

# LES VARIATIONS RYTHMIQUES DANS LES DIALECTES ARABES

HAMDI Rym  
Laboratoire Dynamique Du Langage  
Institut des Sciences de l'Homme  
14, avenue Berthelot  
69363 Lyon Cedex 07  
e mail : ri-hamdi@etu.univ-lyon2.fr

## ABSTRACT

Speech rhythm in the different Arabic dialects investigated has been consistently described as stress-timed. At the same time, there is preliminary evidence from perceptual experiments that listeners use speech rhythm cues to distinguish speakers from North Africa from those of the Middle East. In an attempt to elucidate the apparent contradiction, an acoustic investigation of the proportion of vocalic intervals and the standard deviation of consonantal intervals in six dialects (Morocco, Algeria, Tunisia, Egypt, Syria and Jordan) was carried out using procedures put forth by [Ram99]. The results show that complex syllables and reduced vowels in the Western dialects, and longer vowels in the Eastern dialects seem to be the main factors responsible for differences in rhythmic structures. This paper also raises questions about the discrete or continuous nature of rhythm types.

## INTRODUCTION

Les dialectes arabes peuvent être décrits comme représentant un continuum le long duquel l'intelligibilité mutuelle se dégrade progressivement au fur et à mesure que la distance géographique entre les locuteurs augmente. La communication entre les arabes éduqués est normalement accomplie via l'utilisation de l'arabe standard ou moyen, deux niveaux de langue largement traités en sociolinguistique arabe. Il est aussi bien connu des spécialistes de l'arabe que ce continuum est en quelque sorte bipolaire avec le Moyen Orient représentant les dialectes orientaux et l'Afrique du Nord abritant les dialectes Maghrébins. Bien que cette division soit une sur-simplification de la dialectologie arabe, elle est largement acceptée par la communauté linguistique et peut être appuyée par les données linguistiques. [Bar99] rapporte que, lors d'une expérience perceptuelle sur de la parole synthétique, des sujets naïfs, natifs de différentes régions du monde arabe, ont été capables de distinguer avec succès les locuteurs issus du Maghreb de ceux issus du Moyen-Orient dans 97% des cas. Pour la même tâche, une population de sujets non-arabophones ont pu distinguer les deux régions dans 56% des cas. Les résultats étant

statistiquement significatifs ( $p < 0.05$ ).

## 2. LE RYTHME COMME FACTEUR DISCRIMINANT

Les critères discriminants mentionnés par la plupart des sujets que l'arabe maghrébin semblait plus rapide et saccadé ou "plus haché" que les dialectes orientaux, impression qui -si traduite en linguistique phonétique- est en corrélation avec les caractéristiques rythmiques de la langue considérée. Selon [Ram99] le rythme joue un rôle déterminant dans la discrimination entre les langues. Ils ont mis en place des expériences où des phrases anglaises et japonaises étaient re-synthétisées et transformées de différentes manières afin de contrôler les variables du rythme et de la fréquence fondamentale. Ils rapportent qu'en gardant un  $F_0$  plat et en remplaçant toutes les consonnes d'origine par [s] et les voyelles par [a], préservant ainsi la structure syllabique, les sujets français adultes ont obtenu les scores les plus élevés pendant les tests de discrimination entre le japonais et l'anglais. Ils ont conclu que le rythme est un indice nécessaire et suffisant pour la discrimination entre les langues dans leur expérience.

Les recherches publiées sur le rythme de la parole en arabe ont classé tous les dialectes étudiés comme étant accentuels (stress-timed) par opposition à syllabiques (syllable-timed) ou à moraiques (mora-timed) [Abe67; Mil84; Ben99; Taj, Zaw & Kit, 99...]. Les épreuves ont été essentiellement d'ordre perceptuel.

Selon [Roa82] les mesures obtenues à partir du signal acoustique ne peuvent pas constituer une preuve suffisante pour classer des langues (y compris l'arabe) comme étant accentuelles par opposition à syllabiques, si "stress-timing" veut dire isochronie des accents.

[Dau83] qui est parvenue à des conclusions similaires, explique la perception des différents types de rythme essentiellement par les différences de structure syllabique, de réduction vocalique et de type d'accent. [Ram, Nes et Meh99] soutiennent que les classes de rythme perçues, et correspondant aux propriétés phonologiques mentionnées par [Dau

83,87], peuvent être mesurées au niveau du signal acoustique, [Ram, Nes et Meh99] proposent des corrélats acoustiques pour mesurer le degré de la complexité syllabique et la réduction vocalique dans chaque langue. Il s'agit d'implémenter le modèle phonologique du rythme à l'aide d'une segmentation phonétique, et de calculer les proportions moyennes des durées d'intervalles vocaliques (%V) et l'écart type des durées d'intervalles consonantiques ( $\Delta C$ ) qu'ils définissent comme ceci:

- les intervalles vocaliques: toute séquence ininterrompue de voyelles ( du début de la première voyelle de la séquence à la fin de la dernière).
- Les intervalles consonantiques: toute séquence ininterrompue de consonnes.

### 3. METHODE

#### 3.1. SUJETS

Dans cette étude nous avons examiné la production du rythme dans six dialectes arabes : Marocain (4 locuteurs), Algérien (2 locuteurs) et Tunisien (2 locuteurs) représentant l'arabe maghrébin (MA) et Jordanien (2 locuteurs), Syrien (3 locuteurs) et Egyptien (1 locuteur) représentant le Moyen-Orient ( M.O ).

#### 3.2. CORPUS

Le corpus consiste en enregistrements de l'histoire de "la bise et le soleil" . Chaque sujet écoutait en français, phrase par phrase, avant de traduire chaque phrase spontanément dans son dialecte d'origine. Les extraits ont ensuite été digitalisés à 22KH/16bits/mono. Le corpus final se compose de 140 phrases arabes (10 phrases/sujet) avec une durée moyenne de 2.5 secondes par phrase.

#### 3.3. ANALYSES

Suivant les procédures expérimentales proposées par [Ram, Nes & Meh99], nous avons utilisé le logiciel Multi-Speech pour identifier et classer les segments en tant que consonne ou voyelle en calculant :

- La durée de chaque phrase
- La durée de chaque séquence de voyelles (i.e. intervalles vocaliques)
- La durée de chaque séquence de consonnes consécutives (i.e. intervalles consonantiques).

Nous avons ensuite mesuré la proportion d'intervalles vocaliques et consonantiques (respectivement V% et C%) dans chaque phrase, et

l'écart type des intervalles vocaliques et consonantiques à l'intérieur de chaque phrase ( $\Delta C$  respectivement). Ramus, Nespor et Mehler (1999). proposent que la variable %V est en corrélation avec sont en corrélation avec

### 4. RESULTATS

Maghreb	%V	$\Delta C$	Moyen Orient	%V	$\Delta C$
Maroc	32.38	8.52	Egypte	36.98	3.89
Algérie	33.84	6.75	Jordanie	41.09	4.76
Tunisie	34.97	5.25	Syrie	43.66	4.54

Table 1 : Valeurs de %V et  $\Delta C$

La table 1 montre les proportions moyennes des intervalles vocaliques (%V) et l'écart type moyen des intervalles consonantiques ( $\Delta C$ ) pour les sujets dans chacun des 6 dialectes étudiés :

Ces résultats montrent que bien que la proportion d'intervalles vocaliques représente moins de 50% de la durée totale d'une phrase dans tous les dialectes, elle est plus importante dans les dialectes du Moyen-Orient que dans le Maghreb. En fait il y a une augmentation graduelle de %V et une baisse de  $\Delta C$  au fur et à mesure qu'on se déplace de l'Ouest vers l'Est.

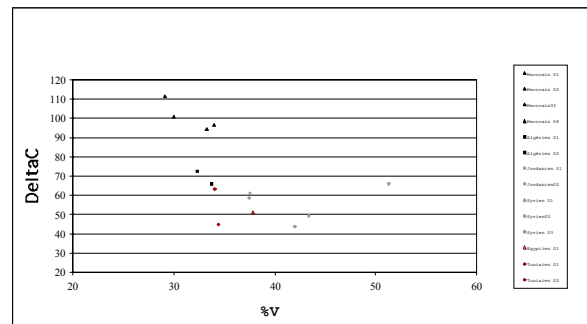


Figure 1: Distribution des dialectes sur le plan %V et  $\Delta C$

La figure I montre la corrélation négative entre %V et  $\Delta C$  et la position géographique du dialecte ( $r = -0,75$ ). Quand on va de l'Ouest vers l'Est,  $\Delta C$  diminue et %V augmente.

Des t-test ont ensuite été réalisés pour étudier statistiquement les écarts observés entre paires de dialectes différents. Il apparaît que le seuil de signification est directement proportionnel à la distance entre les deux régions où ces dialectes sont parlés . Ainsi, quand on compare la Syrie à la Jordanie ou le Maroc à l'Algérie, c'est-à-dire deux

dialectes issus de la même région, les résultats ne sont pas significatifs alors que pour des paires de dialectes situés aux extrémités opposées du continuum, comme c'est le cas du Maroc et de la Syrie, ils s'avèrent hautement significatifs ( $p < 0.001$ ).

Cependant, il est à noter que la Tunisie et l'Égypte, situées vers le centre du continuum montrent des résultats significatifs lorsqu'ils sont comparés au Maroc, un fait que nous essaierons de commenter plus avant.

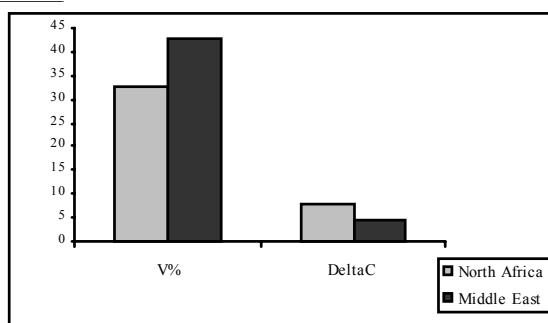


Figure II : Comparaison de %V et  $\Delta C$  entre MA et MO

La figure II est une illustration des valeurs moyennes de la proportion des intervalles vocaliques et de l'écart type des intervalles consonantiques, quand les 3 dialectes de chaque région sont regroupés ensemble. Elle montre clairement que %V est plus élevée dans les dialectes du Moyen-Orient (MO) que dans ceux du Maghreb (MA) ( $p < 0.0001$ ), alors que l'inverse est vrai pour  $\Delta C$ .

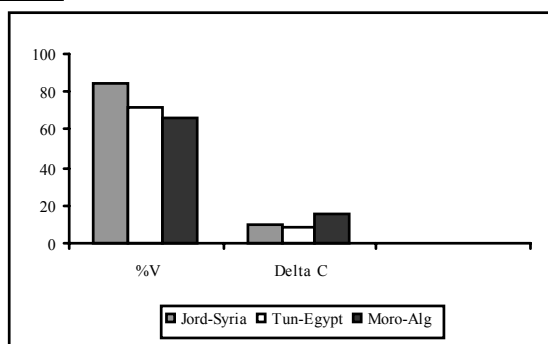


Figure III : Comparaison de %V et  $\Delta C$  entre 3 groupes de dialectes (Algérie + Maroc; Tunisie + Égypte & Jordanie + Syrie).

La figure III est une comparaison entre les différents points du continuum linguistique : le Maroc et l'Algérie représentent l'extrémité du pôle Ouest, la Jordanie et la Syrie l'extrémité Est et la Tunisie et l'Égypte constituent une zone intermédiaire. Cette comparaison confirme la baisse graduelle de %V

d'Est en Ouest avec la Tunisie et l'Égypte montrant des valeurs intermédiaires, mais montre aussi que pour  $\Delta C$  la Tunisie et l'Égypte sont plus proches des dialectes du Moyen-Orient que de ceux du Maghreb.

## 5. DISCUSSION

Les proportions d'intervalles vocaliques plus basses reflètent la présence de voyelles plus brèves. Toutes les études de l'espace vocalique arabe ont souligné le fait que dans les dialectes maghrébins les voyelles phonologiquement longues ou brèves ont une durée moindre que les voyelles correspondantes dans les dialectes moyen-orientaux. Il y a une corrélation positive entre l'écart type des intervalles consonantiques et la diversité et la complexité syllabique [Ram, Nes & Meh99].

L'arabe maghrébin est bien connu pour le processus d'effacement des voyelles brèves en syllabes ouvertes, résultant en différents groupes consonantiques et types de syllabes avec des attaques et marges complexes. Les syllabes complexes associées à des voyelles réduites, spécialement dans les dialectes marocains donnent lieu à cette impression de discours saccadé et haché rapporté par les auditeurs de Barkat [Bar00]. Les différences entre durée vocalique et complexité syllabique semblent être les facteurs principaux responsables des différences dans la structure rythmique. Selon [Ram, Nes & Meh99], les langues avec les  $\Delta C$  les plus hauts et les %V les plus bas, comme l'anglais, sont celles qui étaient traditionnellement classées comme des langues "stress-timed". Dans notre étude les dialectes exhibant ces caractéristiques sont ceux parlés au Maghreb. Puisque les dialectes comme ceux de l'Iraq [Ben99] et la Jordanie [Taj, Zaw et Kit99] ont aussi été classés comme accentuels, nous pourrions peut être envisager une grande variation au sein de la classe "stress-timed". Pour maintenir une catégorie "stress-timing" distincte d'autres types de timing, il devrait exister un ou plusieurs facteurs clés dont la présence induit constamment la perception de stress-timing. Un tel facteur serait la tendance, dans tous les dialectes arabes, des syllabes longues et lourdes à attirer l'accent. Puisque le poids syllabique varie graduellement dans ces dialectes, nous pourrions avoir l'impression de différentes sous-classes du rythme. Il est aussi intéressant de remarquer que les dialectes situés géographiquement entre les pôles sont aussi intermédiaires au point de vue traits phonétiques (figure I). Barkat [Bar00] rapporte que la plupart des erreurs de discrimination faites par ses sujets résultaient de l'incapacité de classer correctement les locuteurs tunisiens. En effet, les locuteurs tunisiens ont un %V similaire aux autres

pays du Maghreb mais un  $\Delta C$  plus proche du Moyen-Orient. Autrement dit, leurs voyelles sont légèrement plus longues et moins réduites que celles des marocains et algériens, mais significativement plus brèves que celles des syriens et jordaniens. Cependant, les locuteurs tunisiens ne montrent pas la même complexité syllabique que les autres sujets maghrébins .

En utilisant de la parole naturelle durant une tâche de discrimination, les locuteurs natifs de l'arabe pourraient aussi utiliser des indices lexicaux et morphosyntaxiques pour réduire la complexité. Une tâche perceptuelle basée sur de la parole re synthétisée d'un corpus multilingue pourrait être utile pour discriminer entre les dialectes. Une telle expérience pourrait révéler significativement si tous les dialectes arabes continuent à être regroupés en tant que langues accentuelles, ou s'ils seraient redistribués dans d'autres catégories rythmiques.

## REFERENCES

- [Abe67] Abercrombie, D., 1967: *Elements of General Phonetics*, Edinburgh University Press.
- [Bar99] Barkat, M: *Détermination des indices acoustiques robustes pour l'identification automatique des parlers arabes. Thèse de Doctorat. Université Lyon 2*, 300 pp.
- [Ben99] Benguerel, A., 1999: Stress-timing vs syllable-timing vs mora-timing: The perception of speech rhythm by native speakers of different languages. *VARIA, Etudes & Travaux n°3*.
- [Dau83] Dauer, R .M., 1983: Stress-timing and syllable-timing reanalyzed, *Journal of Phonetics* pp 51-69
- [Dau83] Dauer, R.M., 1987: Phonetic and phonological components of language rhythm, *Proceedings of the XIth ICPHS*, Tallinn, Estonia, Vol.5, pp. 447-450.
- [Mil84] Miller, M., 1984: On the perception of rhythm, *Journal of Phonetics*, 12, pp. 75-83.
- [Ram, Meh 99] Ramus, F.; Mehler, J., 1999: Language identification with suprasegmental cues. A study based on speech resynthesis. *Journal of the acoustical Society of America*, 105(1), 512-521.
- [Ram, Nes & Meh 99] Ramus, F.; Nespors, M.; Mehler, J., 1999: Correlates of linguistic rhythm in the speech signal. (*sur Internet*).
- [Roa] Roach, P., 1982: On the distinction between 'stress-timed' and 'syllable-timed' languages. *In Linguistic controversies*, D. Crystal Ed, pp. 73-79. London: Edward Arnold.
- [Taj, Zwa, Kit 99] Tajima, K.; Zawaydeh, B.; Kitabara, M., 1999: A comparative study of speech rhythm in Arabic, English and Japanese, *Proceedings of the XIV ICPHS*, San Francisco, U.S.

