

Etude acoustique de la durée de la gémination pour la parole arabe

Mohamed Khairallah KHOUJA^{*}, Mounir ZRIGUI^{**} et Mohamed BENAHMED^{***}

^{*} *Laboratoire RIADI (unité de Monastir)*

khairallah_k@yahoo.fr

^{**} *Laboratoire RIADI (unité de Monastir), Faculté des Sciences de Monastir*

mounir.zrigui@fsm.rnu.tn

^{***} *Laboratoire RIADI, Ecole Nationale des Sciences Informatiques*

mohamed.BenAhmed@riadi.rnu.tn

Résumé: Dans ce papier, nous présentons les résultats d'une étude expérimentale des propriétés acoustiques des consonnes géminées de l'arabe. Nous visons à déterminer les durées correspondantes, pour une séquence VCCV, à la consonne géminée ainsi qu'à la voyelle qui la précède. Nous comparons ces valeurs à celles mesurées pour une séquence VCV correspondante. Les résultats ont prouvé que la durée de la consonne simple était sensiblement différente de celle géminée, ainsi que la durée de la voyelle précédant la consonne. A la base, ce travail est entrepris dans un but de déterminer les caractéristiques de durée permettant la distinction entre une consonne simple et autre géminée dans un système de reconnaissance de la parole arabe.

Mots clés: Analyse acoustique, arabe standard, gémination, durée, reconnaissance de la parole.

1. Introduction

L'originalité de la phonétique arabe se fonde, en grande partie sur la pertinence de la durée dans le système vocalique et sur la présence de consonnes emphatiques et du trait de gémination. Ces aspects particuliers jouent un rôle fondamental dans le développement morphologique nominal et verbal [Selouani, 2000].

La gémination est définie comme étant la succession de deux consonnes identiques prononcées consécutivement. En arabe, la gémination est exprimée à l'aide du symbole « ش » (الشدة). Ce symbole joue un rôle important dans la définition et le sens de certains mots. Pour la langue arabe le paramètre de durée est très important tant au niveau sémantique qu'au niveau grammatical. Il caractérise non seulement les voyelles, mais également les consonnes géminées. Cette caractéristique compense la pauvreté du système vocalique arabe. Concernant ce trait, un double problème se pose en reconnaissance automatique de l'arabe : il faut déceler les phonèmes

allongés tout en s'assurant que ce prolongement est pertinent, c'est à dire en le distinguant des allongements dus au débit d'élocution, à un accent particulier du locuteur, etc.

En effet si l'on observe l'exemple du mot « صلى » /salla:/ (prier) dérivé de la racine « صلا » qui ne s'oppose que par la gémination de la consonne « ل » au mot « صلي » /sala:/ (griller) dérivé de la racine « صلي », nous mesurons à la fois l'importance et la difficulté d'un système automatique à déceler ce trait. La gémination se manifeste par le renforcement de l'articulation et une prolongation de la fermeture de la plosive ou du continuant des autres consonnes. Là aussi, l'école traditionaliste s'impose par le fait qu'elle considère la gémination comme un simple dédoublement de la consonne [Barkat, 2000]. Il est évident qu'il existe une différence de durée notable entre la consonne géminée et son homologue simple.

Plusieurs études similaires à ce travail ont été présentées pour d'autres langues, ou la gémination est considérée comme un trait remarquable, notamment celle pour l'italien [Giovanardi & al., 1998], le grec

[Arvaniti & al., 2000] et l'indien [Samudravijaya, 2003]. Pour l'arabe standard aucune étude n'est faite à nos connaissances dans le cadre de la parole spontanée. Les travaux de Jomaa [Jomaa, 1993] et Allatif [Allatif & al., 2004], sont intéressés par l'effet de la quantité de contraste pour les séquences VVC et CCV et à la quantité vocalique pour l'arabe dialectale.

Nous commençons dans le premier paragraphe par décrire nos outils de travail et nous présenterons ensuite nos résultats.

2. Données et mesures

2.1 Acquisition des données

Vu l'absence d'un corpus annoté, nous étions amenés à effectuer nos propres enregistrements à l'aide du logiciel PRAAT¹ et d'un microphone professionnel mis à la même distance de la bouche de chacun des quatre locuteurs (2 hommes et 2 femmes) participant aux enregistrements, tous dans la même salle pratiquement isolée. On a demandé aux locuteurs de lire avec une vitesse moyenne tout en assurant une bonne articulation et en évitant les perturbations dus aux hésitations, les reprises, les respirations, ..., dans de tel cas le locuteur est invité à reprendre sa lecture. La fréquence d'échantillonnage des enregistrements est 22050 Hz. Des phrases contenant les mots choisis pour l'étude sont lus les uns après les autres sans arrêt. La durée totale des enregistrements est de 1 h 27 min. Les mots sont choisis afin d'avoir au moins, si possible, deux fois la couverture de toutes les consonnes dans le cas de gémiation et simple. Le nombre total des mots analysés a atteint 520 mots.

2.2 Analyse acoustique

Les données étudiées ont été extraites manuellement à l'aide du logiciel PRAAT. Le fait que tous les relevés aient été opérés par la même personne, garantit l'unicité de la méthode et des principes de base et donc l'homogénéité des données. Notre découpage a été fait manuellement sur la base d'indices visuels (spectre, amplitude, formant) le contrôle étant perceptif. La figure 1 montre un exemple du signal et du spectre généré par PRAAT pour le mot « ذهب » /Dahaba/.

2.3 Techniques de mesure

L'appareil vocal n'est pas capable de produire de plosives qui auraient une durée importante. L'effort produit lors de l'articulation ne peut être infini, la séquence est toujours la même : effort et relâchement. Cet effort et ses variations sont véhiculés à travers le signal, ils permettent une segmentation temporelle de la parole en unités. Dans notre cas les unités à mesurer sont les phonèmes.

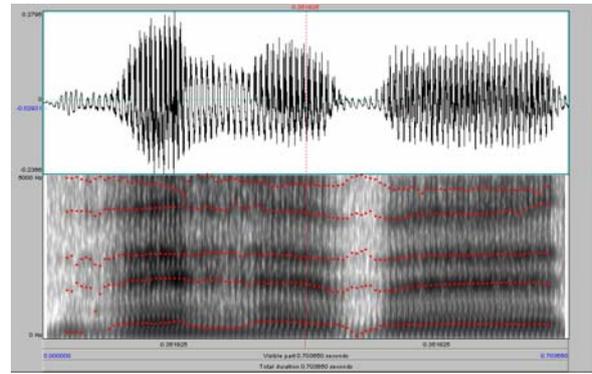


Figure 1. Exemple de la représentation du signal, du spectre et des formants généré par PRAAT pour le mot « ذهب » /Dahaba/.

Les travaux de Klatt [Klatt, 1976] sont à la base de beaucoup de modèles actuels. Ils s'appuient sur la connaissance des durées intrinsèques (D_{Int}) de chaque phonème ainsi que la durée minimum (D_{Min}). Calculer la durée d'un phonème consiste dans ce modèle à ajouter à la durée minimum, une durée qui dépend du contexte dans lequel se trouve ce phonème. Les durées ne peuvent pas descendre en dessous d'un seuil D_{Min} , propre à chaque phonème. Cependant, Klatt distingue l'état accentué ou non d'un phonème ; si nous prenons pour référence D_{Min} d'un phonème p non accentué, la valeur minimale pour ce même phonème accentué est $2 * D_{Min}$. PRCNT exprime un rétrécissement qui dépend du contexte phonétique et syntaxique. La relation utilisée par Klatt est donnée par l'équation (1) :

$$Durée = D_{Min} + \frac{((D_{Intr} - D_{Min}) * PRCNT)}{100} \quad (1)$$

Pour chaque mot extrait des enregistrements on commence par en extraire les unités à mesurer donc la voyelle, la consonne ou la consonne géminée. Chaque unité est éditer (signal + spectre + formants) séparément afin de pouvoir raffiner les limites et calculer la durée. Une moyenne de la durée pour chaque unité est calculée sur les quatre locuteurs.

3. Résultats

Les durées des diverses unités de la parole ont été calculées et employées pour étudier les caractéristiques en durée des consonnes et des voyelles qui les précèdent. Les changements observés des durées dues à la gémiation sont présentés et discutés dans cette section.

3.1 Durée de la voyelle

La Figure 2 présente la durée moyenne pour les différents tests sur les trois voyelles courtes. Pour chaque couple de mots (simple et géminé) nous calculons la durée de la voyelle précédant la consonne sujet de la gémiation. La durée pour les voyelles

¹ Logiciel développé à l'institut des sciences phonétiques de l'université d'Amsterdam par P. Boersma et D. Weenink.

longues est calculée sur d'autres mots composant le corpus, elle est donnée à titre indicatif.

Nous constatons que la durée de la voyelle précédant une gémination diminue par rapport à son homologue (consonne simple) pour 86% des mots étudiés. En moyenne on a une diminution de 11 ms pour le /a/, 13 ms pour /u/ et 12 ms pour /i/.

Cette diminution peut être expliquée par une tendance du locuteur à insister sur la gémination plutôt que sur la voyelle qui la précède. Ce résultat est trouvé pour d'autres langues tels que l'italien et l'indien (diminution de 10 à 15 ms).

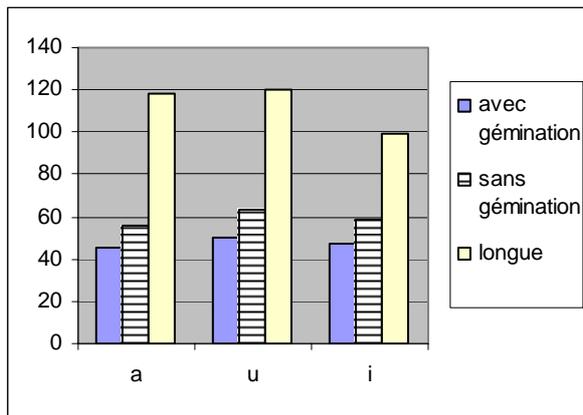


Figure 2. Représentation des durées moyennes (ms) des voyelles précédant une consonne simple ou géminée et celle d'une voyelle longue.

3.2 Durée des consonnes :

L'un des exercices les plus périlleux du traitement de la parole consiste à déterminer les frontières des différentes unités phonétiques contenues dans un énoncé. Cette difficulté tient à la nature même de la parole continue : les unités sont fortement coarticulées, et l'on passe souvent de l'une à l'autre de manière continue.

Comme précisé en 2.1, nous avons mesuré la durée pour chaque consonne simple et son homologue géminée en conservant la même séquence qui précède et qui suit la consonne. Cela est possible puisque nous partons d'une racine verbale (فعل) et nous la dérivons en (فعل), en plus de quelques noms et leurs dérivés.

Les résultats des moyennes de quelques unes de ces durées sont donnés par le tableau 1. Les valeurs sont arrondies et exprimées en ms.

Le rapport (G/S) donne une idée sur la différence entre une consonne simple et la même géminée. La thèse classique considérant ce rapport égal à 2, [Bonnot, 1979], reste comme une moyenne mais n'est pas vraie pour toutes les consonnes puisque les résultats expérimentaux obtenus donnent un rapport variant entre 1,48 et 2,24.

Consonne	Simple	Gémigné	G/S	mots
/d/ ب	58	99	1,71	4
/t/ ت	62	105	1,69	3
/θ/ ث	65	123	1,89	2
/ʃ/ ج	70	121	1,73	2
/z/ خ	65	110	1,69	3
/ʒ/ ح	63	102	1,62	2
/f/ د	69	145	2,10	3
/F/ ذ	68	139	2,04	2
/r/ ر	64	95	1,48	4
/ / ز	67	150	2,24	2
/s/ س	70	130	1,86	2
/ʃ/ ش	72	131	1,82	2
/t/ ط	69	134	1,94	2
/ʃ/ ع	61	113	1,85	1
/h/ ف	62	129	2,08	3
/s/ ق	52	101	1,94	2
/m/ ك	67	129	1,93	2
/h/ ل	68	141	2,07	4
/o/ م	62	116	1,87	4
/p/ ن	60	124	2,07	3
/ʒ/ ه	59	116	1,97	2

Tableau 1. Extrait des moyennes des durées des consonnes simples et géménées, leurs rapports et le nombre de mots étudié pour chaque consonne

Pour mieux cerner les caractéristiques en durée des consonnes nous avons établi la figure 3, qui présente les durées suivant le type des consonnes.

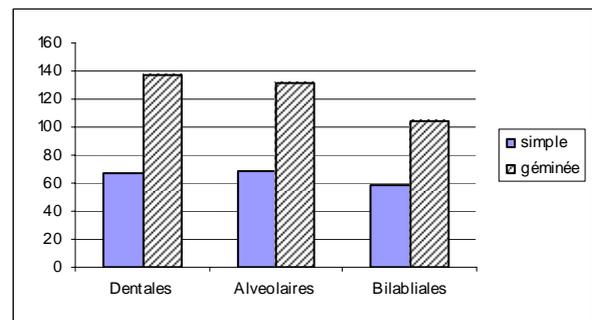


Figure 3. Représentation des durées (ms) pour certains type de consonnes dans les cas simples et géménées.

Le classement des consonnes suivant leurs types donne une vision plus claire sur la notion de durée. Cette classification montre bien la différence entre les différentes classes de consonnes puisque le rapport gémigné/simple est plus caractéristique : 2,04 pour les dentales, 1,94 pour les alvéolaires et 1,78 pour les bilabiales.

Ces résultats restent préliminaires puisqu'une étude plus complète sur un corpus plus important et présentant les caractéristiques formantiques sera plus consistante.

Conclusion

Les résultats statistiques obtenus montrent bien que nous pouvons caractériser une consonne simple d'une gémée d'un point de vue durée. Les résultats ont prouvé que la durée de la consonne gémée était sensiblement différente que celle simple avec les rapports présentés, ainsi que la durée de la voyelle précédant cette consonne. En effet, la durée de la consonne et de la voyelle qui la précède peuvent être une source de décidabilité pour la gémation. Pour pouvoir résoudre le problème de la gémation, dans un système de reconnaissance automatique de la parole continu pour la langue arabe, il faudrait différencier entre une gémation et une consonne simple suivie d'une voyelle longue. Cela passe par une bonne approche automatique acoustique pour la détermination des frontières entre consonne et voyelle.

Références

- [Allatif & al., 2004] O. Allatif, C. Abry, Adaptabilité des paramètres temporels et spectraux dans l'opposition de quantité vocalique de l'arabe de Mayadin (Syrie), *JEP'2004*, 2004.
- [Arvaniti & al., 2000] A. Arvaniti, G. Tserdanelis, On the phonetics of geminates: evidence from cypriot greek, *In proceedings of 6th International Conference on Spoken Language Processing*, volume 2, p: 559-562. Beijing, China, 2000.
- [Barkat, 2000] M. Barkat, Détermination d'indices acoustiques robustes pour l'identification automatique des parlers arabes, Thèse de doctorat, Université Lumière LYON 2, 2000.
- [Bonnot, 1979] J.F. Bonnot, étude expérimentale de certains aspects de la gémation et de l'emphase en arabe, travaux de l'institut phonétique de Strasbourg, N°11, pp. 109-118, 1979.
- [Giovanardi & al., 1998] M. Giovanardi, M-G. Di Benedetto, Acoustic Analysis of Singleton and Geminate Fricative in Italian, *The European Student Journal of Language and Speech*, 1998.
- [Jomaa, 1993] M. Jomaa, Effect of quantity contrasts on the temporal regulation of mandibular movements in Arabic, *Actes du colloque du 2^{ième} congrès : Langue arabe et technologies informatiques avancées*, pages : 141-169, 1993.
- [Klatt, 1976] D.H. Klatt, Linguistic uses of segmental duration in English: acoustic and perception evidence, *Journal of the Acoustical Society of America*, 59, pp. 1208-1221, 1976.
- [Samudravijaya, 2003] K. Samudravijaya, Durational Characteristics of Hindi Stop Consonants, *EUROSPEECH 2003*, 2003.
- [Selouani, 2000] S-A. Selouani, Reconnaissance automatique de la parole par des techniques multi-agents, connexionnistes et hybrides, Thèse d'état, Université des Sciences et Technologie Houari Boumediène, 2000.