

Adaptabilité des paramètres temporels et spectraux dans l'opposition de quantité vocalique de l'arabe de Mayadin (Syrie)

Omran Allatif & Christian Abry

Institut de la Communication Parlée, CNRS UMR 5009, INPG, Université Stendhal
B.P. 25 – 38040 Grenoble Cedex 9, France
allatif@icp.inpg.fr, abry@icp.inpg.fr

ABSTRACT

We studied vocalic quantity in Syrian Arabic, near the Euphrates, for the three extreme vowels /a/, /i/, /u/, under conditions where the speaker had to adapt the phonological role of quantity to demand on speech rate and/or speech intonation. Our first objective was to describe the long and short vowels in terms of their durations and values of the first three formants. By examining the factors that allow for distinction between the long and the short members, we aimed at determining whether short vowels had a proper different acoustic quality from the long ones, or if they were simply long vowels shortened in their realisation. Results set the short vowels (with a mean 60% duration of the long) in a reduced triangle inside the long point vowels. Throughout adaptive variations, quantity and quality account for the distinctiveness of A/AA pair, though quantity plays a greater role. The same for U/UU, but quality is here predominant. Quality alone can account for the distinctiveness of I/II. This last distance is perceptually important enough for allowing a contrast shift from quantity to quality, as confirmed by subsequent (unpublished [1]) perceptual experiments.

1. INTRODUCTION

L'opposition de durée des voyelles à des fins phonologiques (la quantité) est la seconde des stratégies de multiplication des éléments vocaliques, à partir des qualités de base, juste après la nasalité, dans les systèmes phonologiques des langues du monde: elle est adoptée par 51 langues sur 451 de la base UPSID [2]. Présente en arabe standard et dans ses aboutissants dialectaux, cette opposition a été l'objet de plusieurs études acoustiques (depuis [3], citons [4] et [5]). En formes de citation, les voyelles longues ont une durée supérieure à celle des brèves correspondantes, toute la question restant de savoir si cette différence se maintient ou non lorsque le locuteur doit adapter sa performance à des variations de débit d'élocution ou d'intonation (interrogation, focus, etc.). Ces différences, si elles sont statistiquement significatives, dépassent-elles ou non le seuil différentiel de durée? Quel est le rôle des différences dites de qualité vocalique, autrement dit spectrales, conséquences de l'adaptation aux conditions rythmiques et intonatives? Enfin cette différence spectrale est-elle contingente ou indispensable à la création d'un système de plus de 3 qualités vocaliques à partir du système quantitatif hérité de l'arabe classique?

2. CORPUS ET ENREGISTREMENT

Notre corpus a été réalisé par 3 locuteurs natifs de la ville syrienne de Mayadin (vallée de l'Euphrate). Les voyelles opposées sont les 3 brèves /a/, /i/, /u/ et les 3 longues correspondantes /a:/, /i:/, /u:/. Notre corpus était formé de 3 paires de six mots monosyllabiques: *zar* (il a boutonné), *zaar* (il a visité), *zir* (bouton), *ziir* (nom propre), *zur* (boutonne!), *zuur* (visite!). La phrase porteuse est du type « ga:l zir » (il a dit ...).

Les phrases sont émises selon 2 schémas mélodiques, déclaratif et interrogatif, et selon 2 débits d'élocution, normal et rapide, soient 4 conditions expérimentales. Le débit rapide a pour effet de réduire la durée et de centraliser les voyelles dans l'espace vocalique; l'interrogation a pour effet d'augmenter la durée et de « décentraliser » ou périphéraliser ces mêmes voyelles. Le but est de savoir si les brèves se confondent au niveau de leurs durées ou au niveau de leurs spectres sous une condition expérimentale ou sous une autre avec les longues.

Ceci nous donne au total 6 (phrases) x 12 (répétitions) x 2 (schémas) x 2 (débits) x 3 (locuteurs) = 864 phrases enregistrées. Conventions: A = /a/ bref; AA = /a/ long; I = /i/ bref; II = /i/ long; U = /u/ bref; UU = /u/ long. N = débit normal; R = débit rapide; D = déclaratif et I = interrogatif (p.ex. IINI = /i/ long, débit normal, schéma interrogatif).

L'enregistrement a été fait dans la chambre sourde de l'ICP, avec un DAT Sony DECK-ZE700. En éliminant les réalisations douteuses nous avons pu retenir 10 répétitions sur 12. La numérisation des phrases du corpus pour analyse sur PRAAT a été faite avec le logiciel Cool Edit 96 (44.1 kHz sur 16 bits).

3. RESULTATS

3.1. Durée

Nous pouvons classer comme suit nos moyennes (en ms) des voyelles, depuis les moins longues jusqu'aux plus longues, en fonction de leur condition expérimentale de réalisation: RD < RI < ND < NI.
RD: A=119/AA=201/I=95/II=148/U=104/UU=168;
RI: A=122/AA=212/I=102/II=161/U=112/UU=183;
ND: A=143/AA=237/I=120/II=196/U=125/UU=209;
NI: A=161/AA=277/I=127/II=211/U=130/UU=233.

En débit normal avec schéma déclaratif, soit notre référence ND, les voyelles brèves (b) par rapport à la durée de la longue (l) [soit (D_b/D_l)] durent environ 60 % de celle-ci. Le débit rapide réduit la durée des voyelles (longues et brèves confondues) de 18,93 % [$(D_{nd}-D_{rd})/D_{nd}$] parvenant ainsi à donner aux voyelles la durée la plus courte qu'elles peuvent avoir parmi les 4 conditions. L'interrogation augmente la durée des voyelles (longues et brèves confondues) de 11 % [$(D_{ni}-D_{nd})/D_{ni}$]. La condition RI soumet les voyelles à 2 effets opposés, le débit rapide et l'interrogation. Et c'est l'effet du débit qui l'emporte: par rapport à ND, la durée des voyelles (longues et brèves confondues) est réduite de 13,4 % [$(D_{nd}-D_{ri})/D_{nd}$].

Différences statistiques. Pour le schéma déclaratif, l'Anova montre un effet significatif de l'interaction longueur x débit ($F(1, 29)=22.75, p<0.001$). Cette interaction nous permet de savoir si les voyelles brèves ont une différence temporelle statistiquement significative par rapport aux longues en débit normal, et si cette différence persiste toujours sous l'effet du débit rapide. Les voyelles brèves sont plus courtes que les voyelles longues, l'écart entre brève et longue en débit normal étant significatif. Le débit rapide modifie significativement les valeurs des durées du débit normal: l'étendue de l'écart entre brève et longue diminue en passant du débit normal au débit rapide. Mais cet écart demeure significatif en débit rapide. Avec le schéma interrogatif, l'Anova montre un effet significatif de la longueur et du débit ainsi que de l'interaction longueur x débit, ($F(1, 29)=33.65, p<0.001$). Là aussi, le débit rapide modifie significativement les valeurs des durées du débit normal, l'écart entre brève et longue diminuant en passant du débit normal au débit rapide, mais restant significatif dans les deux conditions. En résumé, aussi bien sous une intonation déclarative qu'interrogative, le débit rapide change les valeurs absolues des durées des voyelles du débit normal, mais ce changement n'efface pas l'écart entre les longues et les brèves, cet écart résistant à l'effet perturbateur du débit rapide et demeurant statistiquement significatif.

Différences perceptives. Il reste à savoir si l'écart, en ms, est perceptible par l'oreille humaine. Le seuil différentiel de perception de la durée vocalique pour les voyelles qui ont une durée entre 130 et 290 ms, selon [6], suit la loi de Weber: $\Delta T/T = 22,5\%$ en moyenne. Après calcul, nous trouvons qu'aussi bien en débit normal qu'en débit rapide, pour les 3 paires de voyelles, la différence est supérieure au seuil, donc perceptivement significative.

3.2. Formants (cf. figure 1 *in fine*)

3.2.1. F1

En termes de $\Delta F1/F1_l$ (avec $\Delta F1=|F1_l-F1_b|$), nous observons les comportements suivants. Toutes voyelles confondues, le débit rapide réduit la distance entre les F1 des voyelles longues et brèves, cela est dû surtout à la centralisation des longues, à l'exception du /u:/ qui

reste pratiquement stable. L'interrogation réduit les distances entre les F1 des longues et brèves; son effet, minime, de rapprochement des F1 sur le triangle vocalique est moins important que celui du débit rapide. La condition RI, par rapport à ND, rapproche aussi les F1.

Différences statistiques. Nous avons fait une analyse de variance par schéma (déclaratif ou interrogatif). L'Anova avait 3 facteurs intra: voyelles (/a/, /i/, /u/), longueur (brève, longue) et débit (normal, rapide). Nous avons mélangé les données de nos 3 locuteurs. Avec 10 répétitions par locuteur, nous avons 30 stimuli. Le plan d'expérience est le suivant: Stimulus₃₀*Voyelle₃*Locuteur₂*Débit₂. Pour le schéma déclaratif, l'Anova montre un effet significatif de l'interaction voyelle x longueur x débit ($F(2, 58)=9.87, p=0.0002$). D'après le test de Scheffé, qui permet de comparer chaque paire de moyennes entre elles, le F1 de chaque voyelle brève est différent de celui de sa longue correspondante. Pour le schéma interrogatif, l'Anova montre un effet significatif de la voyelle, de la longueur ainsi que des interactions voyelle x longueur ($F(2, 58)=103.53, p<0.001$) et voyelle x débit ($F(2, 58)=7.48, p=0.001$). Pour résumer: en débit normal et en débit rapide, les valeurs des brèves sont significativement différentes de celles des longues. Le débit rapide ne change pas significativement les valeurs absolues des F1 des voyelles du débit normal, ni avec un schéma mélodique déclaratif ni avec un schéma mélodique interrogatif. L'écart entre les valeurs des F1 des voyelles longues et les valeurs des F1 des brèves correspondantes est significatif en débit normal et en débit rapide, ce qui veut dire que la robustesse de l'écart des F1 résiste à la perturbation du débit rapide.

3.2.2. F2

Toutes voyelles confondues, le débit rapide réduit la distance entre les F2 des voyelles longues et brèves. Cette réduction est plus importante que celle du F1. L'interrogation augmente les distances entre les F2 des longues et des brèves pour certaines voyelles et la réduit pour d'autres. (cf. figure 1).

Différences statistiques Avec un schéma déclaratif, l'Anova montre un effet significatif de l'interaction voyelle x longueur x débit ($F(2, 58)=19.62, p<0.001$). D'après le test de Scheffé, pour les voyelles [a] et [u], l'écart significatif en débit normal entre la brève et la longue, disparaît en débit rapide. L'écart entre brève et longue pour [i] reste constant d'un débit à l'autre. Avec le schéma interrogatif, l'Anova montre un effet significatif de l'interaction voyelle x longueur x débit ($F(2, 58) = 7.15, p=0.016$). D'après le test de Scheffé: pour la voyelle [a], les valeurs de F2 sont confondues quels que soient le débit et la longueur; pour [i], l'écart significatif en débit normal entre la brève et la longue se maintient en débit rapide; pour [u], l'écart entre brève et longue augmente son seuil de signification en passant du débit normal au débit rapide. Noter que pour le [a] l'opposition n'existe qu'en débit normal déclaratif, condition ND prise pour référence. En résumé, les voyelles n'ont pas un

comportement uniforme, certaines ayant un écart qui résiste à l'effet du débit rapide mais pas d'autres. Le débit rapide change significativement les valeurs absolues du F2 des voyelles du débit normal. L'écart entre les valeurs du F2 des voyelles longues et les valeurs du F2 des voyelles brèves est significatif en débit normal, pour certaines voyelles mais pas pour d'autres, et de même en débit rapide. Le tableau ci-dessous (Table 1) rend compte du comportement de la totalité des F2 de toutes les voyelles avec les 2 débits et avec les 2 schémas mélodiques.

Table 1 : Ecarts significatifs (+) sur F2 pour les 3 oppositions, tous locuteurs confondus.

	F2			
	Déclaratif		Interrogatif	
	Normal	Rapide	Normal	Rapide
A/AA	(+)	(-)	(-)	(-)
I/II	(+)	(+)	(+)	(+)
U/UU	(+)	(-)	(+)	(+)

Distances perceptives Nous nous référons maintenant au travail de [7], puisque nous sommes ici concernés avant tout par les différences de qualité vocalique, c'est-à-dire non par les seuils différentiels mais par les différences plutôt catégorielles dites de qualité. Les résultats de [7] montraient qu'avec une distance égale à 1.2 Bark entre 2 stimuli sur l'axe de F1, 50% des sujets naïfs arrivaient à distinguer 2 voyelles différentes, et qu'avec 1.5 Bark, 60 % percevaient les 2 stimuli en tant que 2 voyelles différentes. Alors qu'une distance de 0.6 Bark était suffisante pour que 60 % des experts en phonétique perçoivent les 2 réalisations comme différentes, et que 1.4 Bark suffisaient pour que presque 100 % des experts perçoivent les 2 réalisations en tant que 2 voyelles différentes. Bien évidemment, le pourcentage est proportionnel à la distance en Bark d'un côté et aux connaissances en phonétique des sujets de l'autre. Il y n'a pas de limite au-delà de laquelle tous les sujets sont d'accord ou non qu'un changement a eu lieu. Les résultats de cette expérience peuvent nous donner une idée approximative des frontières des qualités vocaliques. Nous savons maintenant que c'est vers 1,5 Bark (± 1 Bark, une marge due à un changement associé de F2), qu'un éventuel seuil doit s'établir pour être vraisemblable: c'est une zone critique. Nous allons faire le calcul, d'un côté avec les F1 tous seuls et de l'autre avec F1 et F'2 pour nos voyelles longues et brèves selon les 2 débits. Les résultats sont donnés dans le tableau ci-dessous (Table 2).

Table 2 : Différence (valeur absolue en Bark) sur F1 (1ère col.) et distance euclidienne sur F1+F'2 (2ème col.). Schéma mélodique déclaratif pour les 3 oppositions. Chaque moyenne est calculée à partir de toutes les réalisations de tous les locuteurs.

Normal (F ₁)		Normal (F ₁ +F'2)	
A/AA	0,7	A/AA	0,86
I/II	1,03	I/II	2,86
U/UU	0,7	U/UU	1
Rapide (F ₁)		Rapide (F ₁ +F'2)	
A/AA	0,48	A/AA	0,49
I/II	0,81	I/II	2,55
U/UU	0,68	U/UU	0,7

Pour obtenir ces valeurs, nous avons converti nos valeurs de Hz en Bark [F en Bark = $7 * \text{asinh}(F \text{ en Hz}/650)$]. Puis nous avons calculé F'2 selon le modèle

de [8]. F'2 représente une équivalence perceptive des valeurs de F2, F3 et F4 (F4 est fixé à 3560 Hz pour toutes les voyelles, soit 16,81 Bark). Très pratiquement F'2 intègre ces composantes acoustiques supérieures à F1 pour la caractérisation de la qualité de la voyelle [9]. Nous avons calculé la distance euclidienne entre les 2 termes (voyelles) de chaque paire. Ce calcul nous permet plusieurs choses. (i) Il classe nos voyelles par la distance qu'elles entretiennent entre elles: nous savons maintenant que l'ordre croissant selon cette distance est A/AA, puis U/UU, enfin I/II. (ii) Situés ainsi nos voyelles dans les travaux de [7], nous nous retrouvons bien autour de 1,5 Bark (± 1). La distance I/II dépasse la zone critique et elle est très probablement suffisante pour la création de 2 qualités vocaliques différentes. Par contre l'opposition A/AA ne devrait pas créer une telle différence de qualité, car elle atteint des valeurs inférieures à ce que nous pouvons considérer comme la limite inférieure de la zone critique. (iii) Enfin, nous pouvons apprécier le poids des formants supérieurs (F2, F3 et F4) pour chacune des voyelles, comparé au poids de F1.

4. CONCLUSION

Dans le dialecte arabe syrien de Mayadin sur l'Euphrate, les 3 voyelles extrêmes longues /a/, /i/, /u/ forment les trois sommets d'un triangle acoustique, dans lequel est inscrit le triangle réduit de leurs 3 homologues brèves, qui ne représentent en moyenne que 60% de la durée des longues. Nous avons pour but – en soumettant expérimentalement ce contraste phonologique de quantité aux variations adaptatives de contraintes de débit et/ou d'intonation – de tester si les voyelles brèves étaient des qualités acoustiques différentes des longues, ou bien des longues réalisées brèves.

A cela, il n'y a pas une réponse qui vaille pour toutes les voyelles, chaque paire ayant sa propre stratégie adaptative. Tout d'abord, sous aucune condition expérimentale nous ne trouvons que les voyelles se fusionnent à cent pour cent : sur F2 oui, mais pas sur F durée. Notre lecture des résultats de cette expérience montre en outre que les voyelles n'utilisent pas les mêmes paramètres pour affirmer leurs identités vocaliques par rapport à leurs correspondantes. Nous pouvons donc donner les réponses circonstanciées suivantes: /a/ vs. /a:/ se distinguent par la quantité et la qualité, mais surtout par la quantité qui prédomine; /u/ vs. /u:/ se distinguent aussi par la quantité et la qualité, mais ici c'est la qualité qui est prépondérante; /i/ vs. /i:/ peuvent se différencier par la qualité seule, et ils pourraient être en passe de devenir deux voyelles qualitativement indépendantes. Ce que nos expériences de perception (non publiées [1]) ont montré par ailleurs.

Remerciements: à L.-J. Boë et à J.-L. Schwartz pour les distances vocaliques; à M.A. Cathiard pour l'Anova; à Melissa Barkat pour ses indications bibliographiques.

BIBLIOGRAPHIE

[1] O. Allatif. *Caractérisation de la différence/ressemblance des voyelles cardinales longues et brèves de l'Euphrate arabe irako-syrienne*. Mémoire de DEA en Sciences du langage. Université Stendhal, Grenoble, 2003.

[2] I. Maddieson. *The Size and Structure of Phonological Inventories: Analysis of UPSID*. Experimental phonology, Ed. by Ohala J.J., Academic press, New York, 105-123, 1986.

[3] S.H. Al-Ani. *Arabic Phonology. An Acoustical and Physiological Investigation*. Mouton. La Haye, 1970.

[4] M. Jomaa, *Organisation temporelle acoustique et articulatoire de la quantité en arabe tunisien: comparaisons avec l'arabe standard et l'arabe koweïtien*. Thèse en Sciences du Langage, Université Stendhal, Grenoble, 1991.

[5] M. Barkat-Defradas. « Vers l'identification

automatique des parlers arabes ». *Langue & Linguistique*, 7, 47-95, 2001.

[6] M. Rossi. « Le seuil différentiel de durée ». *Mélanges à la mémoire de P. Delattre*. Mouton. La Haye. 435-449, 1972.

[7] N. Vallée et S. Kandel. « Can we recover vowel gestures from speech sounds? An experimental study based on an original psychophysical paradigm ». *15th ICPHS Barcelona*, 817-820, 2003.

[8] M. Mantakas, J. L. Schwartz et P. Escudier. « Modèle de prédiction du deuxième formant effectif F'2. Application à l'étude de la labialité des voyelles avant du français ». *15ème Journées d'Étude sur la Parole*, 157-161, 1986.

[9] J.-L. Schwartz, L.-J. Boë, N. Vallée et C. Abry. « The Dispersion-Focalisation Theory of vowel systems ». *Journal of Phonetics*, 25, 255-286, 1997.

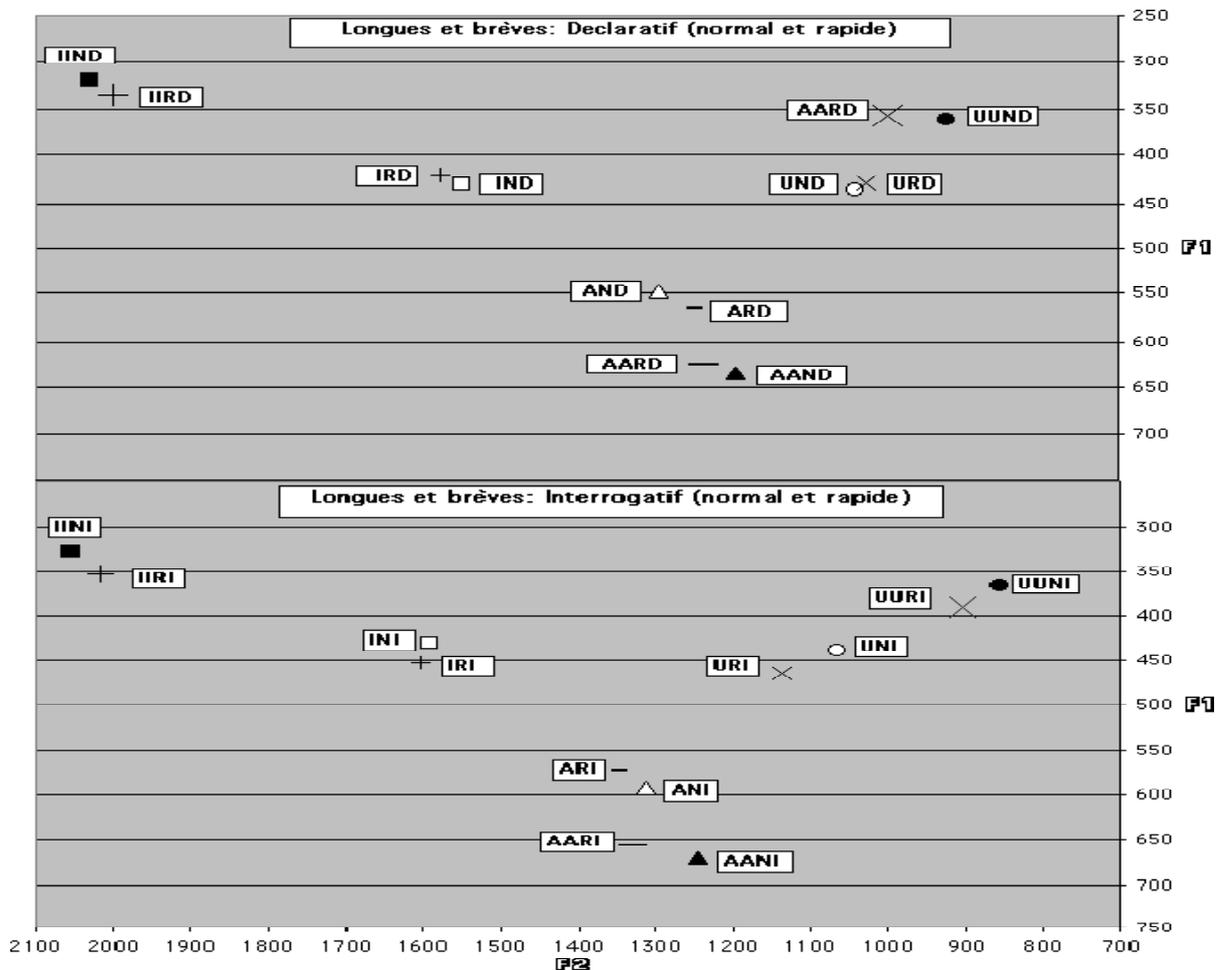


Figure 1. Dispersion des moyennes spectrales (F1-F2) pour les voyelles A, I, U, longues (VV) et brèves (V) en arabe de Syrie (3 locuteurs de Mayadin confondus). En haut la condition intonative déclarative (D); en bas interrogative (I). Dans les deux cas, le débit d'élocution peut être normal (N) ou rapide (R).