

# Variabilité inter-langue et inter-individuelle en production et en perception : étude préliminaire en arabe dialectal et en français

Jalal-eddin AL-TAMIMI, Marion GIRARD et Egidio MARSICO

Laboratoire Dynamique Du Langage (UMR 5596)  
Institut des Sciences de l'Homme

14, avenue Berthelot-69363 Lyon Cedex 07

Tel +00 33 4 72 72 64 59 Fax: +00 33 4 72 72 65 90

e-mail : {Jalal-Eddin. Al-Tamimi, m.girard}@etu.univ-lyon2.fr, Egidio.Marsico@ish-lyon.cnrs.fr

## ABSTRACT

This paper presents a preliminary study of intra-speaker and inter-speaker variability in speech production and perception with an inter-language investigation of acoustic vocalic space according to different systems. This work aims at providing an analytic study based on individual data that might account for individual strategies. We have studied variability in vowel production and perception between six speakers of two languages: French and Arabic. The results of the first part of our work show vocalic spaces larger for perception than for production for French speakers and for Jordanian Arabic speakers. Moreover, inter-language differences in vowel dispersion seem to emerge from these results.

## 1. INTRODUCTION

La variabilité dans la parole est un phénomène peu étudié en tant que tel. Malgré cette variabilité, la communication est possible entre locuteurs d'une langue ayant des parlers très différents. On étudiera ici la variabilité acoustico-phonétique en liaison avec la phonologie au sein de deux langues, pour les segments vocaliques, en production et en perception.

Ce travail s'inscrit dans un projet de recherche plus vaste (Projet Cognitique : "*Variabilité phonétique en production et perception de parole: rôle et limites des stratégies individuelles*", dirigé par René Carré) qui a pour objectif de tenir compte de la variabilité inter-individuelle pour étudier les possibles stratégies individuelles [Hom99]. Dans ce projet trois langues présentant des inventaires vocaliques différents (français, italien et arabe) sont étudiées. On peut formuler l'hypothèse que cette différence du nombre de voyelles entraîne une occupation variable de l'espace vocalique ; la taille de l'espace de réalisation de chaque voyelle dépend ainsi du nombre de timbres vocaliques de la langue. L'espace global peut varier en fonction du nombre de voyelles.

Cette étude est également guidée par le rapport qu'entretiennent production et perception du langage parlé. En effet, l'espace vocalique perceptuel est différent de l'espace vocalique de production, si les positions dans le triangle vocalique des voyelles produites varient, en revanche, les positions des voyelles perçues comme étant "les plus représentatives" sont plus stables et

correspondent à un triangle vocalique de dimension quasi-maximale : les meilleurs prototypes perceptifs correspondent à des voyelles de production 'hyperarticulées' [Joh93].

## 2. METHODOLOGIE

### 2.1 Choix des langues

Pour cette étude préliminaire, deux langues, ayant des inventaires de systèmes vocaliques différents, ont été sélectionnées : le français (parler du centre de la France, 12 voyelles orales) (désormais : Fr) et l'arabe, pour lequel deux dialectes ont été retenus : l'arabe marocain de Casablanca (5 voyelles [Ham91]) (désormais : AM) et l'arabe jordanien (8 voyelles [Ban87]) (désormais : AJ). Ainsi, trois répertoires vocaliques différents sont représentés.

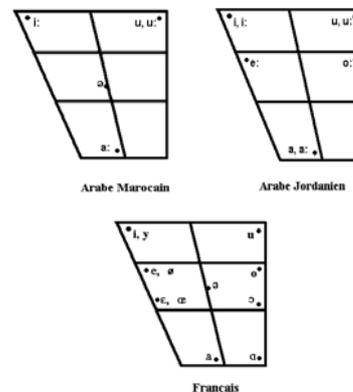


Figure 1: Systèmes phonologiques de l'arabe marocain, l'arabe jordanien et du français standard.

Dans la présente étude, nous avançons trois hypothèses :

- H.1. L'espace de perception vocalique de chaque sujet est plus large que l'espace de production.
- H.2. L'espace de perception inter-sujet est plus stable que l'espace de production.
- H.3. L'emplacement des différentes voyelles et la taille de l'espace vocalique dépendent du nombre d'unités : plus la langue présente d'unités et plus l'espace vocalique est large.

### 2.2 Protocole

Les protocoles s'organisent en deux tâches : 1) production et 2) perception, et sont réalisées via un logiciel spécialement conçu pour cette expérience. Les sons utilisés en perception sont des sons de synthèses d'une durée de 200ms. Les enregistrements en production sont effectués en chambre insonorisée, sur ordinateur PC à 22 kHz, 16 bits, mono.

#### Tâche de production :

Après une phase de familiarisation avec le logiciel, le sujet gère lui-même le déroulement des enregistrements grâce à l'interface du logiciel. Une phrase porteuse met le mot en contexte afin qu'il n'y ait aucune ambiguïté quant à sa prononciation en cas d'homographie. Il est demandé au sujet de prononcer le mot puis la syllabe comportant la voyelle, et enfin la voyelle en isolation, en la prononçant telle qu'elle est produite dans le mot (e.g. "domaine", "do", "o"). Le logiciel établit une présentation aléatoire des stimuli qui sont répétés 5 fois.

#### Tâche de perception :

Cette tâche permet d'étudier le triangle vocalique perceptuel des sujets ayant effectué la tâche de production. Il s'agit d'une MOA (*Method Of Adjustment*) sur les deux premiers formants [Joh93]. L'espace continu F1/F2 est synthétisé grâce à un logiciel de synthèse à formants. Les sujets doivent trouver la « meilleure » voyelle correspondant au mot inscrit à l'écran. Les mots issus du corpus de production sont présentés au sujet en ordre aléatoire. Pour chaque mot, il y a 5 présentations de synthèse avec un F0 à 120 Hz et 5 présentations avec un F0 à 240 Hz. Le F3 est fixé en fonction de la voyelle à identifier (valeurs de référence des F3 de Fr de [Cal89], valeurs de référence des F3 de AM et AJ issues d'un pré-test de production dont ont été extraites des moyennes de F3).

### 2.3 Corpus

Un corpus a été élaboré pour Fr (164 mots) et deux corpus pour l'arabe (AJ : 51 mots, AM : 35 mots). Ces corpus contiennent 7 contextes consonantiques pour chaque voyelle de AM, 9 contextes pour Fr et 7 pour AJ (nous n'avons pas trouvé systématiquement d'items lexicaux correspondant à certaines associations consonne-voyelle).

Les voyelles intermédiaires du système vocalique en français présentent une grande variation en production entre les différents parlers français (Nord, Sud ou centre) en fonction du type de syllabe dans lesquelles elles sont rencontrées. Afin de pouvoir contrôler l'effet perceptuel de ce phénomène, les voyelles intermédiaires apparaissent dans le corpus en syllabe fermée et en syllabe ouverte selon la prononciation du centre de la France, qui est la production dite "standard". La distinction /a/ vs. /ɑ/ postérieure apparaît donc dans notre corpus.

Pour l'arabe, nous avons adapté deux dialectes de tradition orale à l'écrit en utilisant le système d'écriture

de l'arabe standard afin de pouvoir utiliser notre protocole expérimental. On peut formuler une réserve concernant la vocalisation des voyelles brèves : en effet, elle devrait être mentionnée à l'écrit par des symboles diacritiques différents pour chaque voyelle brève. Cependant, nous avons décidé de ne pas les faire apparaître pour éviter que la prononciation des locuteurs ne soit influencée par l'arabe standard. En arabe, la phrase porteuse a donc un rôle essentiel pour identifier le mot grâce au contexte. Notons également que le système phonologique des locuteurs de AJ n'est pas influencé par le système phonologique français, contrairement à celui de AM.

Dans la présente étude, une base de données acoustiques a été recueillie par l'enregistrement de 6 locuteurs hommes (2 par représentation dialectale). Les résultats présentés ici se limitent aux trois voyelles /i/, /a/ et /u/ pour les deux dialectes arabes (voyelles longues) et le français, dans un seul contexte consonantique /d/, permettant de disposer de corpus totalement identiques en production et en perception. Dans cette étude préliminaire, les autres voyelles ont été écartées afin d'avoir une comparaison plus claire entre les deux dialectes arabes et le français. En outre, nous considérons que l'analyse de ces trois voyelles est suffisante pour évaluer les dimensions des espaces vocaliques de production et de perception, ces trois voyelles extrêmes délimitant le triangle vocalique.

### 2.4 Analyse acoustique

Une analyse acoustique a été effectuée pour la répétition de la voyelle considérée dans le contexte *mot* (et pas dans la syllabe ou en isolation car le corpus, de taille importante, n'est pas encore totalement analysé). En ce qui concerne la perception, nous avons analysé uniquement les données pour lesquelles le F0 présenté était de 120 Hz.

Les analyses acoustiques par LPC ont été effectuées au milieu de chaque voyelle (150 pour Fr, 80 pour AJ, 50 pour AM) à l'aide du logiciel Winsnorri® [Lap99]. Les valeurs moyennes et les écarts types des valeurs formantiques F1 et F2 ont été calculées pour chaque voyelle et pour chaque locuteur.

## 3. RESULTATS ET DISCUSSION

Etant donné le peu de données recueillies pour cette étude nous jugeons que l'application de statistiques inférentielles ne serait pas adaptée, seules des statistiques descriptives seront présentées.

**Table 1 :** Valeurs des aires spectrales de production et de perception pour les sujets Fr et AJ.

	Fr1	Fr2	AJ1	AJ2	Moyenne	Ecart type
Aire prod.	239416	272030	194641	193517	224901	38002
Aire perc.	642893	293589	536403	336603	452372	165302

### **3.1 Comparaisons production et perception individuelles : étude intra et inter-sujets**

Les figures n°2 à n°7 montrent une différence intra-sujet importante en ce qui concerne la production et la perception. Ces résultats corroborent notre hypothèse 1. En effet, les sujets Fr (figures 2 et 3 et table 1) utilisent globalement un espace de perception plus large que l'espace de production. Seul le /u/ du sujet Fr 2 en perception est moins postérieur et plus ouvert que celui de production. Les deux sujets, issus de la même région, présentent des espaces de production très proches. Ici, l'espace de perception semble montrer plus de variabilité inter-sujets que l'espace de production. Ceci ne soutient pas H2.

Les résultats en production et en perception des sujets AJ (figures 6 et 7 et table 1) sont en adéquation avec notre hypothèse 1. Les espaces de production et de perception du sujet AJ 2 sont plus larges que ceux du sujet AJ 1. Ces résultats montrent une variabilité interlocuteurs évidente.

Les deux sujets AM (voir figure 4 et 5) ont des espaces de production proches pour les voyelles /u/ et /a/, la voyelle /i/ présente une plus grande variabilité. Pour ces 2 sujets, l'espace de production est plus large que l'espace de perception (aires de production : AM1 : 158400, AM2 : 118293, aires de perception : AM1 : 84682, AM2 : 17740), ce phénomène comparable pour les 2 locuteurs n'entre pas dans nos hypothèses, mais considérant le peu de données, nous ne pourrions pas apporter d'explications fiables à ces résultats.

Pour information, une analyse statistique inter-sujet de type : t-Test Student a révélé que les aires de perception des sujets Fr et AJ sont significativement plus grandes que les aires de production ( $p < .05$ ). Ces résultats sont encourageants pour la suite de notre étude où un nombre plus élevé de données nous permettra de faire des analyses intra-langues.

### **3.2 Variabilité inter-langues en production**

Dans la figure 8, nous effectuons une comparaison inter-langues en production, afin de visualiser les éventuelles différences d'organisation des espaces vocaliques. Les résultats montrent les moyennes des 5 réalisations des locuteurs. On remarque que les triangles vocaliques des sujets Fr et AJ sont proches. Celui des sujets AM est plus centralisé pour la voyelle /u/ (ceci peut être dû au contexte dental, mais la comparaison avec les autres contextes consonantiques révèle la même tendance) et un /a/ moins ouvert que celui des autres dialectes (tendance identique pour les autres contextes consonantiques). Les espaces vocaliques contenus entre les voyelles /i/, /a/ et /u/ des systèmes présentant un nombre élevé de voyelles paraissent être plus larges que ceux du système présentant moins de voyelles (i.e. AM). Ces résultats vont dans le sens de H3.

### **3.3 Variabilité inter-langues en perception**

Dans cette partie, seuls les résultats des sujets Fr et AJ ont été retenus, du fait des résultats au caractère particulier des AM.

La figure 9 montre les espaces de perception des sujets Fr et AJ. On n'observe pas de différence probante entre les espaces vocaliques de perception des deux langues.

Ces résultats en perception ne montrent pas, pour une langue donnée, de plus grande stabilité de l'espace perceptif par rapport à l'espace de production (comparaison figures 8 et 9 et table 1). Ceci va à l'encontre de H2.

## **4. CONCLUSION**

Il est évident que le faible nombre de sujets considérés dans cette étude confère un caractère préliminaire à toutes conclusions. Toutefois nous pouvons établir les observations suivantes : l'hypothèse H1, selon laquelle les cibles perceptuelles sont la représentation d'une production "hyperarticulée" [Joh93] est soutenue par nos résultats. Les voyelles perçues comme étant les meilleures cibles se distribuent sur un espace vocalique très large pour une partie de nos résultats. H2 n'est pas confirmée par nos observations. L'espace de perception inter-sujet pour les résultats de nos sujets n'est pas plus stable que l'espace de production (i.e. grande variabilité inter-sujet).

Nos données en production soutiennent H3. En effet, le système présentant le moins de voyelles présente un espace vocalique moins large que ceux des systèmes présentant un inventaire d'unités plus grand.

Pour la suite de cette étude, dans le cadre du projet Cognitique, des données pour 20 locuteurs par langues (10 femmes, 10 hommes, pour les deux dialectes arabes et chaque forme régionale du français : nord, sud et centre) seront recueillies, et permettront une étude plus approfondie.

## **5. REMERCIEMENTS**

Ce projet est soutenu par le Ministère de la Recherche, programme Cognitique. Les auteurs tiennent par ailleurs à remercier Mélissa Barkat, Christine Meunier et Fanny Meunier pour les échanges scientifiques fructueux intervenus et François Pellegrino pour la réalisation des logiciels d'acquisition du corpus.

## **6. BIBLIOGRAPHIE**

- [Ban 87] BANI-YASIN, R. & J. OWENS (1987), "The Phonology of a Northern Jordanian Arabic Dialect", in *Zeitschrift der Deutschen Morgenlandischen Gesellschaft*, Vol. 137(2), pp. 297-331.
- [Cal89] Calliope, (1989). La parole et son traitement automatique, Masson (collection ENET et ENST) pp. 87-88.

[Ham 91] HAMDY, R. (1991), "étude phonologique et expérimentale de l'emphase en arabe marocain de Casablanca", Thèse de Doctorat : Sciences du Langage : Lyon2. p.172.

[Hom99] HOMBERT, J.M. et CARRE, R. (1999). "Correlation between vowel production and perception for a given speaker", *Journal of the Acoustical Society of America* 106(4), 2152 (abstract).

[Joh93] JOHNSON, K., FLEMMING, E. and WRIGHT, R. (1993). "The hyperspace effect: Phonetic targets are hyperarticulated," *Language* Vol. 69, pp. 505-528.

[Lap99] LAPRIE, Y. (1999), "Snorri, a software for speech sciences", in *Esca workshop, Methode and tool innovations for speech science education*, Matisse, Londres.

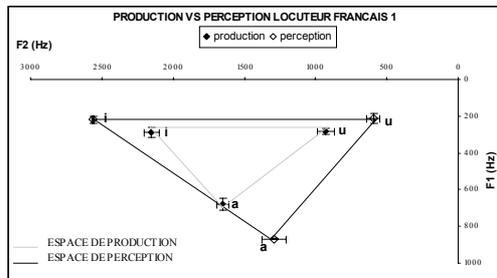


Figure 2 : Espaces vocaliques de production et de perception du locuteur français 1, représenté par /i a u/.

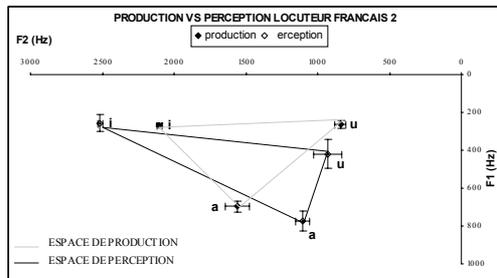


Figure 3 : Espaces vocaliques de production et de perception du locuteur français 2, représenté par /i a u/.

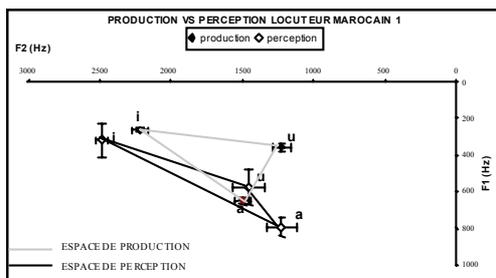


Figure 4 : Espaces vocaliques de production et de perception du locuteur marocain, représenté par /i a u/.

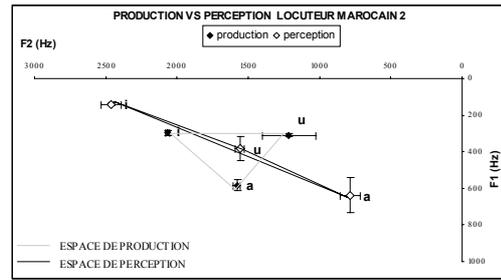


Figure 5 : Espaces vocaliques de production et de perception du locuteur marocain 2, représenté par /i a u/.

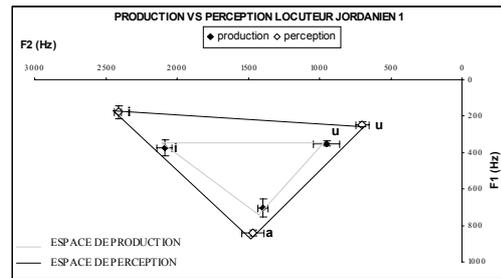


Figure 6 : Espaces vocaliques de production et de perception du locuteur jordanien 1, représenté par /i a u/.

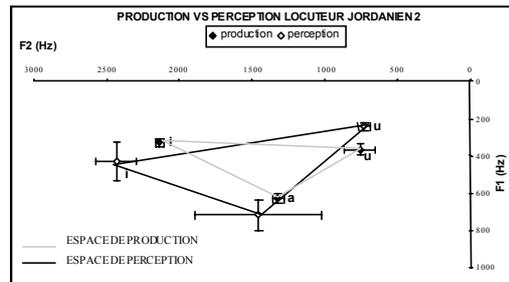


Figure 7 : Espaces vocaliques de production et de perception du locuteur jordanien 2, représenté par /i a u/.

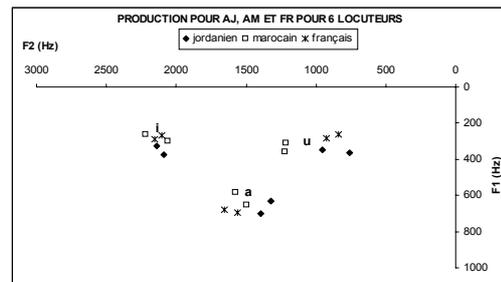


Figure 8 : Espaces vocaliques de production pour AJ, AM, Fr, 6 locuteurs, représenté par /i a u/.

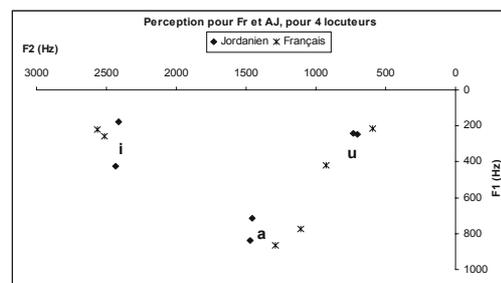


Figure 9 : Espaces vocaliques de perception pour AJ, AM, Fr, 6 locuteurs, représenté par /i a u/.