

L'espace vocalique perceptif dépend de la densité des systèmes vocaliques : Étude translinguistique en arabe marocain, en arabe jordanien et en français.

Jalaleddin Al-Tamimi

School of Education, Communication, and Language Sciences, King George VI Building
University of Newcastle
Newcastle Upon Tyne – NE1 7RU – United Kingdom
Jalal.Al-Tamimi@ncl.ac.uk
<http://www.ncl.ac.uk>

ABSTRACT

The aim of this study was to evaluate the effect of vowel density on the size and on the internal organisation of vowel spaces in perception of speech from two Arabic dialects, Moroccan and Jordanian, and from French. 30 listeners (10 per system) identified the vocalic prototypes of their systems from a MOA design (Method Of Adjustment, Johnson *et al.*, [22]). Results from this work corroborate the hypothesis of an effect of the inventory size on the size and internal organisation of vowel spaces: the larger the vowel inventory, the bigger the perceptual vowel space.

Keywords: vowel space, system density, perception of speech, dialectal Arabic, French.

1. INTRODUCTION

Les langues du monde diffèrent par le nombre et l'organisation de leurs segments vocaliques : de trois à plus de vingt voyelles distinctes (Maddieson [1] ; Maddieson & Precoda [2] ; Vallée [3]). Cette organisation semble dépendre de certaines contraintes articulatoires et perceptives et est fortement corrélée à la densité (i.e., nombre de voyelles) du système vocalique (Maddieson [1] ; Vallée [3]). En se basant sur différents travaux translinguistiques, Stevens [4], [5] et [6] et Liljencrants & Lindblom [7] et Lindblom [8] ont étudié comment le nombre de segments vocaliques dans une langue affecte leur dispersion à l'intérieur de l'espace vocalique. Différentes théories visant à prédire l'organisation des systèmes vocaliques ont ainsi vu le jour s'appuyant sur différents principes universels visant à expliquer cette organisation à travers les langues.

La Théorie Quantique (TQ : Stevens [4], [5] et [6]) propose implicitement qu'il n'y a pas d'effet de densité des systèmes. En effet, cette théorie postule la présence de zones de **stabilité** dans l'espace vocalique où les voyelles extrêmes /i a u/ se situent. Ainsi, la taille et l'organisation des espaces vocaliques ne dépendent pas de la densité des systèmes. Et puisque les trois voyelles extrêmes /i a u/ se situent dans des zones de stabilité, elles présentent moins de variabilité intra-catégorie vocalique que les voyelles non extrêmes.

À l'inverse, dans la Théorie de la Dispersion Adaptée (TDA : Liljencrants & Lindblom [7] et Lindblom [8]) les sons de la parole sont organisés selon une dispersion **maximale** et **adaptée** de leurs éléments dans l'espace acoustique suivant le principe de *Contraste Perceptuel Suffisant*. Par conséquent la taille et l'organisation des espaces vocaliques dépendent de la densité des systèmes.

À partir des travaux de ces deux théories, la théorie de la Dispersion-Focalisation (TDF : Schwartz *et al.*, [9] & [10]) a été développée et propose que les systèmes vocaliques sont régis par deux contraintes en interdépendance : la **dispersion** maximale et adaptée entre les voyelles et la **focalisation** qui évalue la saillance spectrale pour chaque voyelle. On peut supposer que pour cette théorie, la taille et l'organisation des espaces vocaliques peuvent dépendre de la densité des systèmes.

Les effets de la densité des systèmes vocaliques peuvent être prédits en suivant les propositions de chaque théorie et les principes qui les sous-tendent (pour un résumé des propositions de chaque théorie, voir Table 1). Ainsi, si l'on compare deux langues qui diffèrent en nombre de voyelles dans leur système (langue A avec dix voyelles et langue B avec trois voyelles) ; on obtient quatre cas différents pour rendre compte des effets de la densité des systèmes sur la taille et l'organisation des espaces vocaliques (voir Figure 1).

Table 1 : Prédications des trois théories sur les effets de la densité des systèmes.

	Taille espace vocalique	Taille dispersion des voyelles
TQ	Identique → stabilité	Identique → Hot-Spots Variable → autres voyelles
TDA	Variable → dispersion	Variable → toutes les voyelles
TDF	Variable → dispersion	Identique → voyelles focales Variable → autres voyelles

Les quatre cas présentés dans la Figure 1 sont des prédictions théoriques obtenues à partir des propositions des trois théories. L'organisation des voyelles dans les langues du monde est quelque peu différente. Ainsi certaines études ont montré qu'il n'existe pas d'effet de la densité des systèmes sur la taille et l'organisation des espaces vocaliques, suivant

les prédictions de la **TQ** et de la **TDF** (voir par exemple : Engstrand & Krull [12] ; Livijn [13], etc.), tandis que pour d'autres la taille et l'organisation des espaces vocaliques dépendent de la densité des systèmes, suivant les prédictions de la **TDA** (voir par exemple : Bradlow [14] ; Meunier *et al.*, [11] ; Meunier [15], etc.).

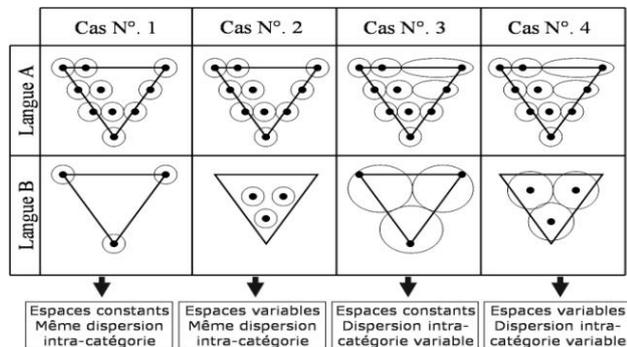


Figure 1 : Schémas hypothétiques des quatre cas d'organisation de l'espace vocalique de la langue A avec dix voyelles et la langue B avec trois voyelles. (Inspiré en partie de Meunier *et al.*, [11]).

Lors de précédentes études, nous avons montré que les espaces vocaliques dans trois langues qui diffèrent en nombre de segments vocaliques dépendent de la densité des systèmes en production des voyelles (voir : Al-Tamimi & Ferragne [16] ; Al-Tamimi [17] et [18]). Les résultats montraient que l'espace vocalique en arabe marocain (avec 5 voyelles) est significativement plus petit que celui en français (avec 11 voyelles) ; celui en arabe jordanien (avec 8 voyelles) se situant entre les deux.

Le but de ce travail est d'examiner les effets de la densité des systèmes sur l'organisation des espaces vocaliques dans les trois langues étudiées en perception de la parole. Il existe en effet peu d'étude sur les effets de la densité des systèmes en perception de la parole. Nous savons que la perception de la parole est un phénomène *many-to-one* où plusieurs stimuli acoustiques sont perçus comme un seul et unique stimulus perceptif (voir Meunier [15]). De ce point de vue, on peut supposer que l'espace vocalique perceptif est homogène entre les êtres humains – de différentes langues maternelles – car on effectue la même normalisation du signal acoustique afin de retrouver le percept d'une catégorie particulière, les voyelles dans notre cas. La seule différence que l'on peut observer dans les langues du monde en perception des voyelles concernerait davantage l'organisation interne des voyelles dans l'espace vocalique et non pas sa taille, car celle-ci dépendrait des connaissances que l'on a de sa langue maternelle et des autres langues (voir également les propositions de Kuhl [19] ; Iverson & Kuhl [20] ; Meunier [15] ; etc.).

2. MÉTHODOLOGIE

2.1. Langues, locuteurs et corpus

Trois systèmes vocaliques ont été comparés : l'arabe marocain de Fès avec 6 voyelles /i: ɪ ɛ a: ɔ u: ʔ/ (Al-Tamimi [18]), l'arabe jordanien d'Irbid avec 8 voyelles /i: e e: a a: o: o u:/ (Al-Tamimi [18]) et le français de Lyon avec 11 voyelles orales /i e ɛ a ɑ ɔ o u y ø œ/ (AM, AJ et FR, respectivement). 30 auditeurs hommes (10 par système), âgés de 20 à 30 ans et ne présentant aucun trouble du langage ni articulaire, ni auditif (testés en audiométrie), ont participé à cette étude. Tous les auditeurs ont la même situation socioprofessionnelle : des étudiants de l'Université, de même milieu socioprofessionnel et vivant tous en ville. Ils sont soit entièrement monolingues (les jordaniens, les français et quelques marocains), soit ont de faibles connaissances d'autres langues.

Un corpus par langue a été proposé afin de faciliter aux auditeurs la recherche des voyelles de synthèse qu'ils doivent effectuer. Les items lexicaux étaient composés de différentes structures syllabiques (C_1V_1 , $C_1V_1C_2$, C_1V_1CVC , etc.). C_1 et/ou C_2 contient l'une des 3 consonnes *phonologiquement* communes aux trois systèmes /b d k/ (ou contexte de présentation) ; qui est suivie de la voyelle étudiée. Chaque stimulus auditif a été présenté aux auditeurs trois fois pour correspondre aux trois contextes de présentation, afin d'examiner la possible présence d'un effet de lieu d'articulation sur le percept choisi par l'auditeur (i.e., le même stimulus auditif pour /i/ en FR a été présenté pour les mots : bilan, dit, qui). La tâche des auditeurs était de trouver la « meilleure » voyelle de synthèse correspondant à la voyelle /i/ comme dans le mot 'dit', par exemple.

2.2. Protocole expérimental

Le test perceptif proposé : test de Recherche des Prototypes, a pour rôle l'élaboration de l'espace acoustico-perceptif des auditeurs de chaque langue. C'est un test basé sur la méthode d'Ajustement de Formants, MOA (Method Of Adjustment) proposée par Johnson *et al.*, [21]. L'espace acoustique $F_1 \sim F_2$ présenté à l'écran est un espace continu et permet donc toutes les « synthèses » possibles à partir des valeurs de F_1 et de F_2 . Les valeurs de F_3 sont fixes et ont été déterminées sur la base des valeurs formantiques moyennes de voyelles produites par dix locuteurs par langue (allant de 2200 Hz pour /u/ jusqu'à 3100 Hz pour /i/, par exemple). La valeur de F_4 est restée constante à 3800 Hz pour toutes les voyelles. Les valeurs des quatre bandes passantes ont été déterminées pour chaque voyelle en s'inspirant des données de la littérature (Klatt & Klatt [22]). Les stimuli synthétiques utilisés sont constitués de : un état stable, avec une durée de 200 ms (et 100 ms pour les voyelles brèves en arabe) et un f_0 montant~descendant de 120-132-96 Hz. Chaque item lexical a été présenté aléatoirement avec 5 répétitions par auditeur. Le synthétiseur à formant a été

conçu par René Carré (d'après Klatt & Klatt [22]). 7500 prototypes vocaliques ont été identifiés (1800 en AM, 2400 en AJ, 3300 en FR). Seules les trois voyelles « extrêmes » /i a u/ dans ce travail.

2.3. Analyses et traitement des données

Les valeurs formantiques de F_1 et F_2 ont été converties en Bark (suivant la formule proposée par Schroeder *et al.* [23]). Les moyennes, écart-types et taille des aires de dispersion des espaces vocaliques et des ellipses de chaque voyelle (par la méthode de l'Enveloppe Convexe) ont été calculés. Une MANOVA à trois facteurs (Langue, Contexte de présentation et Voyelles) a été appliquée aux données.

3. RÉSULTATS ET DISCUSSION

3.1. Dispersion des espaces vocaliques

Les espaces vocaliques obtenus dans les trois langues sont significativement différents : le FR obtient l'espace le plus « périphérique » et l'AM le plus « centralisé » (voir Figure 2) : sur F_1 , la voyelle /a/ est plus « ouverte » en FR et plus « fermée » en AM qu'en AJ ; les voyelles /i u/ sont plus « fermées » en FR et plus « ouvertes » en AM qu'en AJ ($F(4, 1321)=98,31$; $p<0,001$). Sur F_2 , la voyelle /a/ est plus « antérieure » en AJ et plus « postérieure » en FR ; la voyelle /i/ est plus « antérieure » en FR qu'en AM ou en AJ et la voyelle /u/ est plus « postérieure » en FR qu'en AM ou en AJ ($F(4, 1321)=67,72$; $p<0,001$). Ces résultats ont été confirmés quel que soit le contexte de présentation.

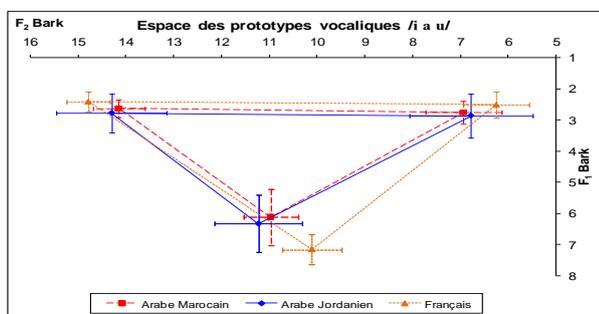


Figure 2 : Espace des prototypes vocaliques /i a u/ dans les trois langues (voyelles longues en arabe).

Les aires de dispersion des voyelles sont significativement plus « grandes » en FR par rapport à celles en AM ou l'AJ : 12,49 Bark² en AM ; 13,79 Bark² en AJ et 20,12 Bark² en FR (entre AM et FR : $F(1, 58)=93,51$; $p<0,001$; entre AJ et FR : $F(1, 58)=30,65$; $p<0,001$). Les mêmes résultats ont été obtenus quel que soit le contexte de présentation (voir Figure 3).

Ces résultats vont à l'encontre de la **stabilité** des espaces vocaliques à travers les langues ; ainsi la densité des systèmes vocaliques semble affecter la taille et la dispersion des voyelles dans les trois langues en perception de la parole. Ces résultats semblent conforter les propositions de la **TDA**.

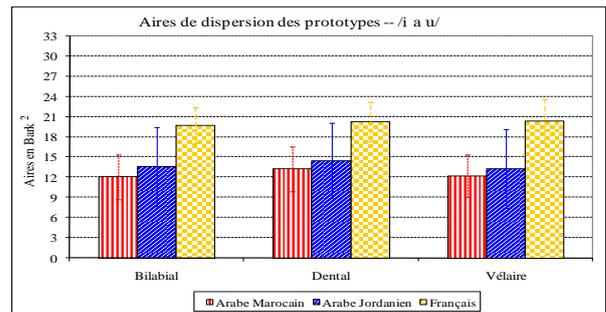


Figure 3 : Aires de dispersion des voyelles prototypes /i a u/ dans les trois langues par contexte de présentation (voyelles longues en arabe).

3.2. Ellipses de dispersion par voyelle

Les prédictions formulées précédemment des effets de la densité des systèmes concernent également la taille des ellipses de dispersion (voir Figure 4). Les résultats obtenus montrent qu'en contexte bilabial, les ellipses obtenues en AJ sont les plus grandes ; celles en FR sont les plus petites (sauf pour /i/) et qu'en contexte dental et vélair, les ellipses sont les plus grandes en AJ et les plus petites en AM. Ces résultats vont ainsi dans le sens des propositions de la **TDA**.

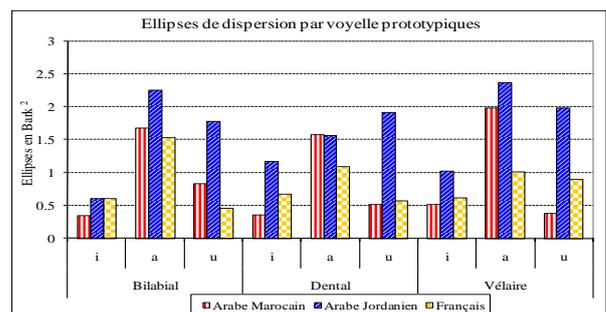


Figure 4 : Ellipses de dispersion par prototypes dans les trois langues par contexte de présentation (voyelles longues en arabe).

4. CONCLUSION

Ce travail avait pour but l'évaluation des effets de la densité des systèmes sur la taille et l'organisation des espaces vocaliques dans trois langues avec différentes densités. Les propositions de « *The Native Language Magnet* » (Iverson & Kuhl [20]) vont dans le sens d'une possible modification de l'organisation des voyelles dans l'espace vocalique et non de sa taille. Les résultats obtenus dans cette étude montrent que les auditeurs des trois langues : AM, AJ et FR semblent identifier des prototypes vocaliques différents pour les trois voyelles extrêmes /i a u/ (voyelles longues en arabe). Ainsi, la taille des espaces vocaliques est significativement différente entre les trois langues, mais également la taille des ellipses de dispersion semble affectée. L'utilisation de tous les prototypes identifiés par les auditeurs ne change pas les résultats et confirme l'expansion de l'espace vocalique en FR par rapport aux deux autres langues. Ces résultats vont dans le sens

des propositions de la **TDA** selon lesquelles, les langues semblent organiser leurs systèmes vocaliques en fonction du nombre d'éléments qui les constituent : plus une langue a de voyelles et plus l'occupation de l'espace est importante, afin de garder le maximum de distinctivité entre elles.

En production des voyelles, l'utilisation de la dynamique semble permettre une meilleure discrimination entre les voyelles dans les trois langues (voir, Al-Tamimi [17] et [18]). L'utilisation des stimuli dynamiques (i.e., associés à une transition consonantique) semble améliorer l'identification des prototypes dans les trois langues. Et, les mêmes effets de densité des systèmes sont observés avec la dynamique (voir Al-Tamimi [18]).

REMERCIEMENT

Je remercie René Carré, Egidio Marsico, François Pellegrino et Emmanuel Ferragne pour leur aide.

BIBLIOGRAPHIE

- [1] I. Maddieson. *Patterns of Sounds*: Cambridge University Press, Cambridge, 1984.
- [2] I. Maddieson and K. Precoda. Updating UPSID. In *UCLA Working Papers in Phonetics*, volume 74, pages 104-111, 1991.
- [3] N. Vallée. *Systèmes vocaliques : de la typologie aux prédictions*. Thèse de Doctorat, Université Stendhal Grenoble, Grenoble, France, 1994.
- [4] K. N. Stevens. The Quantal Nature of Speech: Evidence from articulatory-acoustic data. In *Human communication: a unified view*, E. E. David, and P. B. Denes, (Eds), McGraw-Hill, New York, pages. 51– 66, 1972.
- [5] K. N. Stevens. On the quantal nature of speech. In *J. Phon.* volume 17:1, pages 3-45, 1989.
- [6] K. N. Stevens. Acoustic and Perceptual Evidence for Universal Phonological Features. In *Proc. 15th ICPHS*, pages 33-38, 2003.
- [7] J. Liljencrants and B. Lindblom. Numerical Simulation of Vowel Quality Systems: The role of Perceptual Contrast. In *Language*, volume 48, pages 839-862, 1972.
- [8] B. Lindblom. Phonetic universals in vowel systems. In *Experimental Phonology*, J. J. Ohala, and J. J. Jaeger (Eds), Academic Press, New York, pages 13-44, 1986.
- [9] J. L. Schwartz, L.J. Boë, N. Vallée and C. Abry. Major trends in vowel system inventories. In *J. Phon.* volume 25, pages 233-253, 1997.
- [10] J. L. Schwartz, L.J. Boë, N. Vallée and C. Abry. The Dispersion-Focalisation Theory of Vowel Systems. In *J. Phon.* volume 25, pages 255-286, 1997.
- [11] C. Meunier, C. Frenck-Mester, T. Lelekov-Boissard and, M. Le Besnerais. Production and perception of vowels: does the density of the system play a role?" In *Proc. 15th ICPHS*, pages 723-726, 2003.
- [12] O. Engstrand, O. and D. Krull. Effect of inventory size on the distribution of vowels in the formant space: preliminary data for seven languages. In *PERILUS*, volume 13, pages 15-18, 1991.
- [13] P. Livijn. Acoustic distribution of vowels in differently sized inventories - hot spots or adaptive dispersion? In *Proc. 13th FONETIK*, pages 93-96, 2000.
- [14] A. Bradlow. A comparative acoustic study of English and Spanish vowels. In *J. Acoust. Soc. Amer.* volume 97: 3, pages 1916-1924, 1995.
- [15] C. Meunier. Invariants et Variabilité en Phonétique. In *Phonologie et phonétique*, N. Nguyen, S. Wauquier-Gravelines and J. Durand (Eds), Hermès, Paris, pages 350-374, 2005.
- [16] J. Al-Tamimi and E. Ferragne. Does vowel space size depend on language vowel inventories? Evidence from two Arabic dialects and French", In *Proc. InterSpeech*, pages 2465-2468, 2005.
- [17] J. Al-Tamimi. Analyse dynamique de la réduction vocalique en contexte CV à partir des pentes formantiques en arabe dialectal et en français. In *Actes des XXVI^{èmes} JEP*, pages 357-360, 2006.
- [18] J. Al-Tamimi. *Indices dynamiques et perception des voyelles : étude translinguistique en arabe dialectal et en français*. Thèse de Doctorat, Université Lyon 2, Lyon, France, 2007.
- [19] P. K. Kuhl. Human adults and Human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. In *Perception & Psychophysics*, volume 50:2, pages 93-107, 1991.
- [20] P. Iverson and P. K. Kuhl. Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. In *J. Acous. Soc. of Amer.* volume 97:1, pages 553-562, 1995.
- [21] K. Johnson, E. Flemming and R. Wright. The hyperspace effect: Phonetic targets are hyperarticulated. In *Language*, volume 69, pages 505-528, 1993.
- [22] D. Klatt and L. C. Klatt. Analysis, synthesis, and perception of voice quality variations among male and female talkers. In *J. Acous. Soc. of Amer.* volume 87:2, pages 820-857, 1990.
- [23] M. Schroeder, B. Atal and J. Hall. Optimizing digital speech coders by exploiting masking properties of the human ear. In *J. Acous. Soc. of Amer.* volume 66, pages 1647-1652, 1979.

ⁱ Le statut des voyelles brèves en AM a suscité des controverses. La littérature propose un système bref à 0, 1 ([ə]), 2 ([ə u]) ou 3 ([ɪ a u]) voyelles. Les résultats en production des voyelles obtenus dans Al-Tamimi [18] tendent à montrer que les voyelles brèves en AM obtiennent des ellipses de dispersion très importantes : la voyelle [ə], produites par dix locuteurs dans différents environnements consonantiques, obtient des ellipses allant de la voyelle [ɪ] à la voyelle [ɐ]. Cette variation de qualité de la voyelle est affectée par les différences inter-individuelles. Ce résultat montre que le statut *phonologique* des voyelles brèves en AM n'est pas bien défini et ainsi, j'ai décidé de le tester *perceptivement* en proposant un système vocalique bref à trois voyelles /ɪ ɐ u/.