

# Etude de la coarticulation d'une consonne arabe dans un VCVj à vitesse d'élocution variable

L.Falek ; O.Bouferroum ; A.Djéradi .

USTHB- Faculté d'électronique et informatique- LCPTS

BP32. EL Alia- Bab Ezzouar- Alger

Email : [lilalcpts@yahoo.fr](mailto:lilalcpts@yahoo.fr)

Fax: 213 21247187/21321247607

## Abstract

The anticipatory coarticulation is one of the most important characteristics of coarticulation which during years of studies has not yet got a unique interpretation. These studies have provided some models of speech production in order to understand the organisation of articulatory movements in producing different types of speech utterances. (Kent & Minifie, 1977; Fowler, 1980[6]).

Our study consists of analysing arabic utterances VCVj in brief vocalic context with Vj and speech rate as variables in order to observe the impact of the context 'right' and speech rate on the coarticulation. Thus, we have to look for some invariance in the speech signal explaining the coarticulation phenomenon related to speech rate.

So, we have analysed the formant tracking of the arabic pharyngeal /ʕ/ (or (ع) in arabic) in vocalic context /a/, /u/ and /i/ with variable speech rate (normal, fast and slow) in interrogative sentences. That is in order to confirm the anticipatory phenomenon and observe the influence of speech rate variation on the vocal tract in articulation.

The observed results have shown in our case the existence of anticipatory coarticulation and that it depends on speech rate. (non existence in fast rate however, with more prominence in slow rate).

## 1- Introduction

En production de parole, les mouvements articulatoires ne sont pas d'une transparence évidente à travers laquelle l'unité la plus abstraite serait visible. Les mouvements articulatoires se chevauchent dans le temps d'une manière très complexe. Ce chevauchement crée donc une interaction des sons de parole et rend l'unité linguistique de caractéristiques articulatoires variable. Cette interaction est plus connue sous le terme de coarticulation. La question est de savoir : est-ce qu'il est possible de formuler cette interaction entre les unités linguistiques et de la considérer ensuite comme étant invariable. On pense que si l'invariance existe, elle ne peut être recherchée que dans l'étude de la coordination des mouvements articulatoires (Abry, 1986).

Fowler (1990)[7] pense que le signal acoustique fournit des informations sur les gestes qui l'ont produit et cela malgré les effets de coarticulation. Mais elle conteste l'idée qu'une invariance

quelconque pourrait exister dans le signal acoustique de parole (Fowler, 1994) [8]

Les résonances du conduit vocales ont toujours été d'un grand intérêt pour beaucoup de chercheurs (en production, en synthèse (Denis Klatt [19]) en reconnaissance de la parole). En articuloire, ils ont permis d'expliquer des gestes qui sont très souvent d'une grande complexité comme l'anticipation (fowler) dans la coarticulation d'une consonne en contexte vocalique. Ce travail s'inscrit un peu dans ce domaine. Cette étude consiste à analyser le comportement des suivis des cinq formants d'une pharyngale arabe (/ʕ/ (ou (ع) en arabe), en contexte vocalique bref (/a/, /u/, /i/) à vitesse d'élocution variable (débits normal, rapide et lent) afin de voir leurs résistances dans ces conditions et voir s'il est possible de dégager des invariants.

Nous avons alors mener une campagne d'enregistrement d'un corpus de phrases sur un même locuteur. Ce corpus doit comporter la pharyngale /ʕ/ (ou (ع)), en contexte vocalique /a/, /u/, /i/, produite avec un débit paraissant normal (débit normal) ensuite avec deux autres vitesses d'élocution (rapides et lentes (qui ont paru bien sûr rapide et lente pour le locuteur), avec plusieurs réalisations de la même phrase (dix répétitions).

## 2-Description de l'expérience

Les phrases du corpus ont été choisies de manière à comporter la fricative arabe /ʕ/, en contexte vocalique bref /a/, /u/ /i/ .

Les phrases sont choisies de forme interrogatives pour que le locuteur parle d'une manière naturelle et sensée telles que :

/men sa ʕala/ ; /men sa ʕula/, /men sa ʕila/ pour le /ʕ/ (ou (ع)) en contexte vocalique /a/, /u/ /i/

-Le corpus a été enregistré dans un laboratoire de recherche à des heures calmes de la journée (sans bruit apparent), sur PC muni d'une carte son « sound Blaster »

-Le sujet est un locuteur masculin arabophone et articulant correctement.

- Chaque phrase a été répétée une dizaine de fois (10 réalisations afin d'obtenir des segments temporels moyennés), pour chaque contexte et pour

chaque débit. Comme nous disposons de trois contextes et 3 débits et 10 répétitions, ce qui fait un total de 90 phrases à analyser.

### 2-1 Unités d'analyses exploitées

Il s'agit d'analyser les unités :

/a  $\mathfrak{V}$ j /, ((j= /a/, /i/, /u/), soit:

/a  $\mathfrak{a}$ / /a  $\mathfrak{u}$ / /a  $\mathfrak{i}$ /)

Prélevées chacune sur les phrase interrogatives du corpus précédent

Chacune de ces unités temporelles a été obtenue à l'aide d'une segmentation manuelle [1] en utilisant la représentation temporelle, le spectrogramme, le zoom et l'écoute de chaque segment du signal de parole en utilisant le logiciel winsnoori de Yves Laprie [19].

Le langage de programmation utilisé est le MATLAB 6.5.

### 3- Détermination des trajectoires formantiques

L'analyse en trajectoires formantiques des signaux de parole est indispensable pour la recherche [2]. Il n'existe pas une méthode totalement efficace pour permettre de bonnes estimations de ces trajectoires.

Le but de ce travail n'est pas de faire une étude des méthodes de détection de formants mais de décrire l'outil que nous avons réalisé pour nos besoins d'analyse des séquences /a  $\mathfrak{V}$ j/. A cet effet, nous avons choisie la pharyngale /  $\mathfrak{V}$ / car elle est voisée et donc les trajectoires formantiques seront plus faciles à déterminer.

Les trajectoires formantiques que nous avons déterminées sont le résultat des estimations des fréquences des formants sur le spectre d'une fenêtre glissant le long du signal. Cette fenêtre a une durée choisie d'une part assez courte pour que les caractéristiques spectrales puissent être considérées comme étant stables et d'autre part assez longue pour contenir le maximum d'harmoniques possibles pour augmenter le contraste des formants sur le spectre.

La durée, avec laquelle se déplace le signal, est choisie généralement de 5 à 10 ms. Le nombre de points correspondant à la fenêtre du signal est déterminé avec la fréquence d'échantillonnage du signal.

La méthode de détection des trajectoires formantiques a été inspirée du logiciel modèle de chennoukh (1995) [5]

La durée de la fenêtre d'analyse est de 30ms avec un glissement de 5 ms. Pour une fréquence d'échantillonnage de 16KHz.

Nous avons calculer les formants pour chaque réalisation de l'unité /a  $\mathfrak{V}$ j/. Nous trouvons ainsi pour chaque unité /a  $\mathfrak{V}$ j/ dix trajectoires formantiques.

### 3-1 Synchronisation des signaux

Après avoir analyser les dix différentes trajectoires formantiques (relative à chaque segment temporel /a  $\mathfrak{V}$ j/ et à chaque débit), nous avons remarqué que le débit n'est pas rigoureusement constant (des

extensions et compressions temporelles sont observées). En fait, il n'est pas possible de reproduire des séquences de paroles avec un débit absolument constant. Donc, avant de calculer les moyennes, il est nécessaire de synchroniser les trajectoires formantiques tout en tenant compte de cette élasticité temporelle. En s'inspirant de la DTW (Dynamic Time Warping) [20], un algorithme utilisé en reconnaissance des mots isolés, nous sommes arrivés à repositionner dans le temps tous les points d'une même réalisation /a  $\mathfrak{V}$ j/ (et d'un même débit) par rapport à une réalisation de référence.

Dans cette étape, la réalisation de la séquence représentant le moins de discontinuités a été choisie comme référence. Pour chaque point de cette réalisation, nous cherchons dans le voisinage du point correspondant de la séquence à synchroniser le point le plus proche et lui affecter la même position [20].

Un exemple des résultats obtenus peut être observé sur les figures 1 et 2, (illustrant la différence entre trajectoires formantiques synchronisées et non synchronisées).

Les figures 1 et 2 correspondent aux trajectoires formantiques relevées pour le deuxième formant sur une séquence /a  $\mathfrak{i}$ /, (avec 10 réalisations).

Après avoir synchronisé les différentes réalisations pour chaque unité temporelle /a  $\mathfrak{V}$ i/ ( $\mathfrak{V}$ i=/a/, /u/, /i/), nous avons calculées la moyenne pour avoir une bonne estimation des valeurs formantiques.

Un exemple de trajectoires formantiques moyennées est illustré en figure 3.

Une validation des résultats a été faite en utilisant un synthétiseur à formants de Klatt [11]

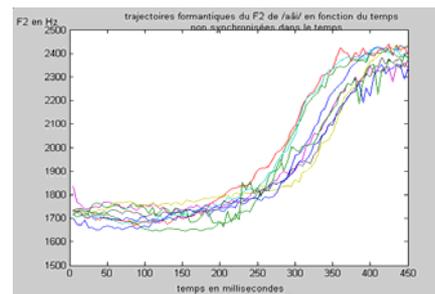


Figure 1 : Séquence /a  $\mathfrak{i}$ /, non synchronisée

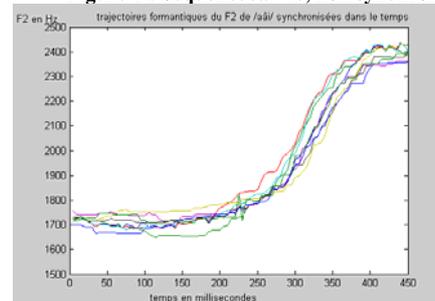


Figure 2 : Séquence /a  $\mathfrak{i}$ /, synchronisée avec algorithme DTW

### 3-2 Modélisation des trajectoires formantiques

Afin de pouvoir stoker facilement ces trajectoires formantiques moyennées, nous avons procédé à une modélisation de ces dernières, à l'aide d'une approche par un polynôme mathématique d'ordre n, par le biais de la bibliothèque Matlab. Une fonction mathématique, sera alors affectée à chaque trajectoire formantique.

Exemple de modèle de trajectoires formantiques obtenues pour la séquence /a ʕ a/ :

$$F1 = -0.7t^2 + 21.3t + 584.2$$

$$F2 = 0.4t^2 - 11.9t + 1731.7$$

$$F3 = 1.5t^2 - 44.8t + 2921.8$$

$$F4 = 0.8t^2 - 31.6t + 3865.9$$

$$F5 = -0.5t^2 + 22.9t + 3902.0$$

Les figures 3 et 4 représentent les trajectoires formantiques avant modélisation pour la figure 3 et une superposition avec le modèle pour la figure 4

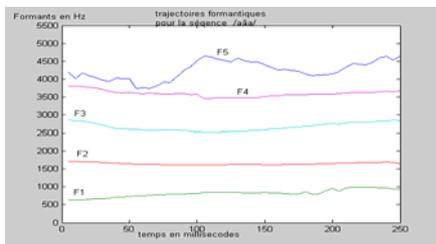


Figure 3 : trajectoires formantiques non modélisées

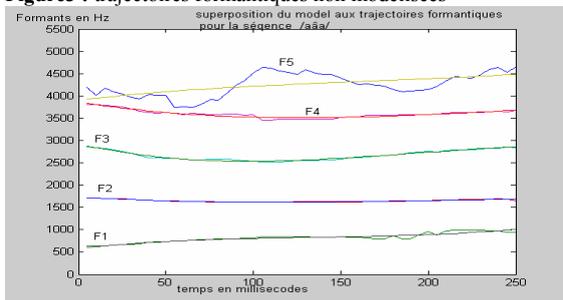


Figure 4 : superposition des trajectoires formantiques non modélisées aux modèles

Nous remarquons que le modèle est pratiquement confondu aux trajectoires non modélisées, ce qui montre que la méthode par approche polynomiale est très intéressante.

Par ailleurs la validation de ces modèles de trajectoires formantiques a été faite par le biais d'un synthétiseur à formants de Klatt. Les résultats obtenus sont pratiquement identiques aux résultats obtenus avec des trajectoires non modélisées, à l'écoute.

### 4- Etude de la résistance des trajectoires formantiques des séquences /a ʕ Vi/ (Vi=/a/, /i/, /u/) aux variations de la vitesse d'élocution

#### 4-1 Description de l'étude

Après avoir mis en place un outil pour la détermination des trajectoires formantiques, nous

l'avons appliqué aux séquences : /a ʕ a/, /a ʕ i/, /a ʕ u/ prélevées dans les phrases interrogatives /men sa ʕ ala/? ; /men sa ʕ ula/? ; /men sa ʕ ila/? (qui veut dire en arabe algérois : qui a toussé ? (à plusieurs temps du passé) à débits variables : normal, rapide et lent. Le temps interrogatif pousse le locuteur à parler d'une façon naturelle malgré la vitesse d'élocution. Les vitesses d'élocutions ont été choisis par le locuteur lui-même (celles qui ont paru normale, rapide et lente pour le locuteur). Après segmentation nous avons donc prélevé au sein de chaque phrase une séquence /a ʕ Vi/ (Vi=/a/, /i/, /u/) donc, les unités étudiées ne sont pas isolées.

Les résultats obtenus sont illustrés par la figure 5 (le /ʕ/ est représenté par /â/). Cette figure illustre les trajectoires des formants (F<sub>1</sub>, F<sub>2</sub>, F<sub>3</sub>, F<sub>4</sub>, F<sub>5</sub>) superposés à leurs modèles

- de gauche à droite : variation du contexte droit de la fricative /ʕ/ : /a/, /u/, /i/)

De haut en bas : variation du débit (lent, normal, rapide)

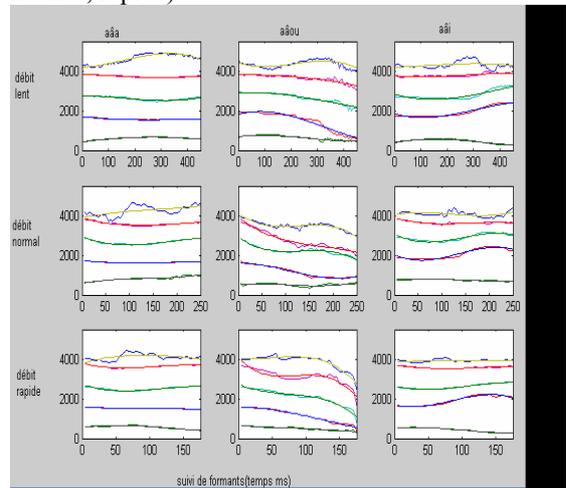


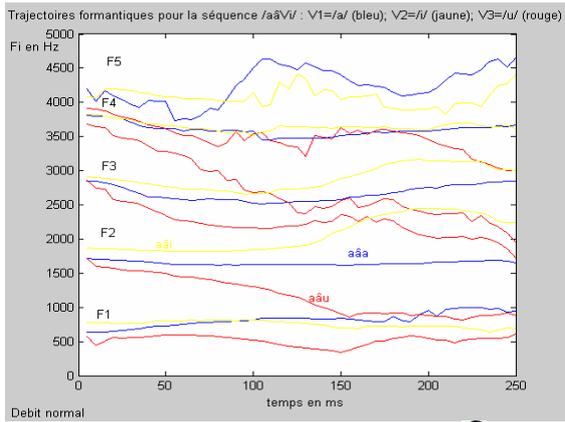
Figure 5 : trajectoires formantiques pour des séquences /a ʕ Vi/ (Vi=/a/, /i/, /u/) à vitesse d'élocution variable

#### 4-2 Interprétation des résultats

a- Les résultats obtenus sont illustrés par les figures 6, 7 et 8. Ces dernières représentent chacune les trajectoires formantiques correspondantes à chaque débit, pour les trois contextes réunis:

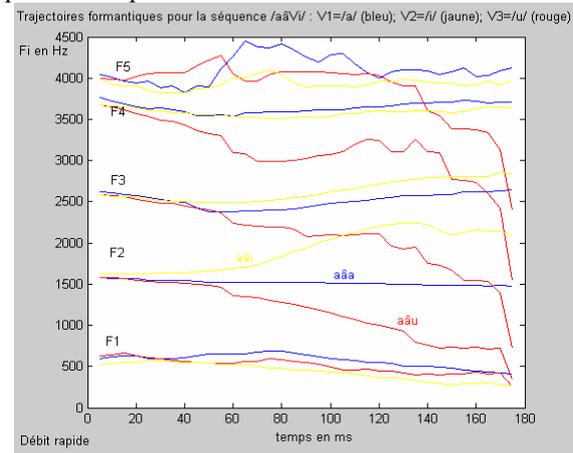
##### Sur la figure 6: débit normal

Si nous considérons par exemple les trois trajectoires F<sub>1</sub> des séquences /a ʕ a/ : en bleu ; /a ʕ i/ : jaune ; /a ʕ u/ : rouge, nous remarquons qu'elles ne commencent pas aux mêmes valeurs de fréquences (les trajectoires formantiques ne se superposent pas), pourtant les trois séquences commencent par la même voyelle /a/. Notons que le même fait est constaté pour les 4 autres trajectoires formantiques (F<sub>2</sub>, F<sub>3</sub>, F<sub>4</sub>, F<sub>5</sub>). Nous pouvons dire alors qu'il y a anticipation articulaire de la voyelle de droite depuis la voyelle de gauche

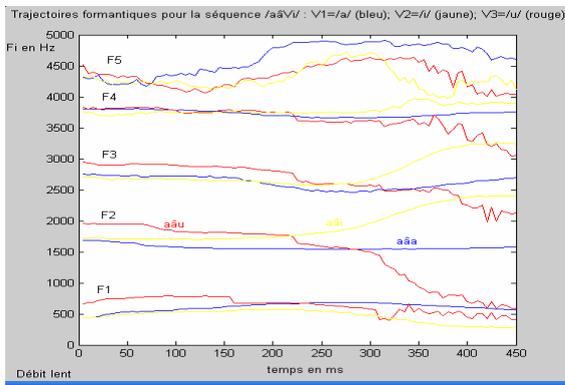


**Figure 6 :** trajectoires formantiques des séquences /a  $\Sigma$ Vi/ (Vi=/a/, /i/, /u/) à vitesse d'élocution variable  
/a  $\Sigma$ a/ : en bleu ; /a  $\Sigma$ i/ : jaune ; /a  $\Sigma$ u/ : rouge  
à débit normal

**-Sur la Figure 7- débit rapide :** pour un même formant, les trajectoires formantiques débutent toutes par la même fréquence (tous les Fi se superposent au début des trajectoires formantiques, et pour les trois séquences). Nous pouvons dire qu'il n'y a presque pas d'anticipation articulatoire.



**Figure 7 :** trajectoires formantiques des séquences /a  $\Sigma$ Vi/ (Vi=/a/, /i/, /u/) à vitesse d'élocution variable  
/a  $\Sigma$ a/ : en bleu ; /a  $\Sigma$ i/ : jaune ; /a  $\Sigma$ u/ : rouge  
à débit rapide



**Figure 8 :** trajectoires formantiques des séquences /a  $\Sigma$ Vi/ (Vi=/a/, /i/, /u/) à vitesse d'élocution variable  
/a  $\Sigma$ a/ : en bleu ; /a  $\Sigma$ i/ : jaune ; /a  $\Sigma$ u/ : rouge  
à débit lent

**- Sur la Figure 8- débit lent :** pour un même formant, les trajectoires formantiques ne débutent pas toutes par la même fréquence (les trois premiers formants ne se superposent au début des trajectoires formantiques, et pour les trois séquences). Nous pouvons dire qu'il y a une anticipation articulatoire de la voyelle de droite depuis la voyelle de gauche. En conclusion nous pouvons dire, que d'après ces trois figures, nous constatons d'abord que l'anticipation existe (ce qui confirme ce qu'a affirmé S.Öhman [14][15][16] en proposant son modèle en 1966 [15] et qui a été ensuite contredit par Gay en 1978 [9][10] en proposant un autre modèle. S.Chennouk et R.Carré en 1995 ont montré en fait que les deux modèles coexistaient (en 1995) [4 ][5].

Cependant, le fait nouveau que nous avons constaté est l'influence de la vitesse d'élocution sur l'anticipation articulatoire. Nous remarquons chez ce locuteur, que l'anticipation articulatoire diminue lorsque la vitesse d'élocution augmente, donc il est possible que les personnes qui parlent naturellement vite (par rapport à la majorité des personnes) anticipent moins. Ce fait pourrait être dû aux mouvements des muscles qui suivent de moins en moins lorsque la vitesse augmente. Ce qui peut justifier la coexistence du modèle de S. Öhman avec celui de Gay.

### 5-Conclusion

D'après les résultats obtenus, nous pouvons dire qu'au niveau d'un même locuteur, nous avons constaté :

- une influence de la vitesse d'élocution sur l'anticipation articulatoire (elle diminue avec la vitesse d'élocution),
- une influence de la vitesse d'élocution sur les dimensions des cavités résonantes du conduit vocal.

### Bibliographie

[1]	Abry .C. "Proposition pour la segmentation et l'étiquetage d'une base de donnée des sons du Français"
[2]	Carré, R., Lindblom, B. and MacNeilage, P. (1995). "Rôle de l'acoustique dans l'évolution du conduit vocal humain," Comptes Rendus de l'Académie des Sciences, Paris t. 30, série IIb, 471-476.
[3]	Carré, R. and Mody, M. (1997). "Prediction of Vowel and Consonant Place of Articulation," Proceeding of the Third Meeting of the ACL Special Interest Group in Computational Phonology, SIGPHON 97, (Madrid), pp. 26-32.
[4]	Carré, R., Marsico, E. and Serniclaes, W. (2003). "Formant transition duration versus

	prevoicing duration in voiced stop identification," Proc. of the 15th ICPhS, (Barcelona).
[5]	S. Chennoukh, 1995, ``Modélisation du Conduit Vocal en Régions Distinctives. Synthèse d'ensembles Voyelle-Voyelle et Voyelle-Consonne-Voyelle," Doctorate Dissertation, ENST, Paris
[6]	Fowler C., 1980, "Coarticulatory and theories of extrinsic timing", J. of Phonetics. 8, pp. 113-133.
[7]	Fowler C., 1990, "Listener-talker Attunements in Speech", Haskins report SR-101/102, pp. 110-129.
[8]	Fowler C., 1994, "Invariants, specifiers, cues: An investigation of locus equations as information for place of articulation", Perception & Psychophysics 55, pp. 597-610.
[9]	Gay T., 1977, "Articulatory movements in VCV sequences", J. Acoust. Soc. of Am. 62, pp. 183-193.
[10]	Gay T., 1978, "Articulatory units: Segment or syllables?», Haskins Labs. Status Report SR-54, pp. 53-63.
[11]	Klatt D.H.. 'Software for a cascade/parallel Formant synthesizer '. Acoustical Society America, 67(3), 971-995. (1980).
[12]	Mrayati M., Carré R. & Guérin B., 1988, "Distinctive regions and modes: a new theory of speech production", Speech Communication 7, pp. 257-286.
[13]	Mrayati M., Carré R. & Guérin B., 1990, "Distinctive regions and modes: articulatory-acoustic-phonetic aspects", Speech communication 9, pp. 231-238.
[14]	S. Öhman. Numerical model of coarticulation. Journal of the Acoustical Society of America. 41:310-320, 1967.
[15]	Öhman S., 1966, "Coarticulation in VCV utterances: Spectrographic measurements", J. Acoust. Soc. of Am. 39, pp. 151-168.
[16]	Öhman S. & Zetterlund S., 1974, "On symmetry in the vocal tract", Proc. Speech Communication Seminar, Stockholm.
[17]	Serniclaes, W. and Carré, R. (2002). "Contextual effects in the perception of fricative place of articulation: a rotational hypothesis," Proc. of the Int. Cong. on Speech and Language Processing, (Denver).

[18]	Serniclaes, W., Bogliotti, C. and Carré, R. (2003). "Perception of consonant place of articulation: phonological categories meet natural boundaries," Proc. of the 15th ICPhS, (Barcelona)
[19]	winsnoori. Version 1.32 Logiciel d'analyse de la parole .Equipe parole- INRIA. Lorraine- Loria- 2002
[20]	Wrigley, Stuart N. Speech Recognition by Dynamic Time Warping, 1998. URL <a href="http://www.dcs.shef.ac.uk/~stu/com326">http://www.dcs.shef.ac.uk/~stu/com326</a> . Visited April 26 th 2001