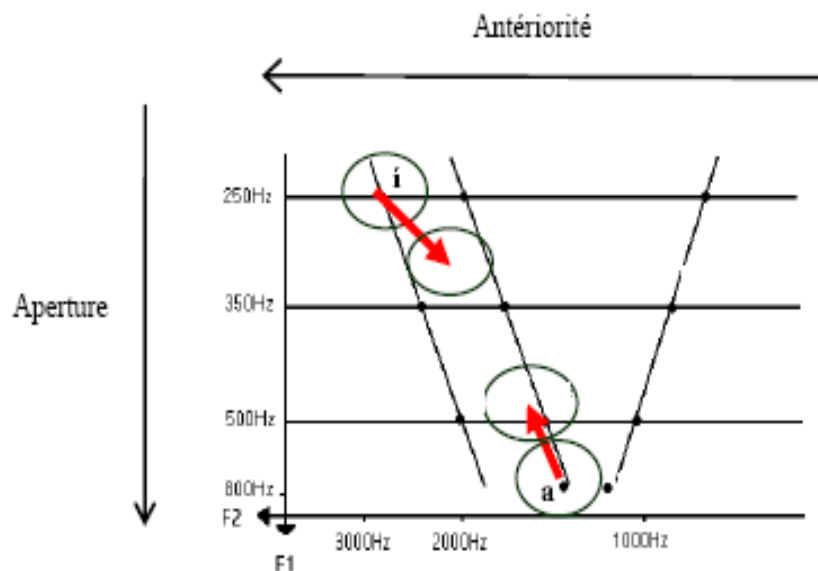


Figure 3. 9. Exemple de centralisation des patrons articulatoire-acoustiques lors de la production des voyelles à l'étude

En guise de constatation, on remarque que la valeur de F1 augmente avec le degré d'aperture et celle de F2 avec le degré d'antériorité. Articulairement, une centralisation du /i/, voudrait dire que le [i] deviendrait moins antérieur et plus ouvert alors que le [a] serait un peu moins postérieur et un peu moins ouvert.

Ces changements influenceraient les valeurs formantiques, et une centralisation des voyelles impliquerait :

Une augmentation de F1 et une diminution de F2 pour [i];

Une augmentation de F2 et une diminution du F1 pour [a].

Une augmentation du F1 et du F2 pour le /u/.

Nous allons donc à présent étudier les variations des valeurs moyennes des formants F1 et F2 pour chaque locuteur en fonction du débit, afin de voir s'il y a centralisation des voyelles /a/, /i/, /u/.

3.5 Calcul des variations des formants F1 et F2 en fonction du débit d'élocution

Les tableaux 3.4 illustrent les valeurs moyennes des formants F1 et F2 de chaque voyelle V_i ($i=/a/, /i/, /u/$), en contexte consonantique /Cj/ : $j=/_a/, /_h/, /_h/_$ et ce pour chaque locuteur, dans le /CjVil/ soit : /a_al/, /a_il/, /a_ul/, /hal/, /hil/, /hul/, /h_al/, /h_il/, /h_ul/

Les abréviations ont été attribuées pour :

F1N : valeur moyenne du formant F1 en débit normal
 F2N : valeur moyenne du formant F2 en débit normal
 F1L : valeur moyenne du formant F1 en débit lent
 F2L : valeur moyenne du formant F2 en débit lent
 F1R : valeur moyenne du formant F1 en débit rapide
 F2R : valeur moyenne du formant F2 en débit rapide

Nous trouvons dans ce tableau des colonnes réservées aux différences des formants F1 et F2 :
 Différence des F1 entre le débit normal et le débit rapide : **F1N-F1R**
 Différence des F2 entre le débit normal et le débit rapide : **F2N-F2R**
 Différence des F1 entre le débit lent et le débit normal : **F1L-F1N**
 Différence des F2 entre le débit lent et le débit normal : **F2L-F2N**

Les flèches montrent le sens de déplacement de la voyelle, qui nous permet d'affirmer s'il y a centralisation

En contexte vocalique /a/

Voyelle /a/	F1La	F2La	F1Na	F2Na	F1Ra	F2Ra	F1N-F1R	F2N-F2R	sens	F1L-F1N	F2L-F2N	sens	centralise
Locuteurs	En contexte consonantique /a_/												
FE	0.707	1.433	0.754	1.420	0.710	1.402	0.0440	0.0180	↗	-0.0470	0.0130	↘	non
CH	0.674	1.353	0.715	1.354	0.763	1.350	-0.0480	0.0040	↘	-0.0410	-0.0010	↗	non
SA	0.691	1.537	0.652	1.571	0.736	1.600	-0.0840	-0.0290	↗	0.0390	-0.0340	↘	non
MA	0.452	1.5	0.516	1.5	0.5195	1.529	-0.0035	-0.0290	↗	-0.0640	0		non
LI	0.812	1.727	0.8261	1.765	0.7923	1.759	0.0338	0.0058	↗	-0.0132	-0.0386	↘	non
NA	0.808	1.823	0.817	1.825	0.7887	1.751	0.0283	0.0740	↗	-0.0090	-0.0020	↘	non
Locuteurs	En contexte consonantique /h_/												
FE	0.671	1.397	0.666	1.396	0.731	1.410	-0.0650	-0.0140	↗	0.0050	0.0010	↘	non
CH	0.647	1.430	0.717	1.425	0.761	1.476	-0.0440	-0.0510	↗	-0.0700	0.0050	↘	non
SA	0.606	1.610	0.603	1.624	0.645	1.718	-0.0420	-0.0940	↗	0.0030	-0.0140	↘	non
MA	0.524	1.5140	0.5174	1.482	0.4764	1.494	0.0410	-0.0119	↗	0.0071	0.0315	↘	non
LI	0.742	1.8021	0.7451	1.833	0.7778	1.358	-0.0327	-0.0022	↗	-0.0024	-0.0315	↘	non
NA	0.912	1.834	0.8251	1.794	0.6920	1.731	0.1331	0.0627	↗	0.0869	0.0399	↘	non
Locuteurs	En contexte consonantique /h_/_/												
FE	0.765	1.328	0.720	1.355	0.751	1.452	-0.0310	-0.0970	↗	0.0450	-0.0270	↘	non
CH	0.676	1.335	0.743	1.436	0.689	1.432	0.0540	0.0040	↗	-0.0670	-0.1010	↘	non
SA	0.655	1.528	0.647	1.557	0.808	1.772	-0.1610	-0.2150	↗	0.0080	-0.0290	↘	non
MA	0.550	0.1539	0.5193	1.529	0.5049	1.565	0.0144	-0.0364	↗	0.0307	-1.3754	↘	non
LI	0.854	1.7855	0.8060	1.819	0.7151	1.821	0.0909	-0.0021	↗	0.0486	-0.0338	↘	non
NA	0.783	1.7379	0.754	1.420	0.8341	1.788	-0.0801	-0.3688	↗	0.0297	0.3179	↘	non

En contexte vocalique /i/

Voyelle /a/	F1La	F2La	F1Na	F2Na	F1Ra	F2Ra	F1N-F1R	F2N-F2R	sens	F1L-F1N	F2L-F2N	sens	centralise
Locuteurs	En contexte consonantique /a_/												
FE	0.707	1.433	0.754	1.420	0.710	1.402	0.0440	0.0180	↗	-0.0470	0.0130	↘	non
CH	0.674	1.353	0.715	1.354	0.763	1.350	-0.0480	0.0040	↘	-0.0410	-0.0010	↙	non
SA	0.691	1.537	0.652	1.571	0.736	1.600	-0.0840	-0.0290	↙	0.0390	-0.0340	↖	non
MA	0.452	1.5	0.516	1.5	0.5195	1.529	-0.0035	-0.0290	↙	-0.0640	0		non
LI	0.812	1.727	0.8261	1.765	0.7923	1.759	0.0338	0.0058	↗	-0.0132	-0.0386	↙	non
NA	0.808	1.823	0.817	1.825	0.7887	1.751	0.0283	0.0740	↗	-0.0090	-0.0020	↙	non
Locuteurs	En contexte consonantique /h/												
FE	0.671	1.397	0.666	1.396	0.731	1.410	-0.0650	-0.0140	↙	0.0050	0.0010	↗	non
CH	0.647	1.430	0.717	1.425	0.761	1.476	-0.0440	-0.0510	↙	-0.0700	0.0050	↘	non
SA	0.606	1.610	0.603	1.624	0.645	1.718	-0.0420	-0.0940	↙	0.0030	-0.0140	↖	non
MA	0.524	1.5140	0.5174	1.482	0.4764	1.494	0.0410	-0.0119	↖	0.0071	0.0315	↗	non
LI	0.742	1.8021	0.7451	1.833	0.7778	1.358	-0.0327	-0.0022	↙	-0.0024	-0.0315	↙	non
NA	0.912	1.834	0.8251	1.794	0.6920	1.731	0.1331	0.0627	↗	0.0869	0.0399	↗	non
Locuteurs	En contexte consonantique /h_/												
FE	0.765	1.328	0.720	1.355	0.751	1.452	-0.0310	-0.0970	↙	0.0450	-0.0270	↖	non
CH	0.676	1.335	0.743	1.436	0.689	1.432	0.0540	0.0040	↗	-0.0670	-0.1010	↙	non
SA	0.655	1.528	0.647	1.557	0.808	1.772	-0.1610	-0.2150	↙	0.0080	-0.0290	↖	non
MA	0.550	0.1539	0.5193	1.529	0.5049	1.565	0.0144	-0.0364	↖	0.0307	-1.3754	↖	non
LI	0.854	1.7855	0.8060	1.819	0.7151	1.821	0.0909	-0.0021	↖	0.0486	-0.0338	↖	non
NA	0.783	1.7379	0.754	1.420	0.8341	1.788	-0.0801	-0.3688	↙	0.0297	0.3179	↗	non

En contexte vocalique /u/

Voyelle /a/	F1La	F2La	F1Na	F2Na	F1Ra	F2Ra	F1N-F1R	F2N-F2R	sens	F1L-F1N	F2L-F2N	sens	centralise
Locuteurs	En contexte consonantique /a_/												
FE	0.707	1.433	0.754	1.420	0.710	1.402	0.0440	0.0180	↗	-0.0470	0.0130	↘	non
CH	0.674	1.353	0.715	1.354	0.763	1.350	-0.0480	0.0040	↘	-0.0410	-0.0010	↙	non
SA	0.691	1.537	0.652	1.571	0.736	1.600	-0.0840	-0.0290	↙	0.0390	-0.0340	↖	non
MA	0.452	1.5	0.516	1.5	0.5195	1.529	-0.0035	-0.0290	↙	-0.0640	0		non
LI	0.812	1.727	0.8261	1.765	0.7923	1.759	0.0338	0.0058	↗	-0.0132	-0.0386	↙	non
NA	0.808	1.823	0.817	1.825	0.7887	1.751	0.0283	0.0740	↗	-0.0090	-0.0020	↙	non
Locuteurs	En contexte consonantique /h/												
FE	0.671	1.397	0.666	1.396	0.731	1.410	-0.0650	-0.0140	↙	0.0050	0.0010	↗	non
CH	0.647	1.430	0.717	1.425	0.761	1.476	-0.0440	-0.0510	↙	-0.0700	0.0050	↘	non
SA	0.606	1.610	0.603	1.624	0.645	1.718	-0.0420	-0.0940	↙	0.0030	-0.0140	↖	non
MA	0.524	1.5140	0.5174	1.482	0.4764	1.494	0.0410	-0.0119	↖	0.0071	0.0315	↗	non
LI	0.742	1.8021	0.7451	1.833	0.7778	1.358	-0.0327	-0.0022	↙	-0.0024	-0.0315	↙	non
NA	0.912	1.834	0.8251	1.794	0.6920	1.731	0.1331	0.0627	↗	0.0869	0.0399	↗	non
Locuteurs	En contexte consonantique /h_/												
FE	0.765	1.328	0.720	1.355	0.751	1.452	-0.0310	-0.0970	↙	0.0450	-0.0270	↖	non
CH	0.676	1.335	0.743	1.436	0.689	1.432	0.0540	0.0040	↗	-0.0670	-0.1010	↙	non
SA	0.655	1.528	0.647	1.557	0.808	1.772	-0.1610	-0.2150	↙	0.0080	-0.0290	↖	non
MA	0.550	0.1539	0.5193	1.529	0.5049	1.565	0.0144	-0.0364	↖	0.0307	-1.3754	↖	non
LI	0.854	1.7855	0.8060	1.819	0.7151	1.821	0.0909	-0.0021	↖	0.0486	-0.0338	↖	non
NA	0.783	1.7379	0.754	1.420	0.8341	1.788	-0.0801	-0.3688	↙	0.0297	0.3179	↗	non

Tableau 3. 4. Variation des formants avec le débit, en fonction du contexte consonantique, pour chaque locuteur et centralisation

Tableau récapitulatif des tableaux 3

	/a/	/i/	/i/	/u/	/u/		
Locuteur		non	oui	non	oui	Total de oui	Total de non
FE	0	3	0	1	2	2	4
CH	0	0	3	0	3	6	0
SA	0	2	1	1	2	3	3
MA	0	3	0	2	1	1	5
LI	0	0	3	0	3	6	0
NA	0	2	1	1	2	3	3

Tableau 3. 5 : Nombre de cas de centralisation des voyelles (/a/, /i/, /u/) en fonction du locuteur

Sur Les tableaux 3.4, les flèches indiquent le sens de déplacement des points M(F2, F1) relatifs à chaque voyelle, lors de la variation des formants, dans l'espace vocalique (F2, F1), lorsque nous passons du débit lent vers le débit normal (LN) ou du débit normal vers le débit rapide (NR). Notons qu'en cas de centralisation, le déplacement des points M(F2, F1) en situation (LN) doit être dans le même que celui du cas (NR) (en tenant compte de la figure 3.9).

Nous pourrions dire alors que d'après le tableau 3.5, nous notons que:

- pour la voyelle /a/ : Aucun locuteur ne centralise
- Pour les voyelles /i/ et /u/:
CH et LI centralisent quelque soit le contexte consonantique
SA et NA centralisent moins que CH et LI et plus que FE et MA
MA et FE sont ceux qui centralisent le moins

3.6 Conclusion

Dans cette étude nous avons analysé les patrons temporels et acoustiques de locuteurs ayant des parlés qui vont en s'éloignant de la langue seconde (ici l'arabe littéral) du corpus de phrases qu'ils ont produites, en contrainte de débit

Les résultats obtenus semblent :

- d'une part rejoindre ceux de certains chercheurs en production de parole perturbée, par rapport à la réduction vocalique, la centralisation la variation des patrons temporels et l'absence de centralisation pour la voyelle /a/. En effet, nos résultats montrent qu'aucun locuteur ne centralise pour le /a/, ce qui rejoint les résultats trouvés par Lindblom [13] et Ferbach-Hecker [8] qui n'ont pas observé de centralisation pour le [a].
- d'autre part, être liés au parlé des locuteurs. En effet :
- Les libanais CH et LI, dont la langue maternelle est proche de l'arabe littéral, centralisent le plus et gardent des durées relatives constantes. Cela pourrait s'interpréter par une compensation de la vitesse par une centralisation des voyelles, pour atteindre leurs cibles acoustiques. Ce qui est synonyme d'un bon contrôle moteur [9] en contrainte de débit
- Les algérois FE et MA, dont la langue maternelle est moins proche de l'arabe littéral que le libanais, viennent en deuxième position : Ils présentent des patrons temporels variables avec

le débit et ils centralisent le moins. Donc on peut considérer qu'ils ont besoin d'une seule compensation en contrainte de débit pour atteindre les cibles acoustiques. Comme ils centralisent très peu, nous pouvons dire qu'ils ont un contrôle moteur moins bon [9] que les libanais en contrainte de débit.

- Les kabyles SA et NA ont besoin de deux compensations (temporelle et acoustique) pour atteindre leurs cibles acoustiques, donc deux fois plus d'effort en contrainte de débit. Ce qui est synonyme de contrôle moteur relativement mauvais en contrainte de débit.

Nos résultats montrent en outre que:

CH et LI suivent le schéma d'une parole neutre prononcée par un locuteur ayant un contrôle fin. Ils se trouvent vraiment dans le premier cas de la stratégie de Gay. (stratégie d'amplitude (*amplitude strategy*))

- FE et MA centralisent très peu, donc ils se trouvent dans le 2ème cas de la stratégie de Gay (ou stratégie de vitesse (*velocity strategy*))., il y a variation des patrons temporels en variation de débit mais les patrons acoustiques varient très peu. Ils ont un bon contrôle moteur mais moins fin que LI et CH en contrainte de débit.

- SA et NA font varier leurs patrons temporels et leurs patrons acoustiques en variation de débit, donc, ce sont ceux qui font le plus d'efforts pour atteindre leurs cibles acoustiques. Cela laisse à supposer qu'ils ont un mauvais contrôle moteur en contrainte de débit.

Les résultats obtenus dans cette étude nous ont permis de dégager les stratégies adoptées en langue seconde par six locuteurs, en utilisant comme paradigme de recherche le débit d'élocution ; cependant pour que ces résultats soient plus robustes, il serait intéressant :

- de faire la même étude sur la langue maternelle de chaque locuteur et de vérifier si les stratégies employées pour leur langue première sont les mêmes que pour la langue seconde (l'arabe littéral) en contrainte de débit.

- d'augmenter le nombre de locuteurs (ici six).

- de changer la langue seconde afin de voir si les mêmes stratégies sont adoptées par les locuteurs pour d'autres langues secondes.

Par ailleurs, la méthodologie développée dans cette étude pourrait :

- être appliquée à d'autres types de contraintes artificielles pour viser d'autres objectifs,

- permettre l'analyse de parole pathologique par le biais des patrons temporels et fréquentiels du signal acoustique.

3.7 Bibliographie du chapitre 3

- [1] J.Aubin ; « *Effets acoustiques et articulatoires des perturbations labiales sur la parole des enfants et des adultes* ». Université du Québec à Montréal. Décembre 2005.
- [2] Beller, G., Schwarz, D., Hueber, T. and Rodet, X. (2006), “*Speech rates in french expressive speech*”, in *Speech Prosody*, SproSig, Dresden: ISCA
- [3] G. Beller. « *Influence de l’expressivité sur le degré d’articulation* ». .Actes des VIIèmes RJC Parole, Paris, 5 et 6 juillet 2007. Page 24 IRCAM - Institut de Recherche et Coordination Acoustique Musique
- [4] A.Calabrino. « *Effets acoustiques du débit sur la production de la parole chez des locuteurs enfants et adultes* ». Actes du Xe Colloque des étudiants en sciences du langage : 61-83, 2006.
- [5] L.Falek & O.Bouferroum & A. Djeradi “*Analyze effects of the flow on the vocalic reduction and the coarticulation in sequences CV of Arabic pharyngeal*”. The Journal of the Acoustical Society of America -- May 2008 -- Volume 123, Issue 5, p. 3744.
- [6] L.Falek, R.Djeradi, and A.Djeradi (À paraître) “*Effets du débit sur les paramètres acoustiques et temporels du signal de parole*”. Canadian Acoustics journal. Janvier 2009
- [7] McFarland, D. H., S. R. Baum et C. Chabot. 1996. «*Speech compensation to structural modifications of the oral cavity*». Journal of the Acoustical Society of America, vol 100, no2.p.1093-1104.
- [8] V. Ferbach-Hecker. La résistivité de la qualité des voyelles orales du français. In *SCOLIA*, volume 20, pages 115-134, 2005.
- [9] Gay, T. 1978. “*Effects of speaking rate on vowel formant movements*”. *Journal of the Acoustical Society of America* 63(1):223-230.
- [10] Gendrot, C. and Adda-Decker, M. (2004), “*Analyses formantiques automatiques de voyelles orales : évidence de la réduction vocalique en langues française et allemande*”, in *MIDL*.
- [11] F.Hirsch. « *Étude de la structure formantique des voyelles produites par des locuteurs bègues en vitesses d’élocution normale et rapide* » 2005. Institut de Phonétique de Strasbourg – E.A. 1339 – LiLPa -Composante Parole et Cognition - Université Marc Bloch
- [12] J. A. Jones & K. G. Munhall. 2003. «*Learning to produce speech with an altered vocal tract:The role of auditory feedback*». *Journal of the Acoustical Society of America*, vol.113, no 1, p. 532-543.

[13] Lindblom, B. (1983), “*Economy of Speech Gestures*”, vol. The Production of Speech, Springer-Verlag, New-York.

[14] Lane, H., M. Denny, F. H. Guenther, M. Matthies, L. Ménard, J. S. Perkell, E. Stockmann, M. Tiede, J. Vick et M. Zandipour. 2005. «*Effects of bite blocks and hearing status on vowel production*». *Journal of the Acoustical Society of America*, vol.118, no 3, p.1636-1646.

[15] Perkell, J. S., M. L. Matthies, M. A. Svirsky et M. I. Jordan. 1993. «Trading relations between tongue-body raising and lip rounding in production of the vowel /u/: A pilot “motor equivalence” study». *Journal of the Acoustical Society of America*, vol. 93, no 5, 2948-2961.

[16] Perkell, J. S., F. H. Guenther, H. Lane, M. Matthies, P. Perrier, J. Vick, R. Wilhelms-Tricarico et M. Zandipour. 2000. «*A theory of speech motor control and supporting data from speakers with normal hearing and with profound hearing loss*». *Journal of Phonetics*, vol. 28, p. 233-272.

Chapitre 4

Simulation de phrases à débits variables par un modèle de production de la parole

4.1 Introduction

Les premiers travaux sur la modélisation de la production de la parole se sont attachés à simuler le comportement d'un tuyau acoustique, ayant les dimensions du conduit vocal, discrétisé en une vingtaine d'éléments cylindriques mis bout à bout, excités par des sources sonores ayant les mêmes caractéristiques que la source laryngienne et/ou les sources de bruit.

De 1948 à 1960, ces études acoustiques ont permis d'asseoir une théorie de la production de la parole. La modélisation précise des sources de bruit fait encore l'objet de nombreuses études. Ainsi les voyelles sont caractérisées par une série de résonances principales (les formants) ; au voisinage des consonnes, ces formants présentent des transitions. Les 3 premiers formants suffisent en général à l'auditeur pour identifier une voyelle. Ainsi les études acoustiques ont abouti à la construction de synthétiseur à formants où les n tuyaux de l'analogie du conduit vocal, sont remplacés par des filtres en cascades. Cette représentation de la parole en formants est efficace et explicative tant d'un point de vue acoustique que perceptif. Mais la description de l'évolution temporelle des formants pour la parole continue, est extrêmement complexe et il n'existe pas de modèles généraux intégrant cette particularité. Nous ne reviendrons pas sur les causes de la variabilité de la parole à l'origine des difficultés de passage du domaine acoustique à celui du phonétique, nous dirons seulement que c'est la cause qui pousse les chercheurs à spécifier l'organisation temporelle sur le plan articulatoire. Il faut arriver non seulement à modéliser les déplacements des organes à l'origine des évolutions du conduit vocal, mais aussi inférer des signaux de contrôle de ces organes. Dans cette étape, on utilise largement les nouvelles techniques de l'imagerie médicale, radio cinématographique, scanner, IRM, comme outils d'acquisition de données physiologiques et anatomiques.

Les modèles articulatoires du conduit vocal permettent maintenant de générer la forme du conduit en deux voir trois dimensions, à partir de commandes qui simulent le comportement des principaux articulateurs (langue, mâchoires, lèvres, voile du palais...). Les modèles du conduit vocal sont donc en passe de devenir des robots parlants, et comme tous les robots, ils posent des problèmes de pilotage. Avec un certain nombre de paramètres de commandes, il existe de multiples possibilités de parvenir au même résultat : comment donc arriver à contraindre pour simuler « la bonne solution », c'est à dire celle qui est optimalement utilisée par le locuteur ? Une des réponses passe par une meilleure connaissance des articulateurs et de leurs évolutions parallèles, mais aussi par une meilleure compréhension du système intelligent qui les contrôle.

Pour notre part, nous nous intéressons à la production des fricatives arrières de l'arabe à vitesse d'élocution variable. Nous cherchons à recenser les phénomènes empruntés donc

non naturels, sur les plans du timing et de l'acoustique, que le locuteur amène pour arriver à produire ces sons, afin de les intégrer dans un modèle acoustique de production de la parole. Le but recherché est une meilleure compréhension du phénomène.

Le modèle que nous avons choisi tient compte du caractère fréquentiel et temporel dans la production de la parole, par le biais de ces paramètres de commande en dynamique. Il s'agit du synthétiseur à formants de Klatt (SAF) [7]. C'est un modèle source/filtre.

Le but est d'abord de comprendre le fonctionnement, puis de réaliser ce modèle. Nous analyserons ensuite sa commande pour voir comment y intégrer le facteur vitesse d'élocution.

Pour ce faire, nous nous sommes basés sur les résultats obtenus dans le chapitre 3, qui nous ont permis de caractériser le signal de parole en contrainte de débit. Nous avons en effet recensé deux paramètres importants : la variation de la distribution du temps dans la phrase et la réduction vocalique. Cependant, nous ne nous sommes pas penchés sur le comportement de la fréquence fondamentale dans ces conditions.

Un retour vers le fonctionnement de l'appareil phonatoire humain a attiré notre attention sur l'importance de la variation du débit d'air des poumons à travers la glotte, lors de l'élocution à vitesse variable. En effet, la pression sub-glottique est très sensible à la contrainte de débit. Lorsque nous parlons vite ou lentement, le débit d'air délivré par les poumons varie, ce qui provoque une augmentation ou une diminution de la pression sub-glottique.

Le contrôle du débit d'air est effectué par les cordes vocales qui ont pour rôle de moduler le passage de l'air à travers la glotte.

Le modèle à deux masses des cordes vocales d'ISHIZAKA [6] tient compte de la pression sub-glottique. Il modélise mécaniquement le débit d'air en provenance des poumons à travers la glotte. Nous avons donc jugé intéressant de l'utiliser comme modèle de source, en remplacement du modèle de source de Klatt. Son adaptation au synthétiseur de Klatt est donc nécessaire.

Nous verrons plus loin que la commande du SAF nécessite une dynamique de la fréquence fondamentale F_0 . Un de nos objectifs principaux, dans cette étude, sera donc la génération par le biais du modèle à deux masses d'ISHIZAKA [6] d'une dynamique de la fréquence fondamentale F_0 pour l'intégrer dans la commande du SAF.

Nous exciterons ensuite, l'ensemble des filtres qui servent de modèle de conduit vocal dans le synthétiseur de Klatt, par l'onde de débit glottique $U_g(t)$ générée par le modèle à deux masses.

L'utilisation du modèle à deux masses n'est pas de toute simplicité car l'un des ses problèmes majeurs réside dans son nombre très élevé de paramètres. Ce qui rend sa commande assez délicate.

Dans ce chapitre, nous allons donc :

- dans un premier temps présenter et réaliser le synthétiseur à formants de Klatt (SAF) [7], et étudier sa commande.
- dans un deuxième temps, proposer des solutions pour intégrer le facteur débit dans la commande du modèle de Klatt.
- Enfin, présenter l'étude effectuée sur le modèle à deux masses. Cette dernière a permis d'une part de procéder à la réduction du nombre de paramètres du modèle (pour simplifier sa commande), puis de déterminer une dynamique de F_0 [2] en fonction de la pression sub-glottique.

4.2 Le synthétiseur à formants de Klatt

Un synthétiseur à formants est un outil qui permet de reproduire la parole à partir de paramètres acoustiques du signal de parole (les résonances formantiques du conduit vocal).

En 1980, Dennis Klatt avait proposé un synthétiseur à formants qui incorpore à la fois deux synthétiseurs à formants : série et parallèle (figure 4.1). La qualité de synthèse était ainsi très prometteuse. Le modèle a été incorporé dans plusieurs systèmes TTS (Texte To Speech).

De par son principe, le synthétiseur à formants de Klatt (SAF) simule plus fidèlement le fonctionnement de l'appareil vocal, que beaucoup d'autres types de synthétiseurs comme le vocodeur à canaux. Il est en effet constitué d'un ensemble de filtres résonants dont la courbe de réponse globale en fréquence, reproduit celle du conduit vocal. Les signaux issus d'une source d'impulsions périodiques et d'une source de bruit viennent attaquer ces circuits résonants.

Pour la production des sons non voisés, la source de bruit attaque un canal différent comprenant deux circuits résonants.

4.2.1- Description de la structure du SAF de KLATT

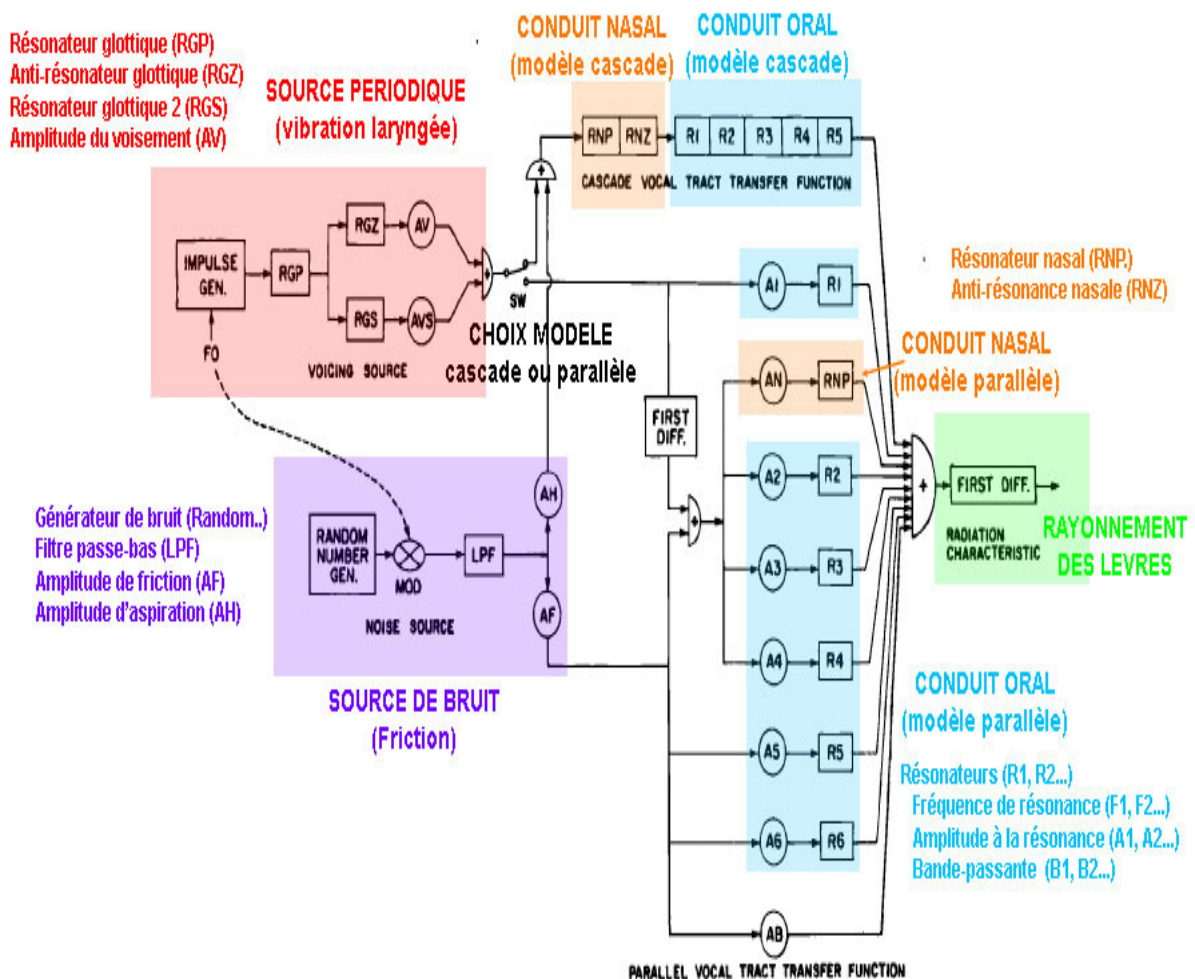


Figure 4. 1 Diagramme bloc du synthétiseur de Klatt. [7]

Ce synthétiseur à formants (**SAF**) est construit à base de filtres numériques qui sont :

-**RGP** (*Resonator of Glottal Pole*) résonateur glottique utilise comme filtre passe-bas pour transformer l'impulsion glottique en une impulsion ayant les caractéristiques ondulatoires et spectrales d'une voix normale ;

-**RGS** (*Resonator of Glottal Sinusoidal*) résonateur glottique sinusoïdal utilisé comme filtre passe-bas pour filtrer la voix normale, obtenue en sortie du résonateur RGP, pour produire une onde glottique quasi-sinusoïdale ;

-**RNP** (*Resonator of Nasal Pole*) résonateur nasal;

-**R1,... R6** résonateurs formantiques;

-**RGZ** (*Resonator of Glottal Zero*) antirésonateur glottique et **RNZ** (*Resonator of Nasal Zero*) antirésonateur du zero nasal;

-**LPF** (*Low Pass Filter*) filtre passe-bas;

-**Diff** et **Read** : filtres passe-haut ;

-**AV** (*Amplitude of Voicing*) amplitude de l'antirésonateur glottique ;

-**AVS** (*Amplitude of Sinusoidal Voicing*) amplitude du résonateur nasal;

-**AN** (*Amplitude of Nasal*) amplitude du résonateur nasal ;

-**AB** (*Amplitude of cascade /parallel Bypass*) amplitude de dérivation.

Chacun des filtres résonants dit « circuit de formants », amplifie une bande de fréquence correspondant à une résonance des cavités en couplage du conduit vocal.

L'entrée du synthétiseur est composée de trois sources d'excitation qui sont :

- la source de voisement
- la source d'aspiration
- la source de friction

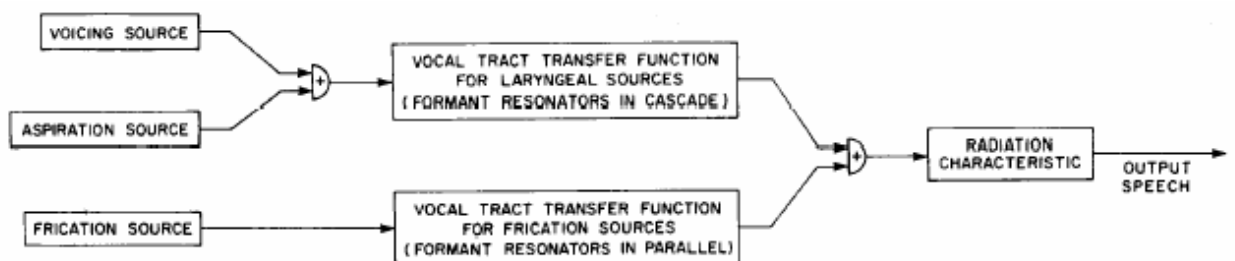


Figure 4. 2. Schéma simplifié du SAF

Ces trois sources vont fournir des signaux d'excitation au conduit buccal et nasal qui va constituer le conduit vocal. Ce dernier est lui aussi constitué de filtres résonants aux fréquences formantiques. En sortie nous trouvons un filtre simulant le rayonnement aux lèvres.

Le synthétiseur permet d'approcher la fonction de transfert du conduit vocal excité au larynx grâce à des résonateurs numériques simulés par une équation aux différences du second ordre. Ce qui introduit un grand nombre de paramètres de commande liés aux caractéristiques de ces filtres.

Pour un résonateur le signal de sortie s'écrit en fonction du signal d'excitation par l'équation:

$$y(nT) = A * x(nT) + B * y(nT - T) + C * y(nT - 2T) \dots (4.1)$$

$$B = 2 * \exp(-\pi * B_w * T) * \cos(-2 * \pi * F_r * T)$$

$$C = -\exp(-2 * \pi * B_w * T)$$

$$A = 1 - B - C$$

$x(nT)$: excitation (entrée du filtre numérique)

$y(nT)$: sortie du filtre numérique

B_w : est la bande passante du filtre

F_r : fréquence centrale du filtre

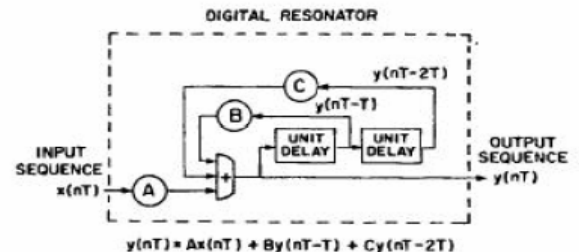


Figure 4.3 résonateur numérique

Chaque filtre dépendra de trois paramètres à régler (amplitude, bande passante et fréquence centrale). Comme le synthétiseur comporte plus de vingt filtres, et que les paramètres de ces derniers vont constituer la commande du SAF, ce qui fait un nombre de paramètres de commande très conséquent.

4.2.2 Réalisation du SAF

Après avoir réalisé les différents filtres constituant les différentes parties de ce modèle nous sommes passés à la partie la plus importante qui est la commande.

4.2.2.1 La commande du synthétiseur

Un ensemble de 34 paramètres (voir tableau 4.1) correspondant aux caractéristiques des filtres numériques, va constituer la commande du SAF. Le choix de ces derniers va décider de la qualité du signal de sortie.

Commander le synthétiseur consiste à trouver un jeu de paramètres acoustiques qui permettent l'obtention d'un son fidèle intelligible qui se rapproche le plus d'un son naturel. Ce qui est d'ailleurs le point sensible de ce synthétiseur.

C'est ainsi qu'après avoir réalisé les différentes parties de ce synthétiseur, nous avons proposé une matrice de commandes dont les colonnes correspondent à différents paramètres, pour un effectif de cinq formants:

Colonnes 1-15 : correspondent aux paramètres F_i , B_i , A_i qui représentent les fréquences et les bandes passantes qui sont en [Hz], et les amplitudes en [dB], ces paramètres sont nécessaires pour la simulation de la fonction de transfert du conduit vocal, et sont calculés à partir du logiciel Winsnoori.

Colonne 16 : correspond à la fréquence fondamentale F_0 [Hz] qui varie selon le sujet.

Colonnes 17-25 : correspondent aux paramètres de la source de voisement qu'on pourra annuler dans le cas d'un son fricatif.

Colonne 26 : amplitude de la source de bruit

Colonnes 27-32 : correspondent aux paramètres de nasalisation.

Colonne 33 : correspond à la bande passante de la source de bruit.

Colonnes 34 : interrupteur sw qui permet de basculer d'une configuration à une autre :

- **sw =0** pour le « fonctionnement parallèle ».
- **sw =1** pour le « fonctionnement cascade » ou série.

N°	V/C	Symbole des différents paramètres	signification	Min	Max	Type
1	V	AV-Db	Amplitude de voisement	0	80	0
2	V	AF	Amplitude de friction	0	80	0
3	V	AH	Amplitude d'aspiration	0	80	0
4	V	AVS	Amplitude quasi-sinus	0	80	0
5	V	F0-Hz	Fréquence fondamentale	0	500	0
6	V	F1	Formant1	150	900	450
7	V	F2	Formant2	500	2500	1450
8	V	F3	Formant3	1300	3500	2450
9	V	F4	Formant4	2500	4500	3300
10	V	FNZ	Fréquence RNZ	200	700	250
11	C	AN-dB	Amplitude nasale	0	80	0
12	C	A1	Amplitude-formant1	0	80	0
13	V	A2	Amplitude-formant2	0	80	0
14	V	A3	Amplitude-formant3	0	80	0
15	V	A4	Amplitude-formant4	0	80	0
16	V	A5	Amplitude-formant5	0	80	0
17	V	A6	Amplitude-formant6	0	80	0
18	V	AB	Amplitude schunt	0	80	0
19	V	B1-Hz	Bande passante de F1	40	500	50
20	V	B2	Bande passante de F2	40	500	70
21	V	B3	Bande passante de F3	40	500	110
22	C	SW	Interrupteur	0(casc)	1(para)	0
23	C	FGP	Fréquence RGP	0	600	0
24	C	BGP	Bande passante de RGP	100	2000	100
25	C	FGZ	Fréquence de RGZ	0	5000	1500
26	C	BGZ	Bande passante de RGZ	100	9000	6000
27	C	B4	Bande passante de F4	100	500	250
28	V	F5	Formant F5	3500	4900	3750
29	C	B5	Bande passante de F5	150	700	200
30	C	F6	Formant F6	4000	4999	4900
31	C	B6	Bande passante de F6	200	2000	1000
32	C	FNP	Fréquence de RNP	200	500	250
33	C	BNP	Bande passante de RNZ	50	500	100
34	C	BNZ	Bande passante de RGS	50	500	100

Tableau 4. 1 Les différents paramètres de commande du SAF

Cependant, il faut noter que chaque type de son nécessite un réglage approprié de tous ces paramètres, pour obtenir une parole intelligible. Une petite étude basée sur l'analyse et l'écoute nous a permis de proposer des matrices de sources spécifiques pour la génération des voyelles, plosives, liquides, nasales et fricatives.

a) Matrice de paramètres de la source pour la simulation des voyelles

Av	Fgp	Bgp	Fgz	Bgz	Avz	Fgs	Bgs	Avs	Af	Fnp	Bnp	Fnz	Bnz	Anp	Anz	Bf
60	0	500	2500	6000	60	0	500	60	-inf	270	90	270	90	0	0	50
60	0	500	2500	6000	60	0	500	60	-inf	270	90	270	90	0	0	50
60	0	500	2500	6000	60	0	500	60	-inf	270	90	270	90	0	0	50
60	0	500	2500	6000	60	0	500	60	-inf	270	90	270	90	0	0	50
60	0	500	2500	6000	60	0	500	60	-inf	270	90	270	90	0	0	50
60	0	500	2500	6000	60	0	500	60	-inf	270	90	270	90	0	0	50
60	0	500	2500	6000	60	0	500	60	-inf	270	90	270	90	0	0	50
60	0	500	2500	6000	60	0	500	60	-inf	270	90	270	90	0	0	50
60	0	500	2500	6000	60	0	500	60	-inf	270	90	270	90	0	0	50
60	0	500	2500	6000	60	0	500	60	-inf	270	90	270	90	0	0	50
60	0	500	2500	6000	60	0	500	60	-inf	270	90	270	90	0	0	50

Tableau 4. 2 Paramètres de commande de la source pour la simulation des voyelles.

Les figures suivantes sont obtenues pour cette simulation.

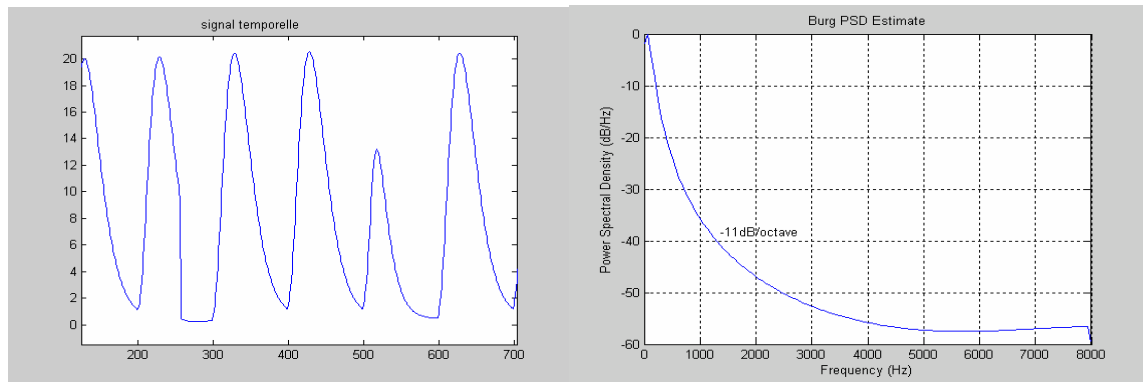


Figure 4. 4 Onde de voisement et son spectre simulés (pour les voyelles).

b) Matrice de paramètres de la source pour la simulation des plosives :

Le tableau (4.3) présente les valeurs des paramètres de commande de la source correspondant à la simulation des plosives (voisées et non voisées).

Av	Fgp	Bgp	Fgz	Bgz	Avz	Fgs	Bgs	Avs	Af	Fnp	Bnp	Fnz	Bnz	Anp	Anz	Bf
-inf	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50
-inf	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50
-inf	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50
-inf	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50
-inf	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50
-inf	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50

Tableau 4. 3 Paramètres de commande de la source pour la simulation des plosives.

Les graphes de la source et de son spectre sont représentés en figure (4.5).

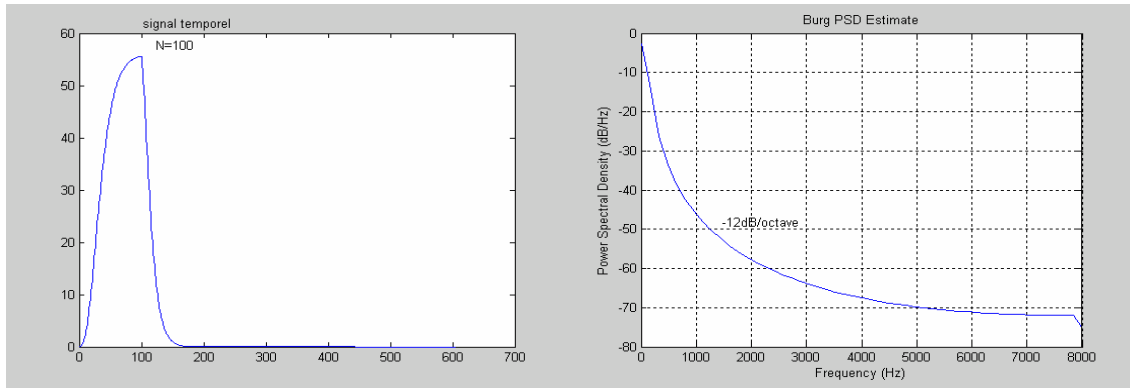


Figure 4. 5 Onde de voisement et son spectre simulés (pour les plosives).

c) Matrice de paramètres de la source pour la simulation des nasales :

Le tableau (4.4) présente les valeurs des paramètres de commande de la source pour la génération des nasales.

Av	Fgp	Bgp	Fgz	Bgz	Avz	Fgs	Bgs	Avs	Af	Fnp	Bnp	Fnz	Bnz	Anp	Anz	Bf
20	0	500	2500	6000	60	0	500	60	-inf	270	90	270	90	0	0	50
20	0	500	2500	6000	60	0	500	60	-inf	270	90	270	90	0	0	50
20	0	500	2500	6000	60	0	500	60	-inf	270	90	270	90	0	0	50
20	0	500	2500	6000	60	0	500	60	-inf	270	90	270	90	0	0	50
20	0	500	2500	6000	60	0	500	60	-inf	270	90	270	90	0	0	50
20	0	500	2500	6000	60	0	500	60	-inf	270	90	270	90	0	0	50

Tableau 4. 4: Paramètres de commande de la source pour la simulation des nasales.

Le résultat obtenu pour cette étude est illustré par la figure (4.6) :

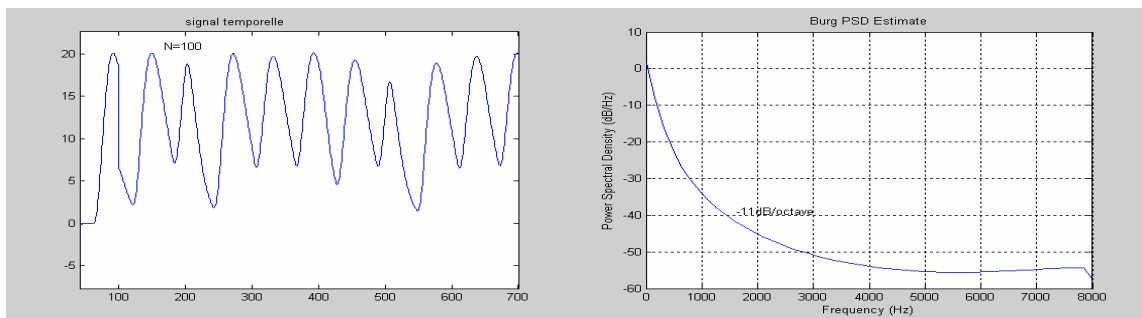


Figure 4. 6 Onde de voisement et son spectre simulés (pour les nasales).

d) Matrice de paramètres de la source pour la simulation des liquides

Le tableau (4.5) présente les valeurs des paramètres de commande de la source, choisies pour la génération des liquides.

Av	Fgp	Bgp	Fgz	Bgz	Avz	Fgs	Bgs	Avs	Af	Fnp	Bnp	Fnz	Bnz	Anp	Anz	Bf
0	0	500	2500	6000	60	0	500	60	60	270	90	270	90	0	0	50
0	0	500	2500	6000	60	0	500	60	60	270	90	270	90	0	0	50
0	0	500	2500	6000	60	0	500	60	60	270	90	270	90	0	0	50
0	0	500	2500	6000	60	0	500	60	60	270	90	270	90	0	0	50
0	0	500	2500	6000	60	0	500	60	60	270	90	270	90	0	0	50
0	0	500	2500	6000	60	0	500	60	60	270	90	270	90	0	0	50

Tableau 4. 5 : Paramètres de commande de la source pour la simulation des liquides.

Les figures obtenues pour cette simulation :

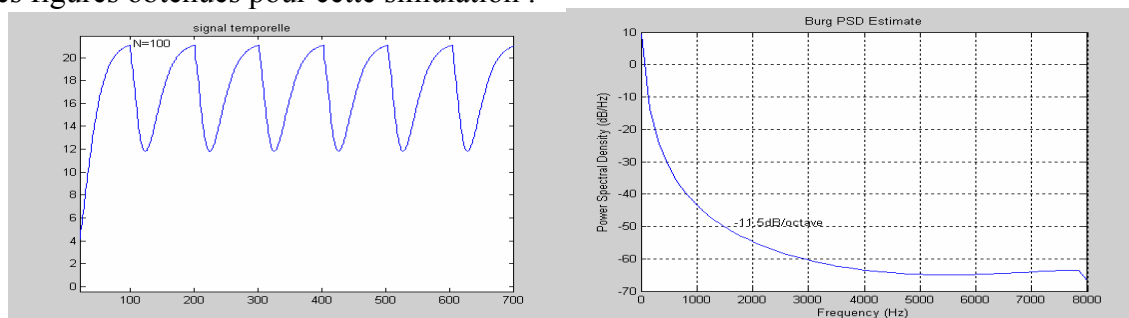


Figure 4.7 : Onde de voisement et son spectre simulés sur Matlab (pour les liquides).

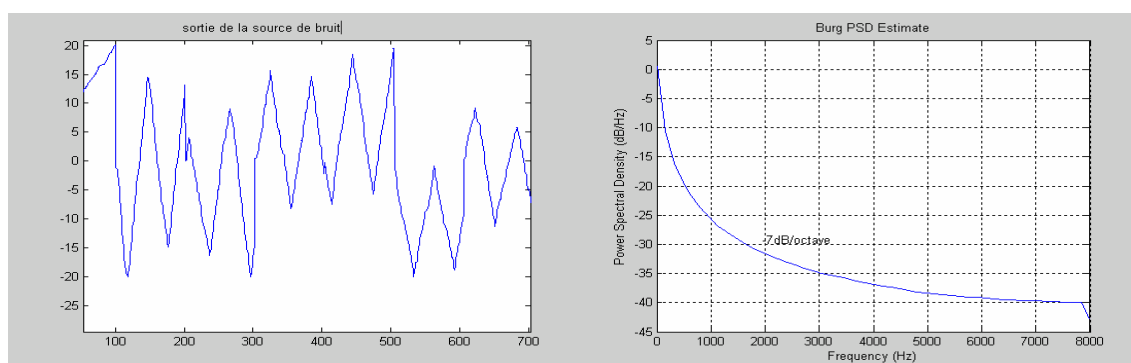


Figure 4. 3 : source de friction (non voisée) et son spectre simulés (pour les liquides).

e) Matrice de paramètres de la source pour la simulation des fricatives - fricatives sonores :

Sur le tableau 4.6 sont présentées les valeurs des paramètres de commande de la source pour la simulation des fricatives sonnantes. Les graphes correspondants aux figures 4.9, ont été tracés à partir de ce tableau.

Av	Fgp	Bgp	Fgz	Bgz	Avz	Fgs	Bgs	Avs	Af	Fnp	Bnp	Fnz	Bnz	Anp	Anz	Bf
0	0	500	2500	6000	20	0	500	20	0	270	90	270	90	0	0	50
0	0	500	2500	6000	20	0	500	20	0	270	90	270	90	0	0	50
0	0	500	2500	6000	20	0	500	20	0	270	90	270	90	0	0	50
0	0	500	2500	6000	20	0	500	20	0	270	90	270	90	0	0	50
0	0	500	2500	6000	20	0	500	20	0	270	90	270	90	0	0	50

Tableau 4. 6: Paramètres de commande de la source pour la simulation des fricatives sonnantes.

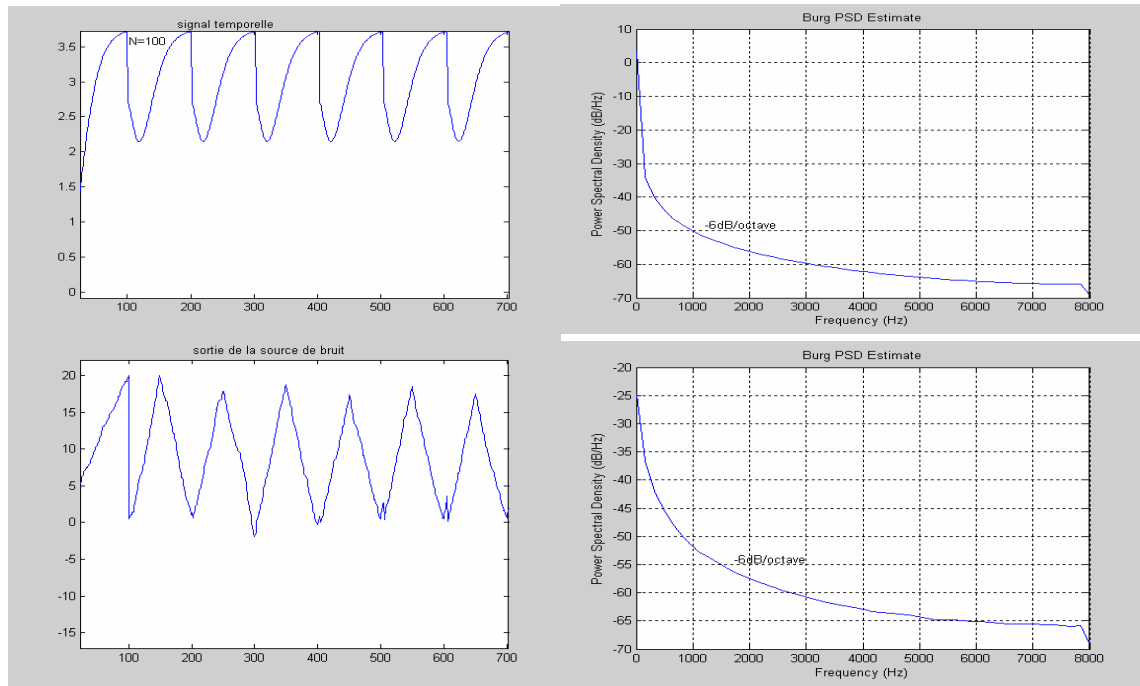


Figure 4. 4 : Signaux de la source d'excitation simulés, pour les fricatives sonnantes (en haut : source de voisement et son spectre, en bas : source de friction et son spectre).

f) Matrice de paramètres de la source pour la simulation des fricatives sourdes

Les paramètres du signal d'excitation pour la génération des fricatives sourdes, sont présentés par le tableau (4.7) :

Av	Fgp	Bgp	Fgz	Bgz	Avz	Fgs	Bgs	Avs	Af	Fnp	Bnp	Fnz	Bnz	Anp	Anz	Bf
-inf	0	500	2500	6000	-inf	0	500	-inf	60	270	90	270	90	0	0	50
-inf	0	500	2500	6000	-inf	0	500	-inf	60	270	90	270	90	0	0	50
-inf	0	500	2500	6000	-inf	0	500	-inf	60	270	90	270	90	0	0	50
-inf	0	500	2500	6000	-inf	0	500	-inf	60	270	90	270	90	0	0	50
-inf	0	500	2500	6000	-inf	0	500	-inf	60	270	90	270	90	0	0	50
-inf	0	500	2500	6000	-inf	0	500	-inf	60	270	90	270	90	0	0	50

Tableau 4. 7 : Paramètres de commande de la source pour la simulation des fricatives sourdes.

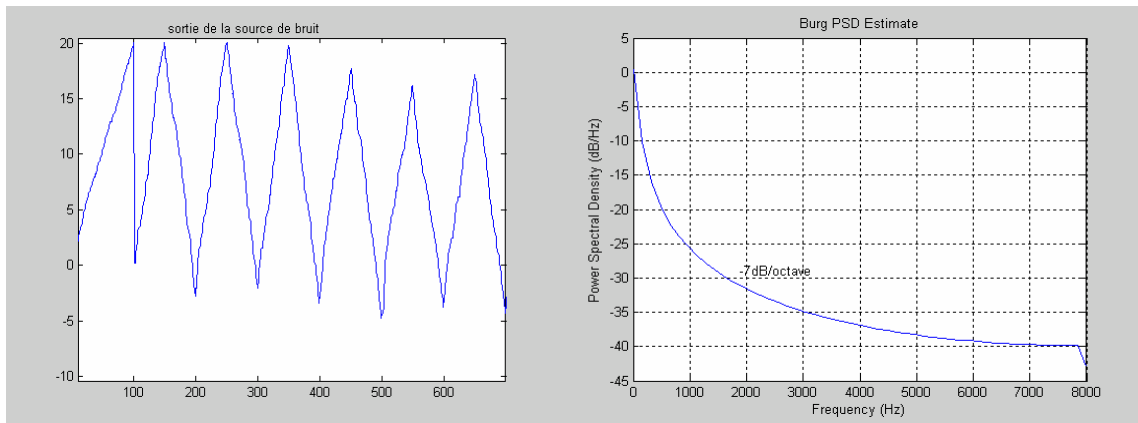


Figure 4. 5 Signaux de la source d'excitation simulés (pour les fricatives sourdes).

Récapitulatif des résultats

D'après les figures obtenues pour les sources, nous remarquons que :

Pour les spectres des sources de voisement:

- L'allure du spectre de l'onde glottique pour les sons voisés présente une pente très proche ou égale à -12dB/octave (comme imposé dans la littérature).
- Pour les voyelles, les nasales et les fricatives sonores la pente est de -11dB/octave (au lieu de -12dB/octave) ;
- pour les liquides -11.5dB/octave ;
- pour les plosives elle est de -12dB/octave .

Pour les spectres des sources de bruit :

Les résultats obtenus montrent que la pente du signal source obtenu pour certains types de sons n'est pas égale à -10dB/octave (comme imposé dans la littérature).

En effet pour :

- les fricatives sourdes et les liquides, la pente est de -7dB/octave ;
- et les fricatives sonores elle est de -6dB/octave .

Ce fait est du au réglage difficile des paramètres.

Pour les allures temporelles des sources :

Les différentes allures des signaux temporels de l'onde de voisement présentent :

- dans le cas des fricatives sonnantes, les liquides et les voyelles, une allure assimilable à un signal quasi sinusoïdal périodique.
- pour les plosives, le signal obtenu ressemble à une impulsion au début puis le reste est nul.
- Pour les nasales, l'onde de voisement présente un signal sinusoïdal périodique.

Ces allures ne ressemblent pas toujours aux allures des sources d'un signal réel. Cependant, les signaux de paroles produits à l'aide de ces modèles de sources sont tout à fait intelligibles.

L'utilité de cet ensemble de matrices de sources proposé est que lorsque nous voulons générer un son, nous prenons la matrice correspondante à la source (qui lui est fixe) et nous lui ajoutons les formants et la fréquence fondamentale (qui sont variables). Ce qui nous simplifie énormément la génération de la matrice de commande.

Cette étude a conduit à la mise en œuvre de la commande de synthétiseur. Pour ce faire, nous avons construit des tables de paramètres qui ont permis de piloter le synthétiseur pour la production de différents types de sons, sur des durées assez importantes.

A l'aide de ces matrices de sources, nous avons pu générer des phrases entières contenant tous les types de sons, après avoir mis au point la matrice de commande adéquate. La qualité n'est pas excellente mais très intelligible. Notons que le but n'est pas de produire un son de qualité mais intelligible, car pour la variation de débit, l'intelligibilité suffit.

Deux exemples de matrices de commande sont illustrés par les tableaux 4.8 (cas théorique) et 4.9 (cas réel)

- La matrice 4.8 permet la génération des trois voyelles de l'arabe: /a/, /i/, /u/.
- La matrice 4.9, permet la génération de la syllabe /ba/ réelle

La matrice du tableau 4.8 a été construite pour des sons qui ne sont pas naturels, c'est pour cela que les valeurs de tous les paramètres sont constantes. Ce qui donne un son impersonnel et robotisé.

Par contre les paramètres de la matrice 4.9 sont variables d'une ligne à une autre. La fréquence F0 est aussi variable car le son est réel.

Chaque ligne de la matrice fournira au synthétiseur, les paramètres des différents filtres le constituant et permettra de produire une portion de parole d'une durée égale à Δt .

La durée du signal de parole obtenue sera égale :

$$\text{Durée du signal} = \Delta t * (\text{nombre de ligne de la matrice})$$

La durée Δt ne doit pas dépasser 5ms car d'après la littérature, les formants n'ont pas le temps de varier sur cette durée.

Donc, pour fixer Δt et reproduire un son à l'aide du synthétiseur, il faut d'abord mesurer la durée de la séquence à produire à l'aide du SAF et déduire le nombre de lignes pour un Δt que nous choisirons inférieur ou égal à 5ms. Autrement la durée de la séquence reproduite sera allongée ou rétrécie par rapport à la durée normale réelle du signal reproduit.

Ce fait est intéressant pour la diminution ou augmentation de la durée de la séquence de parole à produire, si l'on veut introduire le paramètre variation de débit d'élocution dans la commande mais il faudrait tenir compte aussi d'autres facteurs que nous allons voir plus loin.

Après avoir mis en œuvre le SAF et sa commande, nous allons nous intéresser à l'étude de la production de la parole à débit variable à l'aide de ce modèle.

F1	B1	A1	F2	B2	A2	F3	B3	A3	F4	B4	A4	F5	B5	A5	F0	Av	Fgp	Bgp	Fgz	Bgz	Avz	Fgs	Bgs	Avs	Af	Fnp	Bnp	Fuz	Buz	Anp	Anz	Bf	SW
210	14	74	2394	698	31	3478	464	35	4147	555	33	5236	421	32	0	-inf	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
214	12	74	1207	452	41	2555	337	37	3637	443	37	5331	535	31	0	-inf	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
251	89	82	1204	161	65	2767	325	56	3530	395	58	4047	455	54	0	-inf	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
471	165	86	1151	106	80	2784	162	72	3492	256	75	3949	248	74	0	-inf	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
545	148	92	1136	121	87	2809	139	80	3554	260	80	4010	286	77	0	-inf	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
809	62	101	1383	92	94	2877	195	86	3523	234	90	3914	223	88	0	-inf	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
813	53	102	1386	57	98	2887	195	87	3513	264	90	3934	207	89	277	30	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
823	67	101	1386	28	104	2894	344	85	3368	355	89	3890	121	94	277	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
826	67	101	1385	24	105	2923	322	87	3398	435	88	3898	149	93	277	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
828	80	100	1384	28	106	2943	270	88	3493	503	89	3899	242	91	277	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
907	168	96	1386	50	106	2983	309	89	3622	170	97	4067	292	88	279	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
899	226	94	1368	76	105	2951	247	90	3570	230	94	3957	287	89	271	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
887	232	94	1361	87	104	2960	225	91	3562	229	94	3965	198	92	271	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
901	240	93	1359	90	103	2979	251	92	3591	252	95	3979	236	91	271	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
882	329	91	1324	188	100	2960	200	89	3614	299	90	3932	356	85	270	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
923	335	92	1344	183	99	2973	218	89	3706	330	90	3891	623	86	270	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
876	331	90	1319	183	99	2940	179	89	3660	300	89	3930	366	85	267	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
880	346	90	1316	202	99	2931	178	89	3646	255	89	3970	416	82	267	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
908	363	92	1322	235	98	2932	176	89	3637	208	90	4023	371	83	265	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
908	396	91	1299	297	97	1469	676	91	2924	179	88	3609	233	89	265	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
942	383	92	1323	269	97	2935	187	88	3635	196	90	4039	453	80	265	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
942	366	92	1334	248	98	2944	206	88	3650	202	91	3968	459	84	265	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
925	377	92	1317	272	97	2949	210	87	3654	200	91	3917	519	82	265	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
933	379	92	1309	279	97	2949	215	87	3639	198	91	3930	482	82	265	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
965	371	92	1333	316	96	2944	226	87	3635	201	91	3973	434	82	264	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
947	416	91	1287	364	96	1468	661	91	2942	205	87	3634	183	91	264	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
973	396	93	1336	389	96	2940	224	86	3633	195	90	3974	567	79	262	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
997	353	94	1402	406	94	2937	252	85	3590	232	90	3927	362	83	262	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
998	381	94	1162	617	95	1440	405	92	2928	231	86	3563	207	90	261	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
1024	272	94	1399	307	94	2945	253	85	3610	223	90	3891	529	82	261	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
997	311	94	1345	260	94	2941	203	86	3602	148	92	3986	467	80	262	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
997	275	94	1347	201	96	2941	224	85	3617	212	90	3915	501	81	263	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
1000	283	94	1344	276	93	2938	212	85	3557	194	90	3934	212	85	262	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
1003	235	94	1334	218	95	2952	228	85	3575	181	91	3946	268	85	261	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
985	249	91	1309	188	93	2997	225	83	3567	139	92	3942	157	87	260	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
1021	229	91	1337	143	94	2911	268	76	3552	225	81	3974	166	78	265	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
999	220	92	1323	152	94	2945	213	81	3561	180	87	3952	160	85	265	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
1001	212	92	1328	149	94	2964	200	81	3604	167	87	3970	166	83	267	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
1009	225	90	1348	101	93	2929	151	80	3540	241	79	3994	220	75	270	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
1012	233	90	1355	102	93	2951	135	80	3568	241	79	3959	219	77	271	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
776	574	80	1065	233	90	1404	139	90	2941	424	68	3536	334	73	279	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
751	438	80	1083	209	89	1424	145	87	2952	413	67	3569	260	74	279	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
744	337	81	1097	196	89	1436	141	87	3102	554	66	3628	337	74	283	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
897	267	80	1301	201	80	2992	172	71	3783	156	76	3958	685	69	293	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
885	204	79	1342	212	77	3024	247	67	3838	102	74	5201	644	47	297	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
882	119	80	1466	175	73	2952	150	66	3830	116	70	5256	683	44	297	60	0	500	2500	6000	60	0	500	80	-inf	270	90	270	90	0	0	50	1
904	119	79	1455	238	70	2943	181	64	3728	290	62	4259	651	54	305	60	0	500	2500	6000	60	0	500	80	-								

4.3 Etude de la production de la parole à débit variable par le SAF

Comme nous l'avons précisé plus haut, notre souci dans cette étude est la génération de parole à débit variable par le biais d'un modèle.

Pour ce faire, il faudrait dégager les caractéristiques de la variation de débit dans un signal de parole. Ce qui nous amène à l'étude que nous avons présentée dans le chapitre 3.

Cette dernière a montré que la variation de débit affecte plusieurs paramètres du signal de parole:

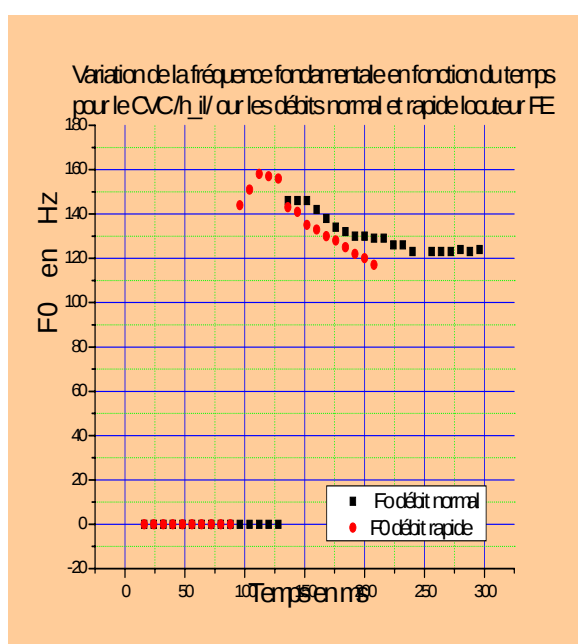
- le premier est évidemment le facteur temps qui sera soit allongé soit rétréci, selon que nous serons en diminution ou augmentation de débit par rapport au débit normal d'élocution. Cette variation du temps n'est pas linéaire. La distribution du temps dans une phrase est spécifique du locuteur (liée à ses origines sociogéographiques). La durée d'élocution par contre peut être commandée.
- Le deuxième paramètre important est le facteur fréquentiel. En effet lorsque nous faisons varier notre vitesse d'élocution, les sons changent tout en restant intelligibles. C'est l'espace vocalique qui varie. L'étude précédente (chapitre 3) a montré une réduction vocalique avec l'augmentation de la vitesse d'élocution pour tous les locuteurs. La seule différence est la centralisation. Tous les locuteurs ne centralisent pas.
- Le troisième paramètre que nous n'avons pas traité dans le chapitre 3 est la fréquence fondamentale. Cette dernière varie obligatoirement car la variation de débit affecte l'intonation (Savariaux (1996) [12]). Donc la mélodie est affectée

4.3.1 Analyse fréquentielle d'une séquence de parole à débit variable

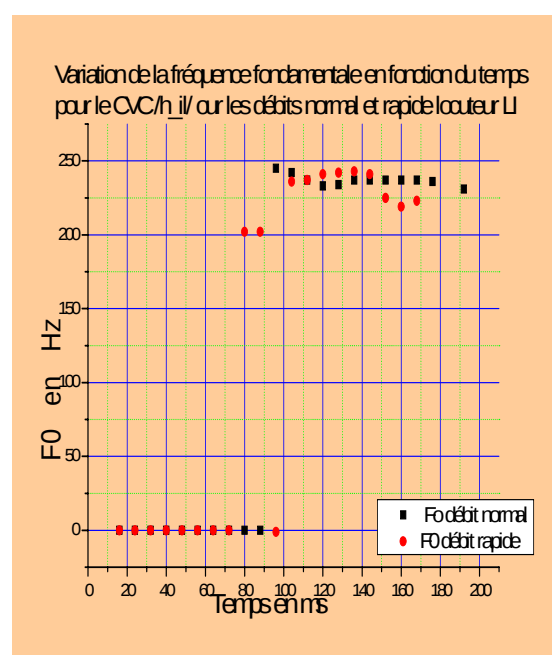
Nous allons observer ici, l'exemple d'une séquence de parole produite par quatre de nos locuteurs dans deux débits différents: le débit normal et le débit rapide.

Nous avons relevé les mélodies de nos locuteurs sur une unité CVC (/h_il/) de parole en figure 4.6.(a,b,c,d)

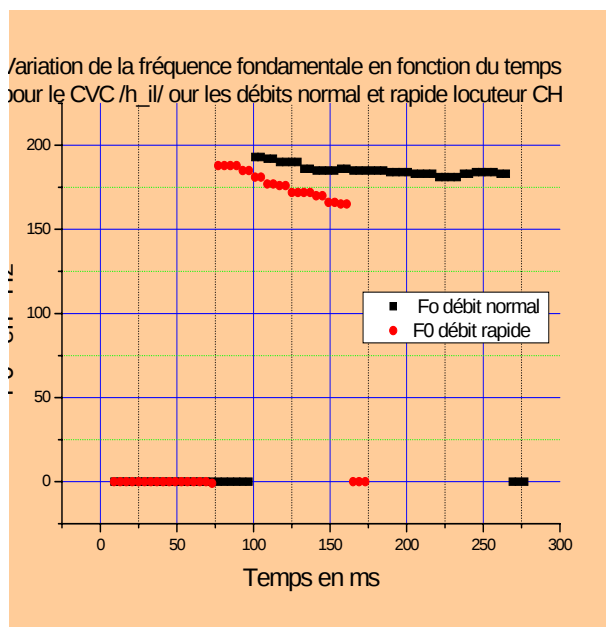
a) Locuteur FE



b) Locuteur LI



c) Locuteur CH



d) Locuteur SA

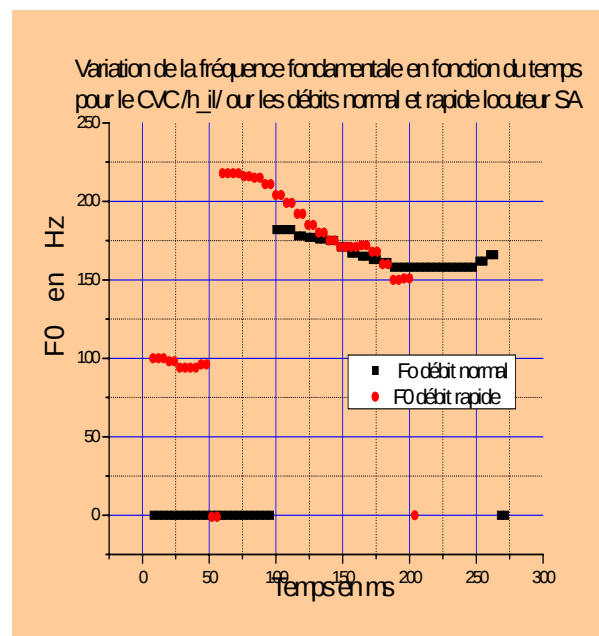


Figure 4.6 .a), b), c), d). Mélodie $F_0(t)$ pour les quatre locuteurs

Les figures 4.6 montrent que :

- l'allure de la mélodie varie avec le débit d'élocution
- le sens de variation de cette mélodie dépend du locuteur.

4.3.2 Production de parole à débit variable par le SAF

Nous voyons à travers cet exemple que le débit affecte les paramètres acoustiques du signal de parole. Cependant la loi de variation de ces paramètres est spécifique du locuteur.

a) au plan temporel

La durée de la séquence à produire « Ds' » peut être réalisée sur la matrice de commande en fixant un nouveau nombre de lignes « Nl' » après avoir fixé le temps « Tr » de rafraichissement des paramètres.

Nous pouvons aussi intervenir sur la durée « Tr ».

la durée du signal à produire à débit variable serait:

$$Ds' = (Nl' \times Tr) = [Nl \times (Tr \pm \Delta t)]$$

Nl: durée de lignes à débit normal

Tr: ne doit pas dépasser 10ms (durée maximale pendant laquelle les formants n'ont pas le temps de varier)

- **b) Au plan fréquentiel:**

- **Pour les formants:**

Il s'agirait de réaliser le modèle de centralisation des formants par le SAF:

En partant des formants en débit normal:

- la centralisation du /i/ conduirait vers une diminution du F2 et une augmentation du F1 d'où $F2 - \Delta F2$; $F1 + \Delta F1$

- La centralisation de la voyelle /u/ conduirait vers une augmentation de F2 et de F1 :

$$D'où : F2 + \Delta F2 ; F1 + \Delta F1$$

- **Pour la fréquence fondamentale F0**

Nous avons remarqué que la mélodie peut augmenter ou diminuer selon le locuteur. C'est en fait la plus difficile à réaliser, car c'est elle qui donne l'intonation. Il faudrait générer une mélodie qui ait l'air d'être naturelle.

Nous avons rappelé plus haut qu'une variation de la pression sub-glottique pourrait être à l'origine. Ce fait est d'ailleurs constaté en variation de vitesse d'élocution. Il serait donc intéressant de générer une nouvelle mélodie en utilisant un modèle des cordes vocales qui intègre la pression sub-glottique. Le modèle à deux masses a justement été réalisé dans cet

esprit. Il permet de modéliser une onde de débit glottique à partir de la variation de la pression sub-glottique de l'air venant des poumons.

Nous avons donc pensé qu'il serait judicieux de remplacer le modèle de source de Klatt par celui d'ISHIZAKA.

Le travail qui va suivre est fait dans ce sens. Il va consister en particulier à déterminer une dynamique de F_0 par le biais d'un modèle des cordes vocales: le modèle à deux masses. Le modèle à deux masses est connu pour fournir une onde glottique très réaliste. Si nous augmentons ou diminuons la pression glottique, nous pouvons trouver une variation de F_0 différente pour un même locuteur (le locuteur étant entre autre caractérisé dans ce modèle, par la masse de ses cordes vocales, qui donnera une gamme de F_0 : homme, femme, enfant)

L'onde de débit glottique $U_g(t)$ de période T_0 , générée par le modèle à deux masses, viendra exciter la batterie de filtres simulant le conduit vocal. Il suffit alors de remplacer dans les équations des filtres (4.1) $x(t)$ par $U_g(t)$

Les valeurs de F_0 correspondantes à cette nouvelle source seront reportées dans la matrice de commande du SAF.

4.4 Etude du modèle à deux masses

Le modèle à deux masses est un modèle mécanique des cordes vocales. Il modélise l'écoulement de l'air à travers la glotte.

4-4-1- Description du modèle à deux masses

Le schéma du modèle est donné sur la figure 4.9. On peut voir, à gauche, la trachée qui amène l'air des poumons. A droite, se trouve le tube du larynx qui réalise la liaison glotte – conduit vocal. Ces deux tubes, trachée et larynx, sont supposés cylindriques et de dimensions invariables.

Les cordes vocales elles-mêmes sont divisées en une partie supérieure et une partie inférieure. Chacune des parties est représentée par un ensemble mécanique comprenant une masse m , d'épaisseur d , un ressort de tension s et un amortissement r [8]. Les deux masses peuvent se déplacer latéralement, d'une valeur notée x . Elles sont reliées par un ressort de raideur k_{12} . Le déplacement de chacune des parties ne se fait donc que dans un seul plan, perpendiculairement à l'axe du système, à partir d'une position de repos (notée x_0). Les différents facteurs notés sur la figure 4.9 sont :

l_g : la longueur effective des cordes vocales,

d_1 et d_2 : l'épaisseur des deux parties d'une corde vocale,

s_1 et s_2 : les tensions équivalentes des deux parties,

r_1 et r_2 : les résistances de viscosité équivalentes,

Ag_{01} et Ag_{02} : les sections transversales de la glotte lorsque les masses m_1 et m_2 sont au repos (aire de phonation neutre),

U_g : le débit d'air à travers la glotte,

P_s : la pression sub-glottique par rapport à la pression atmosphérique.

P_{CV} : la pression à l'entrée du conduit vocal.

Compte tenu de la symétrie du modèle, les aires d'ouverture de la glotte pour les deux masses sont :

$$A_{g1} = A_{g01} + 2l_g x_1 = 2l_g (x_{01} + x_1) \quad (4.8)$$

$$A_{g2} = A_{g02} + 2l_g x_2 = 2l_g (x_{02} + x_2) \quad (4.9)$$

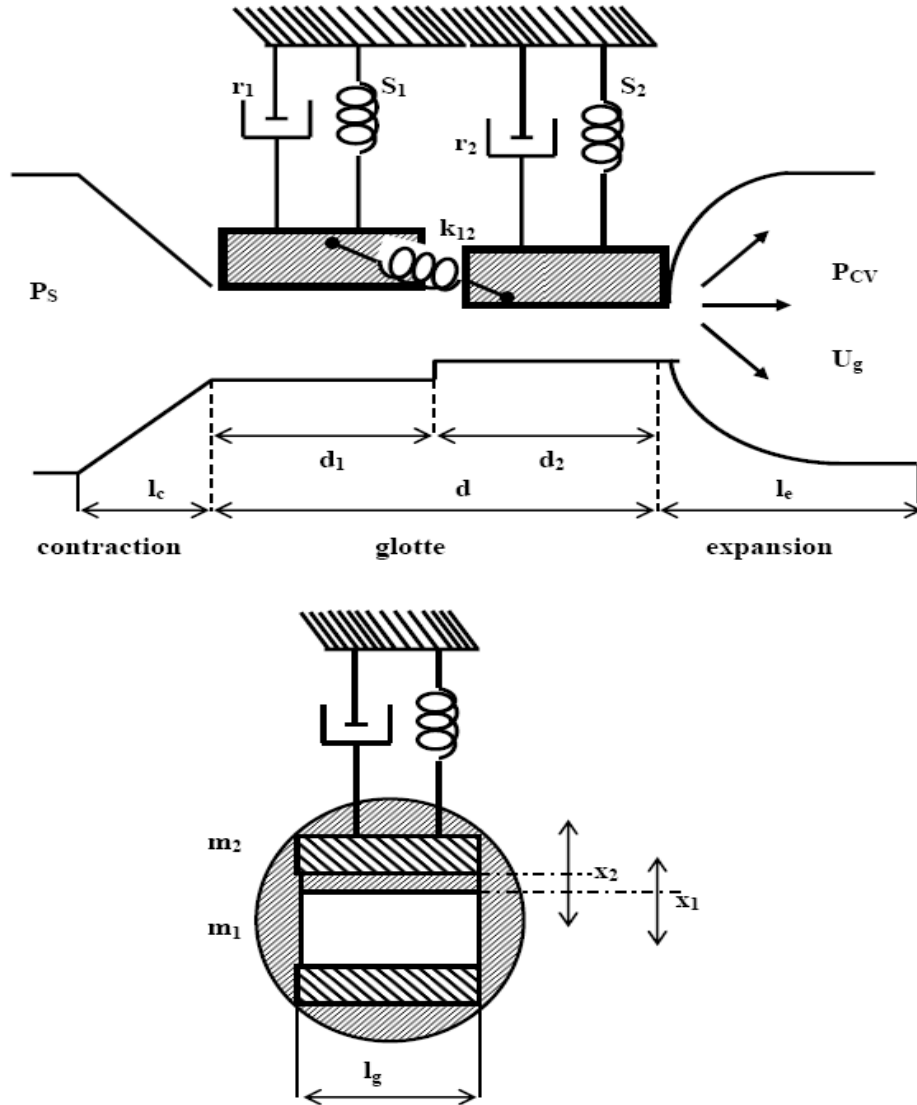


Figure 4. 7 Modèle à deux masses des cordes vocales.

$$m_1 \frac{d^2 x_1}{dt^2} + r_1 \frac{dx_1}{dt} + s_1(x_1, x_{01}) + k_{12}(x_1 - x_2) = l_g d_1 P_{m1} \quad (4.10)$$

$$m_2 \frac{d^2 x_2}{dt^2} + r_2 \frac{dx_2}{dt} + s_2(x_2, x_{02}) + k_{12}(x_2 - x_1) = l_g d_2 P_{m2} \quad (4.11)$$

P_{m1} et P_{m2} sont les pressions agissant sur chacune des masses.

$s_1(x_1, x_{01})$ et $s_2(x_2, x_{02})$ sont les tensions équivalentes dont les relations non linéaires, sont données par:

$$s_i(x_i, x_{0i}) = k_i(x_i - x_{0i})[1 + \eta_{k_i}(x_i - x_{0i})^2] + f_{c_i} \quad (4.12)$$

Avec $f_{c_i} = 0$ si $x_i \succ -x_{0i}$

$$\text{et } f_{c_i} = k_i x_i (1 + \eta_{h_i} x_i^2) \quad \text{si } x_i \prec -x_{0i} \quad \text{pour } i = 1, 2. \quad (4.13)$$

η_{k_i} et η_{h_i} sont des coefficients de non linéarité, le terme f_{c_i} rend compte de relations supplémentaires correspondant à la mise en contact des deux cordes vocales au cours de la phase d'occlusion totale.

Pour nos différentes simulations nous avons retenu les valeurs numériques suivantes. Ces valeurs sont tirées essentiellement de [4.8, 4.12, 4.13].

$$h_i = 3k_i; \quad \eta_{k_i} = 100; \quad \eta_{h_i} = 500; \quad r_i = 2\varepsilon_i \sqrt{m_i k_i} \quad \text{pour } i = 1, 2$$

$$10 \leq k_1 \leq 80; \quad 100 \leq m_1 \leq 380 \text{mg}; \quad 10 \leq m_2 \leq 70 \text{mg}$$

$$2 \leq k_2 \leq 10; \quad l_g = 1.4 \text{cm}; \quad d_1 = 0.25 \text{cm}; \quad d_2 = 0.05 \text{cm}; \quad 6.5 \leq k_{12} \leq 30$$

$$\varepsilon_1 = 0.1 \text{ et } \varepsilon_2 = 0.6 \text{ pour une glotte ouverte; } \varepsilon_1 = 1.1 \text{ et } \varepsilon_2 = 1.6 \text{ pour une glotte fermée.}$$

La résolution du système d'équations différentielles donne les valeurs de déplacement x des masses. On peut alors en déduire l'aire de la glotte:

$$A_{g_i}(t) = A_{g_i} + 2l_g x_i(t) = 2l_g (x_i(t) + x_{0i}) \quad i = 1, 2 \quad (4.14)$$

Dans l'analogie mécanique-électrique, la glotte peut être représentée par un circuit électrique composé de résistances et d'inductances. Le circuit équivalent obtenu est donné dans la figure 4.10 [3]. Le conduit vocal est représenté par son impédance d'entrée réalisée à l'aide de circuits résonants accordés sur les fréquences des deux premiers formants. Dans cet article, nous ne nous intéressons pas au couplage avec le conduit vocal. Aussi, nous avons adopté, pour la suite, un conduit vocal uniforme de section 2cm^2 . Les valeurs des différents éléments sont données par [4] :

$$R_{k1} = 1.37 \frac{\rho}{2} \frac{|U_g|}{A_{g1}^2}; \quad R_{V1} = \frac{12\mu}{A_{g1}^3} d_1 l_g^2; \quad L_{g1} = \frac{\rho}{A_{g1}} d_1;$$

$$R_{k2} = -\frac{\rho}{A_{g2} A_1} \left(1 - \frac{A_{g2}}{A_1}\right) |U_g|; \quad R_{V2} = \frac{12\mu}{A_{g2}^3} d_2 l_g^2; \quad L_{g2} = \frac{\rho}{A_{g2}} d_2$$

$$R_{k12} = \frac{\rho}{2} \left(\frac{1}{A_{g1}^2} - \frac{1}{A_{g2}^2} \right);$$

Où ρ : densité de l'air ($1.14 \cdot 10^{-3} \text{ g/cm}^3$) ; μ : coefficient de viscosité de l'air ($1.86 \cdot 10^{-4} \text{ dyne/cm}^2$) ; A_1 : aire de la première section du conduit vocal (2 cm^2) ;
 U_g : débit de l'air à travers la glotte

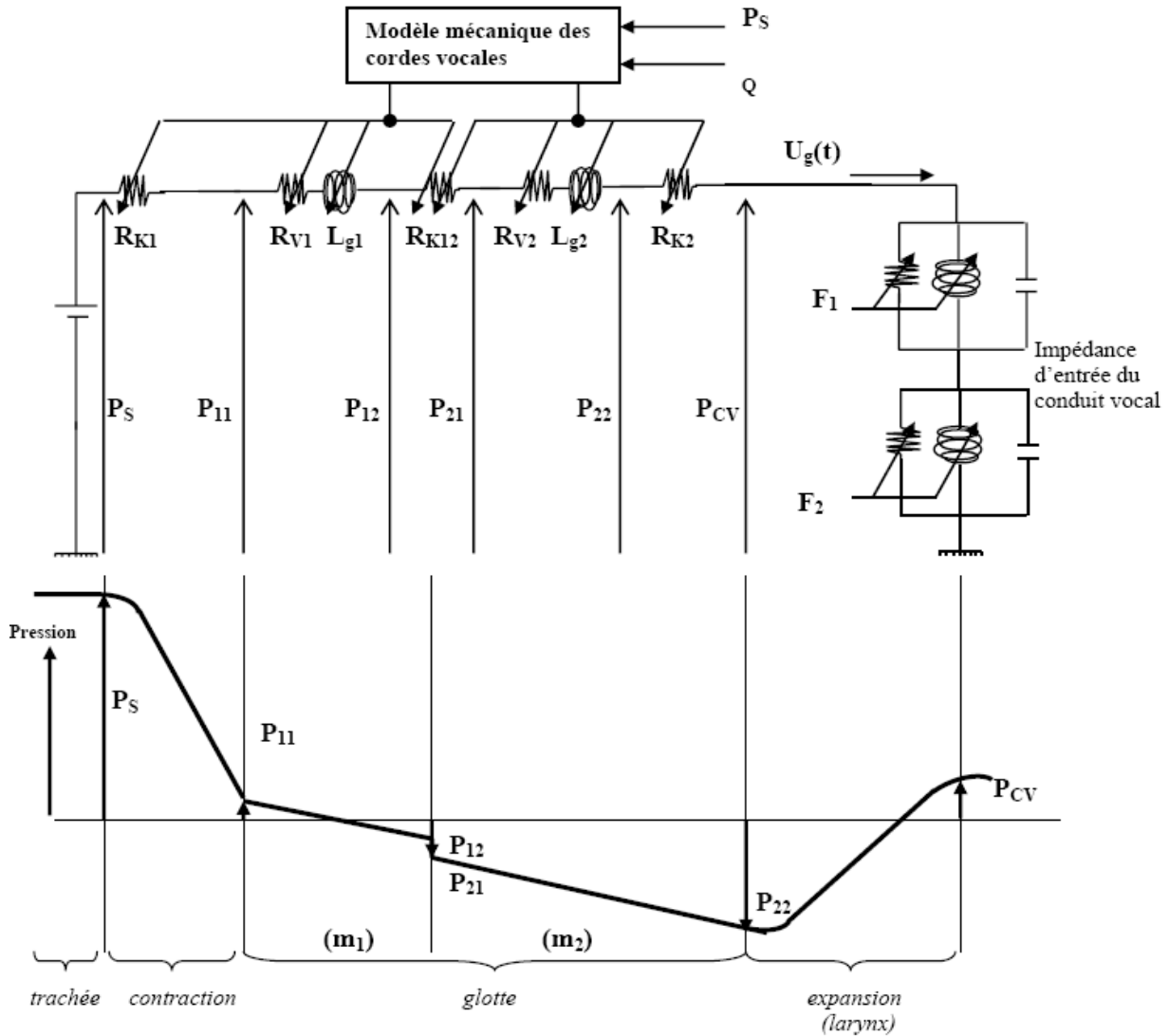


Figure 4. 8. Modèle électrique équivalent de la source vocale à deux masses couplé au conduit vocal et représentation en concordance de la distribution de la pression dans la glotte d'après Guérin.

Sachant que l'équation différentielle qui lie les évolutions temporelles du débit $U_g(t)$, des parties réelles et imaginaires de l'impédance de la glotte et de la pression intra-glottique $P_i(t)$ peut être écrite de la forme :

$$U_g(t) \cdot R_g(t) + \frac{d}{dt}(L_g(t) \cdot U_g(t)) = P_i(t) \quad (4.15)$$

Alors, à l'aide de la figure 4.10, on peut déduire la pression sur chacune des masses et l'équation différentielle donnant $U_g(t)$:

$$P_{m1} = P_S - R_{k1}U_g - \frac{R_{V1}U_g + L_{g1} \frac{dU_g}{dt}}{2} \quad (4.16)$$

$$P_{m2} = P_{CV} + R_{k2}U_g + \frac{R_{V2}U_g + L_{g2} \frac{dU_g}{dt}}{2} \quad (4.17)$$

$$P_S - P_{CV} = (R_{K1} + R_{V1} + R_{K12} + R_{V2} + R_{K2})U_g + (L_{g1} + L_{g2}) \frac{dU_g}{dt} \quad (4.18)$$

Dans ces relations, nous négligeons le terme $U_g \left(\frac{dL_g}{dt} \right)$ qui proviendrait de $\frac{d}{dt}(L_g U_g)$ car nous considérons que l'aire de la glotte ne varie que très lentement.

Pour résoudre le système d'équations mécaniques (équations 4.10 et 4.11), nous avons utilisé la méthode de RUNGE-KUTTA [9, 1] à quatre approximations, le système d'équations étant transformé en quatre équations de premier ordre.

Le calcul s'effectue de la manière suivante : on se donne, au départ, les aires de repos A_{q01} et A_{q02} et on calcule le premier échantillon de $U_q(t)$. Cette valeur permet de calculer P_{m1} et P_{m2} , puis de résoudre le système d'équations mécanique. On dispose alors de $A_{q1}(t)$ et $A_{q2}(t)$ qui permettront d'évaluer un nouvel échantillon de $U_q(t)$, ce qui correspond à la première étape décrite dans ce paragraphe. Le cycle décrit peut alors recommencer. Dans le cadre de ce travail, qui se limite à une étude sur les voyelles orales du Français, la largeur spectrale intéressante est égale à 5Khz, nous prendrons ainsi une fréquence d'échantillonnage de 10 kHz.

Vu le nombre important de paramètres de contrôle du modèle à deux masses, nous proposons dans cette étude, une contribution à la réduction de ceux ci.

4-4-4- Détermination des paramètres de commande du modèle à deux masses

Comme le soulignent Sciamarella et D'alessandro en 2004 [11], l'absence d'une relation simple entre les paramètres de contrôle du modèle et la physiologie de la glotte est une des faiblesses principales des modèles de production vocale. Même si la plupart des paramètres de contrôle sont choisis d'après des mesures physiologiques, ces paramètres doivent néanmoins être réglés pour compenser les simplifications du modèle. La tâche n'est pas simple, d'abord parce que les paramètres de contrôle l'emportent largement en nombre sur les paramètres acoustiques, et deuxièmement parce que la corrélation entre les deux types de paramètres n'a pas été élucidée.

La démarche adoptée par la plupart des utilisateurs du modèle à deux masses consiste à étudier l'évolution des grandeurs de sortie, telles que la fréquence fondamentale d'oscillation, la puissance de l'onde de débit, le quotient d'ouverture des cordes vocales (paramètres acoustiques), en fonction de certains paramètres (physiques) clés qui ont été jugés très significatifs dans la simulation du mouvement des cordes vocales [5]. Les six paramètres de commande concernés sont les suivants : P_s , m_1 , m_2 , k_1 , k_2 , k_{12} , c'est à dire : la pression sous-

glottique, les deux masses et les raideurs. Les aires initiales A_{q0i} ont été fixées à 0.05 cm^2 et les résistances de viscosité identiques à celles précisées par Ishizaka et Flanagan [6].

Pour obtenir des résultats significatifs, il faudrait donner à chaque paramètre une dizaine de valeurs différentes, ce qui conduirait à un nombre de simulations excessif. Guérin [4] a fait varier simultanément les six paramètres. Ce qui a conduit à 200 simulations environs. Les résultats obtenus sont regroupés sur le tableau 4.11.

Ps	m₁	m₂	k₁	k₂	k₁₂
(cm H₂O)	(mg)	(mg)	(kdynes/cm)	(kdynes/cm)	(kdynes/cm)
2-12	100-380	10-70	10-80	2-10	6.5-30

Tableau 4. 10: Domaine d'évolution des paramètres de commande [4].

Une étroite dépendance entre les masses et les raideurs a été constatée. Les régressions suivantes ont été obtenues en fonction de m_1 .

$$\frac{m_2}{m_1} = \frac{k_2}{k_1} = 0.17 \quad k_1 = -0.10m_1 + 57 \quad k_2 = -1.8 \cdot 10^{-2} m_1 + 9.9$$

$$k_{12} = -3.7 \cdot 10^{-2} m_1 + 27$$

A partir de ces relations, nous proposons dans cette étude un paramètre de contrôle unique à partir duquel seront déterminés tous les autres, mis à part Ps qui est indépendant. A cet effet, la masse m_1 est représentée par le rapport entre une masse globale unique et un paramètre masse/tension noté Q , tel que :

$$\begin{cases} m_1 = \frac{M}{Q} \\ k_1 = k \cdot Q \end{cases} \quad (4.19)$$

avec comme valeurs de bases

$$\begin{cases} M = 380 \text{ mg} \\ k = 25 \text{ Kdynes / cm} \end{cases} \quad k_{12} = (10 + 5 \cdot Q) \text{ Kdynes / cm} \quad \frac{m_2}{m_1} = \frac{k_2}{k_1} = 0,17$$

Ces valeurs sont inspirées des travaux de Guérin [1], Ishizaka et Flanagan [4].

Nous voyons que si le paramètre Q augmente, les deux masses diminuent et les raideurs croissent. Il s'agit donc d'un paramètre qui rend compte de l'étirement et de l'effilement des cordes vocales ; c'est-à-dire pour l'essentiel de l'effet observé sur les cordes vocales réelles. Sur la figure 4.11, nous donnons trois exemples de forme d'onde $U_q(t)$ qui n'ont pas été retenues lors de la recherche des valeurs des paramètres de commande (cas A, B et C), ainsi

qu'un exemple de cas de forme d'onde régulière (cas D). Pour chacun des cas, nous donnons la valeur des paramètres variables : m_2/m_1 , k_{12} , P_S et Q .

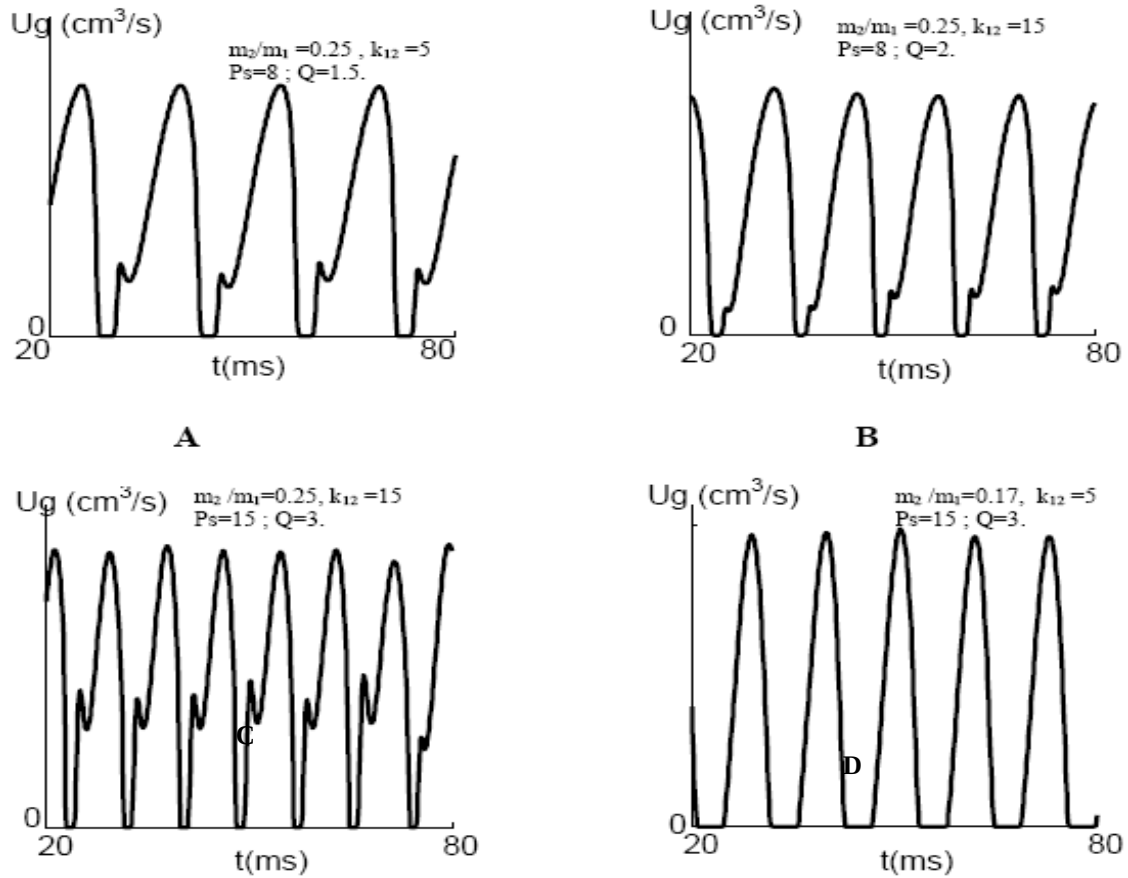


Figure 4. 9 : Exemples de forme d'onde $U_g(t)$.

Nos simulations ont permis dans un premier temps de vérifier les plages de variations données sur le tableau 4.11. Dans ce qui suit, nous allons décrire un certain nombre de simulations qui ont permis de fixer les plages de variations des deux paramètres de contrôle P_S et Q , correspondant à des formes d'ondes régulières.

4-4-5- Variation de la fréquence fondamentale de vibration des cordes vocales en fonction des paramètres Pression et facteur masse/tension

Puisque nous n'avons pas de renseignements assez précis sur la dynamique du paramètre Q , nous proposons dans cette étude une méthodologie permettant de fixer la dynamique de ce dernier. Nous nous plaçons maintenant dans le cadre du système de commande défini au paragraphe 4.4.4 ; c'est à dire $\{Q, P_S\}$. Nous avons alors opéré à partir de 68 simulations avec les valeurs suivantes : $P_S = 3$ à $19 \text{ cm H}_2\text{O}$ et $Q=1.5$ à 3 . Nous donnons sur la figure 4.12 comment se fait la mesure de la fréquence fondamentale de vibration des cordes vocales $F_0=1/T_0$.

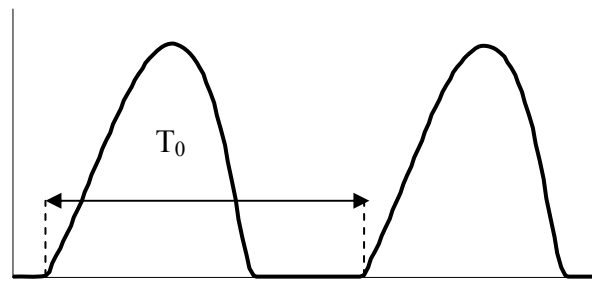


Figure 4. 10 : onde glottique et mesure de F_0 .

Sur les résultats de ces 68 simulations, nous avons effectué une analyse des régressions entre les commandes et la variable de sortie F_0 . Le tableau 4.12 donne les relations obtenues et la figure 4.13 donne les résultats des simulations:

Q	$F_0 = f(P_s)$ en Hz
1.5	$F_0 = 2.00 P_s + 78$
2	$F_0 = 2.15 P_s + 101$
2.5	$F_0 = 2.45 P_s + 123$
3	$F_0 = 2.53 P_s + 146$

Tableau 4. 11 : Relations entre les deux paramètres de commande P_s et Q et la variable de sortie F_0 .

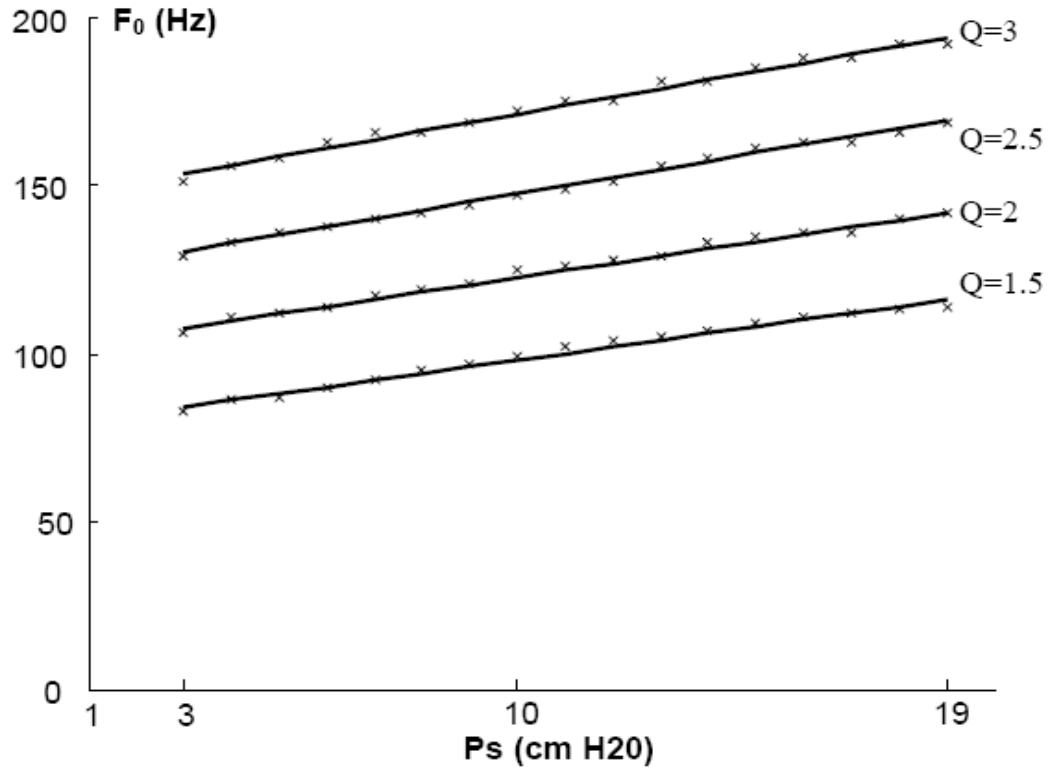


Figure 4. 11 F_0 en fonction de P_s pour différentes valeurs de Q (résultats de 68 simulations)

La relation qui lie F_0 à P_s varie selon Q , de 2 à 2.53 Hz/(cm H₂O), ce qui est très voisin des valeurs relevées en simulation [4,6] (2.5 Hz/(cm H₂O)). Les résultats que nous obtenons montrent que si les masses sont faibles (Q élevé), P_s a une influence plus sensible sur F_0 . Indirectement, il nous semble possible de proposer une dynamique et des valeurs de Q permettant d'obtenir, à l'aide de la relation $F_0 = f(P_s)$, des résultats proches de la réalité. En effet, nous avons relevé, pour 68 simulations, la relation moyenne $F_0 = f(P_s)$ qui s'écrit :

$$F_0 = 2.3 P_s + 48 Q + 1.98 \quad (4.20)$$

Si P_s est de 8 cm H₂O, avec Q variant de 1.5 à 3, F_0 couvre la gamme : 92 à 164 Hz, ce qui correspond bien aux évolutions d'une voix d'homme.

Sur la figure 4.14, nous avons porté les résultats de mesure de F_0 en Hertz en fonction de Q pour différentes valeurs de P_s . Nous remarquons que le choix des plages de variations des paramètres $\{P_s \text{ et } Q\}$ est réaliste (F_0 variable entre 80 Hz et 190 Hz).

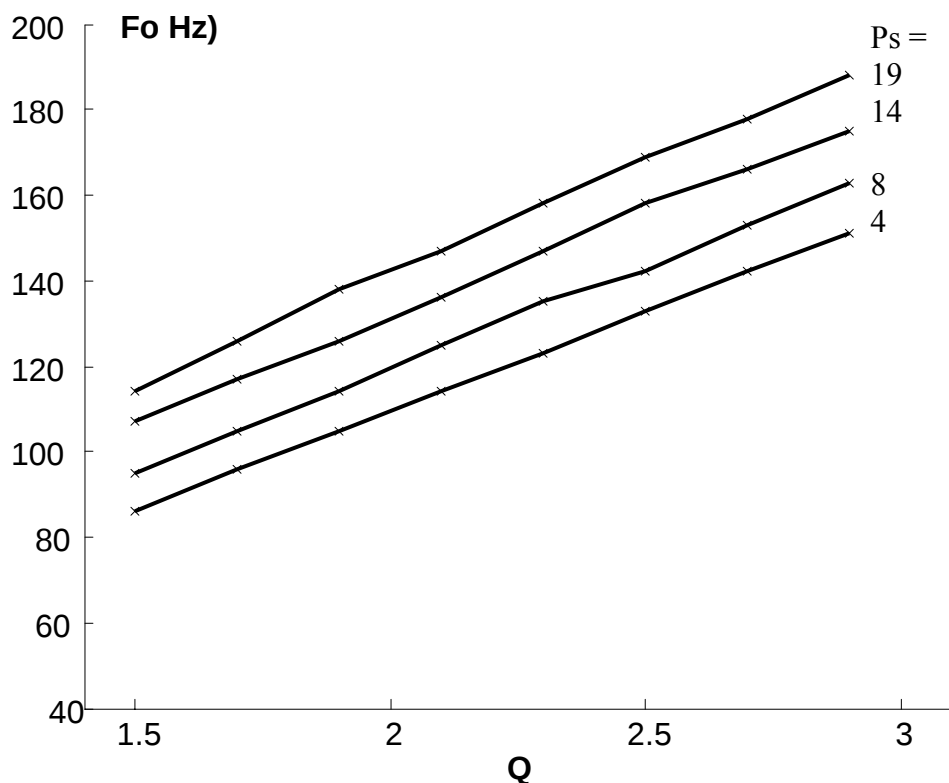


Figure 4. 12 F_0 en fonction de Q pour différentes valeurs de P_s (résultats de 32 simulations)

4. 5 CONCLUSION

Nous avons entrepris de produire de la parole à débit variable à partir d'un synthétiseur à formants de Klatt. Nous avons réalisé ce modèle et mis au point une matrice de commande à l'aide d'un ensemble de sources d'excitation pour les différents filtres servant de conduit vocal au SAF.

L'analyse d'un exemple d'une séquence de parole produite à débit variable par deux locuteurs d'origines sociogéographiques différentes, a montré des différences très notables au plan de la variation formantique et de F_0 .

- Les deux locuteurs sont concernés par la réduction vocalique mais tous les deux ne centralisent pas.

- La fréquence fondamentale est variable pour les deux locuteurs mais pas dans le même sens.

Il sera donc possible de modéliser la centralisation dans la commande du SAF.

La production d'une mélodie ayant l'air naturelle ne sera possible qu'à l'aide d'un modèle de source simulant les variations de la pression sub-glottique avec la variation de débit pour un même locuteur.

Le modèle à deux masses des cordes vocales répond à ces caractéristiques. Cependant la commande de ce dernier est rendue difficile à cause du nombre élevé de ses paramètres. Nous avons procédé à une réduction de ces derniers en utilisant un analogue électrique du modèle mécanique.

Avec le modèle que nous avons utilisé, nous avons défini un paramètre Q dont les variations peuvent être assimilées à celles de la tension passive des cordes vocales. Associé à la pression

sous-glottique P_s , on dispose de deux paramètres physiques aux gammes de variations définies.

Nous avons précisé quantitativement les relations entre l'ensemble $\{Q, P_s\}$ et la fréquence d'oscillation des cordes vocales. La lecture et l'interprétation de nos résultats montrent qu'ils correspondent bien à ceux qui ont été obtenus dans le cas de la parole naturelle [5].

Les résultats ainsi exprimés peuvent être utilisés en synthèse à partir de paramètres physiques et en analyse de la parole avec pour objectif l'identification des commandes au niveau du larynx.

Ce modèle ainsi simplifié sera utilisé pour simuler la source dans le modèle de Klatt. Il permettra de générer une onde glottique correspondant au débit de l'air à travers la glotte $U_g(t)$ qui viendra exciter l'ensemble les filtres mis en parallèle ou en série pour simuler le conduit vocal dans le SAF.

Nous avons obtenu une onde de débit glottique $U_g(t)$ en fonction du temps et une dynamique de F_0 par l'équation 4.20 qui est fonction de la pression sub-glottique P et du coefficient masse tension Q . Le réglage de Q permettra de situer la gamme de fréquence fondamentale correspondant à un locuteur spécifique. L'augmentation ou la diminution de P permettra d'obtenir une mélodie adéquate pour cette pression. Ce qui sera utilisé dans la matrice de commande du synthétiseur pour la simulation d'une variation de la vitesse d'élocution.

Nous avons ainsi proposé quelques solutions pour la possibilité d'introduire le paramètre débit dans le synthétiseur à formants de Klatt. Nous pensons que cette étude apportera un plus dans cette problématique. Les travaux en cours vont permettre de vérifier ces résultats et poursuivre le travail pour proposer d'autres solutions à ce problème car il reste encore beaucoup à faire pour arriver à égaler l'appareil phonatoire humain pour cette problématique.

Bibliographie du chapitre 4

- [1]. G. Fant, J. Liljencratts & Qi-guang Lin. "A Four-Parameter Model of Glottal Flow", French-Swedish Symposium, Grenoble, Apr. 22-24, 1985, STL-QPSR4/1985.
- [2] L. Falek, H. Teffahi, A. Djeradi. (à paraître) "Simulation d'un modèle de la source vocale et détermination des paramètres de commande". Canadian Journal of physics.
- [3] B. Guerin, M. Mrayati., R. Carré., In I.E.E.E. Int. Conf. "On Acoustics, Speech and Signal Processing". p. 47. (1976).
- [4]. B. Guerin. "Contribution aux recherches sur la production de la parole. Etude du fonctionnement de la source vocale. Simulation d'un modèle" Thèse de doctorat, Institut Nationale Polytechnique de Grenoble, France, (1978).
- [5] N. Henrich. "Etude de la source glottique en voix parlée et chantée". Thèse de doctorat de l'Université Paris 6. (2001).
- [6] K. Ishizaka., J.L. Flanagan, "Voice simulation with a body-cover model of the vocal folds" Bell.Syst.Tech. Journal., **51**, p. 1233., (1972).
- [7] Klatt.D.H "Software for a cascade/parallel formant synthesizer" Acoustical society of America, 67(3), (pp-971-995), 1980.
- [8] M. Mrayati., B. Guerin, L.J.Boe. "A two-mass model of the vocal cords: Determination of control parameters and their respective consequences". Acustica., **35**., p. 330., (1976).
- [9] X. Pelorson et al. "Optimal Glottal Configuration for Ease of Phonation", Journal. Acoust. Soc. Am, 96 (6), (1994).
- [10] D. Sciamarella., C. D'alessandro. "A study of the Two-Mass Model in terms of Acoustic parameters", International Conference on Spoken Language Processing ICSLP. p.2313, (2003)

- [11]. D. Sciamarella., C. D'alessandro. "*Modele symétrique des cordes vocales*". Acta Acustica, Vol 90, n°4, 746-761, (2004).
- [12] Savariaux, C Abry, Christian / Cathiard, Marie-Agnès / Abed, R. El / Lallouache, M. T. / Leroy, M.-C. / Perrier, Pascal / Poveda, F. / : "*Silent speech production: anticipatory behaviour for 2 out of the 3 main vowel gestures/features while pausing*", In *SPM-1996*, 101-104.
- [13]. H. Teffahi, *SIMCV*. "*Un simulateur analogue de la propagation du son dans le conduit vocal*". Canadian Journal of Physics, vol **85**, N°12. p. 1325-1341(17) (2007).

Conclusion générale

L'analyse que nous avons effectuée a porté sur la lecture du comportement des muscles articulatoires de la parole en faisant varier le débit, le contexte et le locuteur.

C'est ainsi qu'une campagne de mesure a été menée sur une population de six locuteurs, chacun a répété 10 fois les phrases du corpus.

Ce corpus contient 1620 phrases, chacune étant construite autour des trois fricatives arrière de l'arabe. Chacune de ces fricatives est mise dans trois contextes vocaliques de l'arabe : /a/, /i/, /u/. Nous avons obtenu un matériau linguistique qui nous a permis de faire le reste de notre étude

L'étape première a consisté à analyser ces différents enregistrements. Lors de cette analyse il a fallu tout d'abord segmenter les signaux manuellement et vérifier la cohérence des segments obtenus, par la suite, nous avons retenus les segments temporels nécessaires à notre analyse, tels que les durées absolues de la phrase et les durées des CVC concernées.

Nous avons aussi déterminé les valeurs relatives de ces segments par rapport à la durée de la phrase, sans oublier que nous avons répété cette opération pour trois débits d'élocution de ces phrases.

Le choix de l'espace de représentation des temps est important si l'on veut éliminer les différences liées à l'articulation en absolue.

Nous avons ensuite analysé temporellement, les segments de parole retenus, dans une condition normale d'élocution, afin de faire ressortir l'impact du lieu d'articulation des trois pharyngales étudiées sur le timing de ces segments. Les résultats obtenus ont montré un allongement des durées qui diminue avec le voisement des consonnes fricatives considérées. Le caractère voisement semble nécessiter moins d'effort. Cela pourrait s'expliquer par le fait que la production des consonnes sourdes n'exige pas que les cordes vocales soient accolées. La dépense d'air est nettement plus importante que lors de la production des sons voisés. Donc, plus d'apport d'air des poumons est nécessaire pour produire le /ħ / (/ح/) /h_ / que le /h/ (/أ/) et le /ʕ / (/ع/).

En somme nous dirons que les lieux d'articulation des pharyngales utilisées, en contextes vocaliques, semblent n'avoir aucun effet sur le timing de la parole. Nous pensons que cela est dû au fait que les lieux d'articulation de ces trois fricatives sont très proches. Les effets sont tellement faibles qu'ils sont inappréciables au plan temporel.

Une analyse temporelle et fréquentielle en contrainte de débit, des segments de parole précédents, a été conduite. Il s'agissait de comparer les différentes durées des segments de parole et les variations formantiques (F1, F2) pour les voyelles, pour six locuteurs.

L'analyse des espaces vocaliques a montré une réduction vocalique pour tous les locuteurs. Cependant certains locuteurs centralisent et d'autres pas.

L'analyse des durées relatives a montré des augmentations des durées avec la vitesse pour certains locuteurs.

Les résultats obtenus ont montré les stratégies empruntées par les locuteurs pour atteindre leurs cibles acoustiques. Ces différentes stratégies semblent être liées au caractère sociogéographique des différents locuteurs du corpus. C'est ainsi que nous avons dégagé des patrons temporels et acoustiques de production de la parole en vitesse d'élocution variable.

Dans le but d'une meilleure compréhension des phénomènes mis en jeu en vitesse d'élocution variable, nous avons entrepris la production de la parole à débit variable par le biais d'un modèle. Nous avons alors mis en œuvre un modèle acoustique de production de la parole (le synthétiseur à formants de Klatt), puis étudié et réalisé sa commande, qui nous a permis de comprendre comment intervenir sur ce modèle pour y intégrer la contrainte de débit.

En s'appuyant sur les résultats du chapitre 3 et de l'analyse d'un exemple d'une séquence de parole produite à débit variable par deux locuteurs d'origines sociogéographiques différentes, nous avons recensé des différences très notables au plan de la variation formantique et de F0. En effet, comme trouvé au chapitre 3 :

- Les deux locuteurs sont concernés par la réduction vocalique mais seul un centralise.
- La fréquence fondamentale est variable pour les deux locuteurs mais pas dans le même sens.

- Nous avons retenu un fait important et modélisable, qui est la centralisation des formants des voyelles en débit variable.
- Le problème de la variation de la fréquence fondamentale a été réglé par le remplacement du modèle de source de Klatt par un modèle des cordes vocales. Ce dernier a été choisi car il produit une onde glottique tenant compte d'un facteur intervenant dans la variation de la vitesse d'élocution: la pression sub-glottique. Après une étude sur le modèle à deux masses pour réduire dans un premier temps ses nombreux paramètres (qui rendaient sa commande très difficile), nous avons dégagé une dynamique de la fréquence fondamentale en fonction de deux paramètres importants : le facteur pression sub-glottique et le facteur masse tension. En agissant sur ces deux paramètres, nous pouvons proposer une mélodie variable pour un même locuteur.

C'est ainsi que nous intégrons le modèle à deux masses dans le modèle de Klatt. L'onde de débit glottique $U_g(t)$, produite par le modèle d'Ishizaka, viendra exciter la batterie de filtres servant de modèle de conduit vocal dans le SAF. La dynamique de F0 en fonction de la pression glottique sera intégrée dans la commande du SAF.

- Le facteur variation du temps lié au débit sera intégré dans la matrice de commande en jouant sur le nombre de ligne et la durée de parole produite par une ligne : le paramètre Δt .

Ces trois points importants intervenant dans la production de parole à débit variable par un modèle ont permis de caractériser la production de parole à débit variable.

La mise en œuvre de la production de parole à débit variable est au stade expérimental au laboratoire LCPTS.

Perspectives

- La base de données sur les articulations arrière de l'arabe, pourrait être utilisée pour améliorer les modèles articulatoires existants, qui ne prennent pas en compte ce type d'articulation (le modèle de Maeda).
- L'usage de modèles pourrait donner une meilleure compréhension des relations entre les aspects de coordination des groupes de muscles et leur commande.
- La mise en œuvre du modèle à 2 masses va nous permettre d'améliorer notablement le fonctionnement du modèle de production de Klatt, mais à condition d'élaborer la source de friction pour la production des fricatives.

BIBLIOGRAPHIE GENERALE PAR CHAPITRE

Bibliographie du chapitre 1

- [1] Miller, N., Maruyana, G., Beaber, R.J. & Valone, K. (1976). Speed of speech and persuasion. *Journal of Personality and Social Psychology*, vol. 34. p 615-624.
- [2] Street, R. & Brady, R. (1982). Speech rate acceptance ranges as a function of evaluative domain, listener speech rate, and communication context. *Communication Monographs*, N° 4, p 290-308.
- [3] Caelen-Haumont, G. (1991). Stratégies des locuteurs et consignes de lecture d'un texte: Analyse des interactions entre modèles syntaxiques, sémantiques, pragmatique et paramètres prosodiques, Thèse d'Etat, Aix-en-Provence.
- [4] Blumstein, S. E. 1986. «On acoustic invariance in speech». In *Invariance and variability in speech processes*, sous la dir. de Joseph S. Perkell et Denis H. Klatt, p. 179-201. Hillsdale (NJ): Lawrence Erlbaum.
- [5] Stevens, K. N. et S. E. Blumstein. 1981. «The search for invariant acoustic correlates of phonetic features». In *Perspectives on the study of speech*, sous la dir. de Peter D. Eimas et Joanne L. Miller, p. 1-38. Hillsdale (NJ): Lawrence Erlbaum.
- [6] Lindblom, B. 1988. «Phonetic invariance and the adaptive nature of speech». In *Working Models of Human Perception*, sous la dir. de Ben A.G. Elsendoom & H. Bouma. p. 139- 173. London, UK: Academic Press.
- [7] Lindblom, B. 1996. «Role of articulation in speech perception: Clues from production». *Journal of the Acoustical Society of America*, vol. 99, no 3, p. 1683-792.
- [8] Liberman, A. M. et I. G. Mattingly. 1985. «The motor theory of speech perception revised». *Cognition*, vol. 21. p. 1-36.
- [9] Fowler, C. A. 1991. «Auditory perception is not special: We see the world, we feel the world, we hear the world». *Journal of the Acoustical Society of America*, vol. 89, no 6, p. 2910-2915.
- [10] Fowler, C. A. 1996. «Listeners do hear sounds, not tongues». *Journal of the Acoustical Society of America*, vol.99, no 3, p. 1730-1741.
- [11] Lindblom, B. et J. Sundberg. 1971. «Acoustical consequences of lip, tongue, jaw, and larynx movement». *Journal of the Acoustical Society of America*, vol. 50, no 4, p. 1166-1179.

- [12] Linblom, B., J. Lubker et T. Gay. 1979. «Formant frequencies of some fixed vowels and a model of speech motor programming by predictive simulation». *Journal of Phonetics*, vol. 7, p. 147-161.
- [13] Gay, T., B. Lindblom et J. Lubker. 1981. «Production of bite-block vowels : Acoustic equivalence by selective compensation». *Journal of the Acoustical Society of America*, vol. 69, no 3, p. 802-810.
- [14] D. H. McFarland, et S. R. Baum. 1995. «Incomplete compensation to articulatory perturbation». *Journal of the Acoustical Society of America*, vol. 97, no 3, p. 1865-1873.
- [15] S. R. Baum, D. H. McFarland et M. Diab. 1996. «Compensation to articulatory perturbation: Perceptual data». *Journal of the Acoustical Society of America*, vol. 99, no 6. p. 3791-3794.
- [16] D. H. McFarland, S. R. Baum et C. Chabot. 1996. «Speech compensation to structural modifications of the oral cavity». *Journal of the Acoustical Society of America*, vol. 100, no 2. p. 1093-1104.
- [17] A. Gibson et L. McPhearson 1980. «Production of bite-block vowels by children». *PERILUS Report*, vol. 2, p. 26-43
- [18] D. K. Oller, et P. MacNeilage. 1983. «Development of speech production : Perspectives from natural and perturbed speech». In *The production of speech*, sous la dir. de Peter MacNeilage, p.91-108. New-York: Springer-Verlag.
- [19] S. R. Baum et W. F. Katz. 1988. «Acoustic analysis of compensatory articulation in children». *Journal of the Acoustical Society of America*, vol. 84, no 5, p. 1662-1668.
- [20] J. Edwards, 1992. «Compensatory speech motor abilities in normal and phonologically disordered children». *Journal of Phonetics*, vol. 20. p. 189-207.
- [21] M. M. Campbell, 1999. «Articulatory compensation for a bite block, with and without auditory feedback, in hearing and hearing impaired speakers». Thèse de doctorat, New-York, City University of New-York, 523 p.
- [22] J. R. Green, A. Moore et K. J. Reilly, 2002. «The sequential development of lip and jaw control for speech». *Journal of Speech, Language, and Hearing Research*, vol. 45, p. 66-79.
- [23] B. L. Smith et A. McLean-Muse. 1987. «Effects of rate and bite-block manipulations on kinematic characteristics of children's speech». *Journal of the Acoustical Society of America*, vol. 81, no 3, p. 747-754.
- [24] L. Ménard, H. Loevenbruck et C. Savariaux. (2006). «Articulatory and acoustic correlates of contrastive focus in French: a developmental study». In *Towards a better understanding of speech production processes*, sous la dir. de Harrington, J. and Tabain, M. Psychology Press : New York.
- [25] A. Gibson et L. McPhearson 1980. «Production of bite-block vowels by children». *University of Stockholm. PERILUS Report*, vol. 2, p. 26-43.

- [26] R.Netsell, 1985. «Construction and use of a bite-block for the evaluation and treatment of speech disorders». *Journal of Speech and Hearing Disorders*, vol. 50, p. 103-109.
- [27] C.Savariaux, P. Perrier et J.-P. Orliaguet. 1995. «Compensatory strategies for the perturbation of the rounded vowel [u] using a lip tube : A study of the control space in speech production». *Journal of the Acoustical Society of America*, vol. 98, no 5, p. 2428- 2442.
- [28] C.Savariaux, P. Perrier, J.-P. Orliaguet et J.-L. Schwartz. 1999. «Compensation strategies for the perturbation of French [u] using a lip tube. II. Perceptual analysis». *Journal of the Acoustical Society of America*, vol. 106, no 1, p. 381-393.
- [30] L.Ménard, P. Perrier et C. Savariaux. 2004. « Exploring production-perception relationships during speech development : a study of compensation strategies to a lip-tube perturbation », 147th Congress of the Acoustical Society of America, vol. 115, no 5, p. 2629.
- [31] C.J. Riordan, 1977. «Control of vocal-tract length in speech». *Journal of the Acoustical Society of America*, vol. 62, no 4, p. 988-1002.
- [32] J. S.Perkell, F. H. Guenther, H. Lane, M. Matthies, P. Perrier, J. Vick, R. Wilhelms-Tricarico et M. Zandipour. 2000. «A theory of speech motor control and supporting data from speakers with normal hearing and with profound hearing loss». *Journal of Phonetics*, vol. 28, p. 233-272.
- [33] J. S.Perkell, M. L. Matthies, M. A. Svirsky et M. I. Jordan. 1993. «Trading relations between tongue-body raising and lip rounding in production of the vowel /u/: A pilot “motor equivalence” study». *Journal of the Acoustical Society of America*, vol. 93, no 5, 2948-2961.
- [34] J. A. Jones et K. G. Munhall. 2003. «Learning to produce speech with an altered vocal tract: The role of auditory feedback». *Journal of the Acoustical Society of America*, vol.113, no 1, p. 532-543.
- [35] H.Lane, M. Denny, F. H. Guenther, M. Matthies, L. Ménard, J. S. Perkell, E. Stockmann, M. Tiede, J. Vick et M. Zandipour. 2005. «Effects of bite blocks and hearing status on vowel production». *Journal of the Acoustical Society of America*, vol.118, no 3, p. 1636- 1646.
- [36] F. H.Guenther, & J. S. Perkell. 2004. «A neural model of speech production and its application to studies of the role of auditory feedback in speech». In *Speech Motor Control in Normal and Disordered Speech*, sous la dir. de B. Maassen, R. Kent, H.Peters, P. Van Lieshout, et W. Hulstijn, p. 29-49. Oxford: Oxford University Press.
- [37] J. S.Perkell, M. Zandipour, M. Matthies, et H. Lane. 2002. « Economy of effort in different speaking conditions. I. A preliminary study of intersubject differences and modelling issues», *Journal of the Acoustical Society of America*, vol. 112, p. 1627-41.
- [38] G. J. Borden, 1976. «The effect of mandibular nerve block upon the speech of four-year-old boys». *Language and Speech*, vol. 19, 173-178.

[39] J. A. S.Kelso et B. Tuller. 1983. «Compensatory articulation under conditions of reduced afferent information : a dynamic formulation». *Journal of Speech and Hearing Research*, vol. 26, p. 217-224.

[40] P.Perrier 2005. «Control and representations in speech production». *ZAS Papers inLinguistics*, vol. 40, p. 109-132.

Bibliographie du chapitre 2

[1] A.Djeradi “Analyse et caractérisation des fricatives d’arrière de l’arabe .Acoustique, articulatoire et timing. Thèse de doctorat d’état en électronique. USTHB . Alger. 1992

[2] Stevens K.N & Klatt “Role of formant transitions in the voiced-voiceless distinctions for stops”.*J.Acoust.Soc.Am.*55, 653-659. 1974

[3] Butcher A. Ahmad K.“ Some Acoustic and Aerodynamis Characteristics of Pharyngeal Consonants in Iraqi Arabic”. *Phonetica* 44, 156-172. 1987

[4] C. Zeroual. A fiberscopic and acoustic study of « guttural » and emphatic consonants of Moroccan Arabic. *Proceedings of the XIVth International Congress of Phonetic Sciences*, San Francisco: 997- 1000, 1999.

[5] C. Zeroual. “*Propos controversés sur la phonétique et la phonologie de l’arabe marocain*”. Thèse de Doctorat, Université de Paris 8, 2000.

[6] C. Zeroual C. and L. Crevier-Buchman.“ L’arabe marocain possède des consonnes epiglottales et non pharyngales”*XXIVèmes Journées d’Etude sur laParole*, 24-27 Juin 2002, Nancy, France : 237-240, 2002.

[7] C. Zeroual. “Aerodynamic study of Moroccan Arabic guttural consonants”. *Proceedings of the XVth International Congress of Phonetic Sciences, Barcelona*: 1859-1862, 2003.

[8] Chakir Zeroual, Lise Crevier-Buchman & John H. Esling. “Ajustements laryngaux durant les consonnes gutturales et l’emphatique [S]”simples et géminées de l’arabe marocain. *JEP*2004.

[9] Ramousse R., Le Berre M. & Le Guelte L. « Introduction aux statistiques » - © 1996,

Bibliographie du chapitre 3

[1] J.Aubin ; « *Effets acoustiques et articulatoires des perturbations labiales sur la parole des enfants et des adultes* ». Université du Quebec à Montréal. Décembre 2005.

- [2] Beller, G., Schwarz, D., Hueber, T. and Rodet, X. (2006), “*Speech rates in french expressive speech*”, in *Speech Prosody*, SproSig, Dresden: ISCA
- [3] G. Beller. « *Influence de l’expressivité sur le degré d’articulation* ». .Actes des VIIèmes RJC Parole, Paris, 5 et 6 juillet 2007. Page 24 IRCAM - Institut de Recherche et Coordination Acoustique Musique
- [4] A.Calabrino. « *Effets acoustiques du débit sur la production de la parole chez des locuteurs enfants et adultes* ». Actes du Xe Colloque des étudiants en sciences du langage : 61-83, 2006.
- [5] L.Falek & O.Bouferroum & A. Djeradi “*Analyze effects of the flow on the vocalic reduction and the coarticulation in sequences CV of Arabic pharyngeal*”. The Journal of the Acoustical Society of America -- May 2008 -- Volume 123, Issue 5, p. 3744.
- [6] L.Falek, R.Djeradi, and A.Djeradi (À paraître) “*Effets du débit sur les paramètres acoustiques et temporels du signal de parole*”. Canadian Acoustics journal. Janvier 2009
- [7] McFarland, D. H., S. R. Baum et C. Chabot. 1996. «*Speech compensation to structural modifications of the oral cavity*». Journal of the Acoustical Society of America, vol 100, no2.p.1093-1104.
- [8] V. Ferbach-Hecker. La résistivité de la qualité des voyelles orales du français. In *SCOLIA*, volume 20, pages 115-134, 2005.
- [9] Gay, T. 1978. “*Effects of speaking rate on vowel formant movements*”. *Journal of the Acoustical Society of America* 63(1):223-230.
- [10] Gendrot, C. and Adda-Decker, M. (2004), “*Analyses formantiques automatiques de voyelles orales : évidence de la réduction vocalique en langues française et allemande*”, in *MIDL*.
- [11] F.Hirsch. « *Étude de la structure formantique des voyelles produites par des locuteurs bègues en vitesses d’élocution normale et rapide* » 2005. Institut de Phonétique de Strasbourg – E.A. 1339 – LiLPa -Composante Parole et Cognition - Université Marc Bloch
- [12] J. A. Jones & K. G. Munhall. 2003. «*Learning to produce speech with an altered vocal tract:The role of auditory feedback*». *Journal of the Acoustical Society of America*, vol.113, no 1, p. 532-543.
- [13] Lindblom, B. (1983), “*Economy of Speech Gestures*”, vol. The Production of Speech, Springer-Verlag, New-York.
- [14] Lane, H., M. Denny, F. H. Guenther, M. Matthies, L. Ménard, J. S. Perkell, E. Stockmann, M. Tiede, J. Vick et M. Zandipour. 2005. «*Effects of bite blocks and hearing status on vowel production*». *Journal of the Acoustical Society of America*, vol.118, no 3, p.1636-1646.

[15] Perkell, J. S., M. L. Matthies, M. A. Svirsky et M. I. Jordan. 1993. «Trading relations between tongue-body raising and lip rounding in production of the vowel /u/: A pilot “motor equivalence” study». *Journal of the Acoustical Society of America*, vol. 93, no 5, 2948-2961.

[16] Perkell, J. S., F. H. Guenther, H. Lane, M. Matthies, P. Perrier, J. Vick, R. Wilhelms-Tricarico et M. Zandipour. 2000. «A theory of speech motor control and supporting data from speakers with normal hearing and with profound hearing loss». *Journal of Phonetics*, vol. 28, p. 233-272.

Bibliographie du chapitre 4

[1]. G. Fant, J. Liljencrats & Qi-guang Lin. "A Four-Parameter Model of Glottal Flow", French-Swedish Symposium, Grenoble, Apr. 22-24, 1985, STL-QPSR4/1985.

[2]. L. Falek, H. Teffahi, A. Djeradi. (à paraître) "Simulation d'un modèle de la source vocale et détermination des paramètres de commande". *Canadian Journal of physics*.

[3]. B. Guerin, M. Mrayati, R. Carré, In I.E.E.E. Int. Conf. "On Acoustics, Speech and Signal Processing". p. 47. (1976).

[4]. B. Guerin. "Contribution aux recherches sur la production de la parole. Etude du fonctionnement de la source vocale. Simulation d'un modèle" Thèse de doctorat, Institut Nationale Polytechnique de Grenoble, France, (1978).

[5]. N. Henrich. "Etude de la source glottique en voix parlée et chantée". Thèse de doctorat de l'Université Paris 6. (2001).

[6]. K. Ishizaka, J.L. Flanagan, "Voice simulation with a body-cover model of the vocal folds" *Bell Syst. Tech. Journal.*, **51**, p. 1233., (1972).

[7]. Klatt, D.H. "Software for a cascade/parallel formant synthesizer" *Acoustical society of America*, 67(3), (pp-971-995), 1980.

[8]. M. Mrayati, B. Guerin, L.J. Boe. "A two-mass model of the vocal cords: Determination of control parameters and their respective consequences". *Acustica.*, **35**, p. 330., (1976).

[9]. X. Pelorson et al. "Optimal Glottal Configuration for Ease of Phonation", *Journal. Acoust. Soc. Am.*, 96 (6), (1994).

[10]. D. Sciamarella, C. D'alessandro. "A study of the Two-Mass Model in terms of Acoustic parameters", *International Conference on Spoken Language Processing ICSLP*. p. 2313, (2003)

[11]. D. Sciamarella, C. D'alessandro. "Modele symétrique des cordes vocales". *Acta Acustica*, Vol 90, n°4, 746-761, (2004).

- [12] Savariaux, C Abry, Christian / Cathiard, Marie-Agnès / Abed, R. El / Lallouache, M. T. / Leroy, M.-C. / Perrier, Pascal / Poveda, F. / : "*Silent speech production: anticipatory behaviour for 2 out of the 3 main vowel gestures/features while pausing*", In *SPM-1996*, 101-104.
- [13]. H. Teffahi, *SIMCV*. "*Un simulateur analogue de la propagation du son dans le conduit vocal*". Canadian Journal of Physics, vol **85**, N°12. p. 1325-1341(17) (2007).