INTER-DIALECTAL AND INTER-INDIVIDUAL VARIABILITY IN PRODUCTION AND PERCEPTION: A PRELIMINARY STUDY IN JORDANIAN AND MOROCCAN ARABIC

Jalal-Eddin AL-TAMIMI & Melissa BARKAT-DEFRADAS
Laboratoire Dynamique Du Langage, Lyon

ABSTRACT

This paper presents a preliminary study of intra-speaker and inter-speaker variability in speech production and perception with an inter-dialect investigation of acoustic vocalic space according to different phonological 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 for 20 speakers of two Arabic dialects: Jordanian Arabic and Moroccan Arabic. Results show on the one hand, that vocalic spaces larger for perception than for production for speaker of both Arabic dialects; and on the other hand that the vocalic space in production for Moroccan Arabic seems more centralized than for Jordanian Arabic.

1. INTRODUCTION

Speech variability is a phenomenon that is fairly studied. Despite of speech variability, communication is still possible between speakers speaking different dialects of the same language. We will study here the phonetic variability in production and perception of vocalic segments, in two Arabic dialects: Jordanian Arabic (henceforth JA) and Moroccan Arabic (hencefort MA).

This work is part of a larger research project (Projet Cognitique : "*Variabilité phonétique en production et en perception de parole : rôle et limites des stratégies individuelles*", directed by René Carré, ENST Paris) which aim is to study inter-speakers' variability, so as to find possible individual strategies (Hombert 1999). In this project, 3 languages attesting different vocalic inventories are studied (French with 11 oral vowels, Italian with 7 vocalic segments and Arabic dialects with respectively 8 vowels for JA and 5 vowels for MA). We formulate the hypothesis that vocalic space distribution depends on the number of vocalic segments attested in the language.

AIDA 5 PROCEEDINGS XXX-XXX
CÁDIZ 2003

In this work, we will try to understand the relation between production and perception in speech. We think that perceptual vowel space is different from that of production. Indeed if the position in the vocalic triangle of produced vowels is different, the position of prototypic perceived vowels will correspond to "hyper-articulated" produced vowels (Johnson 1993).

Our goals in this preliminary study are to:

- Understand the relation between production and perception of vowels,
- Study speech variability in production and perception of vowels in two Arabic dialects: JA and MA,
- Observe the differences between vocalic realizations of male and female,
- Study the distinction between long and short vowels within Arabic dialects,
- Observe the effect of pharyngealization on the adjacent vowels both in JA and MA

Each of these different points being analysed both on the production and the perception point of view.

2. SPEECH VARIABILITY

Speech variability may be due to various factors: physiological differences (in articulation and/or audition) speakers' emotional state, sociolinguistic differences (variation between male and female phonetic realizations), inter-speakers' variability (regional and dialectal differences), etc.

2.1. Variability in production

Peterson and Barney (1952) studied the distribution of American English vowels produced by 77 speakers (males, females and children). They found that the acoustic dispersion, corresponding to each vocalic segment, undergoes an important inter-speakers' variability, (figure 1).

We can see that ellipses corresponding to different vowel categories are overlapping; the size of ellipses being different for each vowel. These differences are essentially due to physiological differences (i.e. vocal tract size between males, females and children). This is the reason why we commonly observe that female' frequencies present in average 15% higher than males' one. In a study based on males' and females' vocalic realizations, Henton (1995) found that, after normalization females' data, vocalic space is larger than males' (mostly on F1 axis).

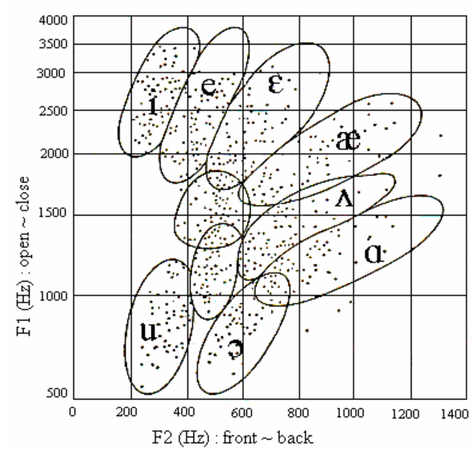

Illustration 1: Ellipses of 10 vowels of American English produced by 77 speakers. from Peterson and Barney (1952).

Intra-speaker variability can also be produced because of coarticulation. When produced in isolation, a back vowel such as [u] is not influenced by any phonetical context that may causes differences on formant frequencies. But when the subject is asked to produce the same back vowel within a consonantal context, vowel position changes considerably (cf. illustration 2).

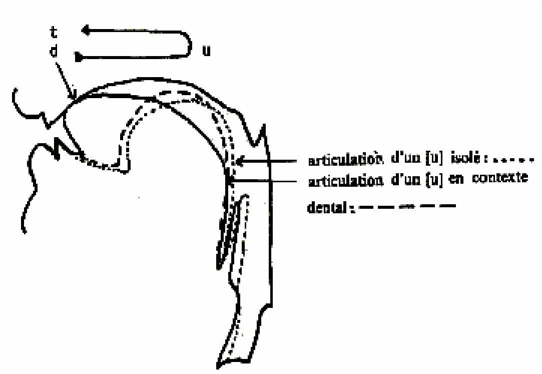

Illustration 2: Back rounded vowel [u] is fronted in dental context, (from Calliope 1989).

The same type of coarticulation phenomenon is observable in Arabic. Vowels are backed in presence of pharyngealized context, the backing -articulation being a remarkable effect on the second formant (Ghazali 1977, Elgendi 2001). In CVC context vowels are centralized (Ghazali 1979).

2.2. Variability in perception

On the basis of a set of perceptual experiments, Hombert & Puech (1984) studied the differences of Swahili speakers' vocalic perceptual representations. Hearers listened to 53 synthetic vocalic stimuli before choosing a representative category for each stimulus they heard. Results revealed that for the same vocalic system, speakers have different perceptual areas, in other words the same stimulus can be characterized as belonging to the same vowel category or not, illustration 3).

Hearer 1 **Hearer 2**

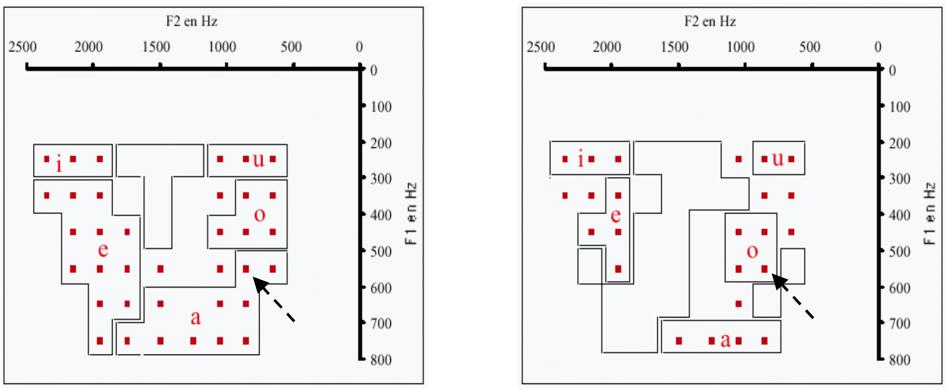

Illustration 3: Vocalic categories for two hearers of Swahili. Arrows show the segments that are categorized differently by hearer 1 and hearer 2, adapted from Hombert & Puech (1984).

The fact that stimulus can be assigned to different vowel categories by different speakers is due to the fact that speakers organize their perceptual areas differently depending on their perceptual attitudes.

2.3. Relation between Production and Perception

Several linguists have tried to explain the relationships existing between production and perception. Some theories were based on the articulation, such as the Motor Theory of speech perception revised by Liberman & Mattingly (1985). According to authors, hearers use their phonetic knowledge in order to recognize the gestural model corresponding to the acoustical reality of a specific stimulus. Twadell (1952), Hockett (1955) & Dellatre (1958) described speech perception as depending on speakers' articulatory habits.

Kluender & al. (1987) claimed perception is based on audition only. They proposed that categorical perception does not depend on speakers' articulatory attitudes but on memorization processes.

Fry (1966) took developed the notion of '*categorical perception'*. The author showed there are differences between the categorical perception of consonants and vowels. According to Fry (1966), speakers were able to categorize limits between /b d g/ (cf. illustration 4), while they were not able to establish exactly the limits between /i e a/ (cf. illustration 5). Actually, vowel categories merge due to the continuum space from which they proceed. This is not the case for consonants that are formed in a discontinuous space.

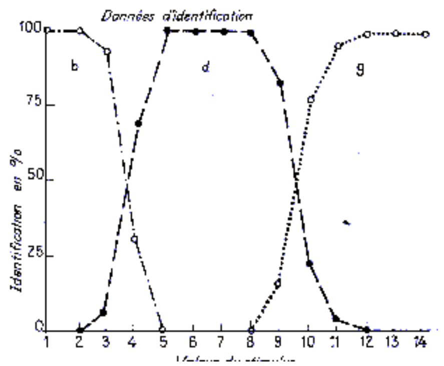

Illustration 4: Stimuli identification for /b d g/, from Fry (1966).

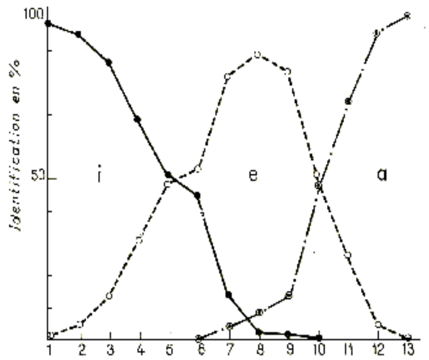

Illustration 5: Stimuli identification for /i e a/, from Fry (1966).

As mentioned earlier, Johnson & al. (1993) demonstrated prototypic representations of vowels that may correspond to a "hyper-articulated production". These experiences were based on a MOA experiment (Method of Adjustment), which revealed hearers' expectations for the sounds of their mother tongue. In this method, speakers manipulate a vocal synthesizer until the machine produces the "best vocalic target". Johnson & al. (1993) compared the vocalic spaces obtained in production and in perception. They found that vocalic space in perception is larger than productions' corresponding to a "hyper-articulated production".

2.4 Dialectal Variability in Arabic

Arabic language has a variety of regional and dialectal differences. Indeed, it is commonly accepted that the Arabic speaking world can be divided into two dialectal areas: the Middle-Eastern area (i.e., Mashreq) vs. the Western area (i.e., Maghreb) (illustration 6). The part including the west of Egypt, the south of Tunisia and Libya is considered as an intermediate area where one can find mixed-languages (Ghazali et al. 2002).

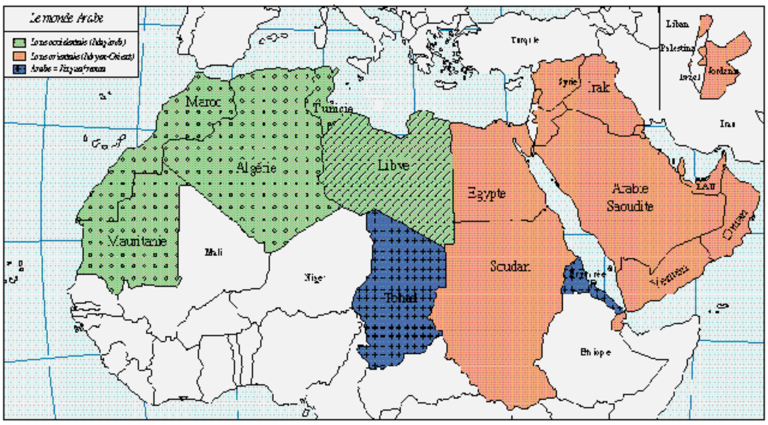

Illustration 6: In grey with dots: this area corresponds to the Western zone (i.e. Maghreb), in plain grey: this area corresponds to Middle–Eastern area (Mashreq), hatched area refers to a small part of the dialectal intermediate zone whereas crossed section in dark grey corresponds to the countries where Arabic is spoken as a *lingua franca*.(from Barkat (2000)).

Differences between dialects of Middle East and Maghreb can be found at all linguistic levels (i.e., rhythmic, lexical phonetic and phonological levels). It has been shown for example that vocalic dispersion in the dialects of Maghreb is much more centralized than in Middle Eastern dialects (Barkat 2001). As far as rhythm is concerned, dialects of Maghreb have been described as more speed and halting (i.e., jerky rhythm) than their Middle Eastern counterparts. (Ghazali et al. 2002).

3. MATERIAL AND METHODS

In this paper two dialects are studied: Jordanian Arabic of Amman and Irbid (JA henceforth) and Moroccan Arabic of Casablanca (MA henceforth). Vocalic systems of these two dialects are slightly different in terms of number of items (cf. illustration 7).

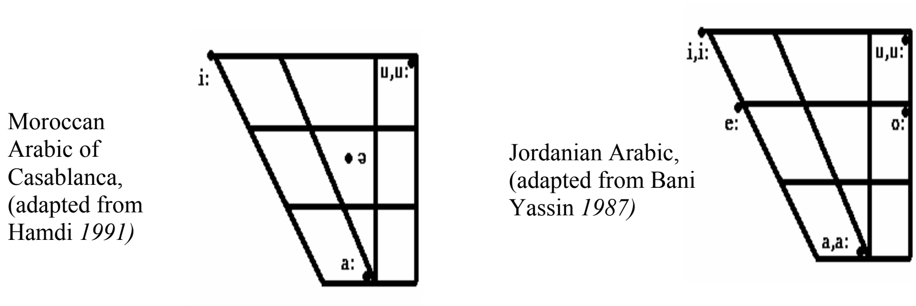

Illustration 7: Vocalic systems of JA and MA

In our corpus, vowels of JA and MA are preceded by 9 consonants /b,d,k,$t^{ʕ}$, $d^{ʕ}$, q, $s^{ʕ}$, w/ in production and /d $d^{ʕ}$/ in perception. In production, 51 items (for JA) and 35 (for MA) were presented five times to native speakers (i.e. 255 items for JA and 175 for MA). Items were recorded on PC Computer, at 22 KHz, 16 bits, mono. The experiment in perception was based on a method of formant adjustment of isolated synthetic vowels (Johnson 1993). In this experiment, two F0 are used: F0 at 120 Hz (male voice) and at 240 Hz (female voice). Subjects had to find here the best prototypic vocalic representations for 16 items (for JA) and 10 (for MA). These words were presented five times to speakers (i.e. 160 items for JA and 100 for MA). We recorded 16 speakers (8 males and 8 females) in JA and 19 speakers (10 males and 9 females) in MA.

4. DATA ANALYSIS

In production, we have extracted values of F1 (corresponding to the opening vs. closing of the jaw) and F2 (corresponding to the position of the tongue in the mouth, i.e., front vs. back) at the middle part of the vowels using *Winsnoori*® (Laprie (1999)), within two consonantal contexts i.e., /d $d^{ʕ}$/. Then, we applied the following formula to normalize female data (so as to counterbalance physiological peculiarities) ***6*ASINH(Data/600) - 1***[1] (Henton *1995)*. Then, we calculated the average and the standard deviation of all vowels' formant values.

In perception, we just converted vowel formant values from Hertz to Bark. Then we calculated the average and the standard deviation of F1 an F2 values for all vowels.

5. RESULTS

Five different types of results will be presented below both for JA and MA: Male vs. Female distinction, Distribution of long vs. short Vowel, effect of Pharyngealization on vocalic distribution, Perceptual Vocalic space is "hyper-articulated" and cross-dialectal comparison. Our working hypothesis were:

H.1. The size of females' vowel space in production is larger than males'.
H.2. In Production, short vowels are more centralized than long ones.
H.3. In production, adjacent vowels in pharyngealized context are backed and more open than in 'neutral' context
H.4. Perceptual vocalic space corresponds to a hyper-articulated space.
H.5. Both short and long vowels in MA are more centralized than JA vowels.

In perception, we think *a priori* results would follow the same patterns than in production

5.1. Male vs. Female Opposition

According to Henton (1995) (cf. part 2.1.), there is an important physiological difference between males' and females' size of vocal tract. We will study in this section this difference in production and in perception.

5.2. Male vs. Female Opposition

According to Henton (1995) (cf. part 2.1.), there is an important physiological difference between males' and females' size of vocal tract. According to the author, male vs. female differences are not only due to physiological differences, but also to sociolinguistic reasons.

---

[1] *ASINH* corresponds to the reversed hyperbolic Sinus of a number; *Data* is the formant value of F1 or F2 in Hz.

5.2.1. In Production

We found that for both JA and MA: males' vowel space is more centralized than females' ($p<0.0001$) (after normalizing females' vocalic space) (cf. illustrations 8 & 9).

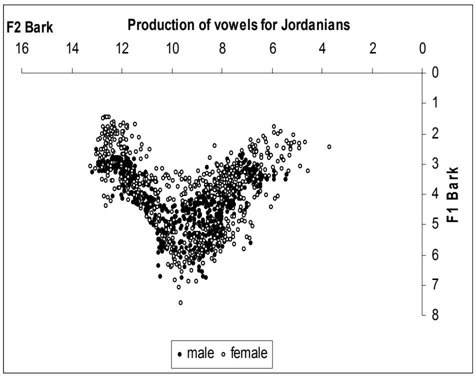

Illustration 8: Female vs. Male vocalic space for JA in production.

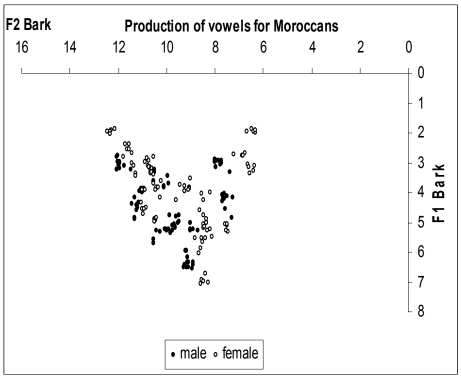

Illustration 9: Female vs. Male vocalic space for MA in production.

5.2.2. In Perception

We have applied the same statistical analysis on perceptual data and found both for JA and MA, females' vowel space is posteriorized as compared to males' ($p<0.001$) (cf. illustrations 10 & 11).

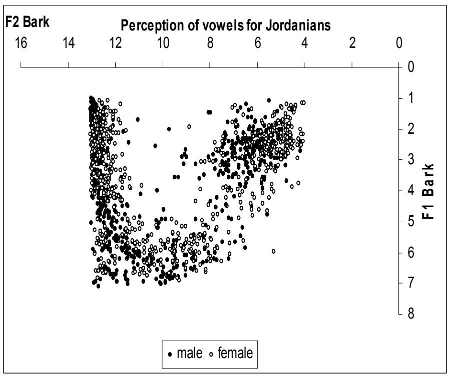

Illustration 10: Female vs. Male vocalic space for JA in perception.

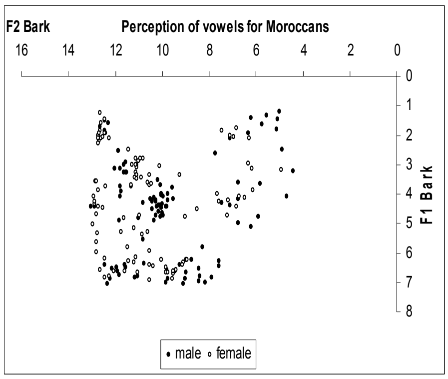

Illustration 11: Female vs. Male vocalic space for MA in perception.

5.3. Long vs. Short Vowel Opposition

Several works have studied the opposition between long and short vowels, among them: *Ghazali* 1977 and Ghazali 1979, Barkat 2000 and Barkat-Defradas 2001, Al-Tamimi 2002 and Al-Tamimi & al. 2002. Authors regularly noticed that short vowels are more centralized than long ones'. We will present here our results in production and in perception.

5.3.1. In Production

Statistical analysis confirmed short vowels are more centralized that long ones' ($p<0.0001$) both for JA and MA (cf. illustrations 12 & 13).

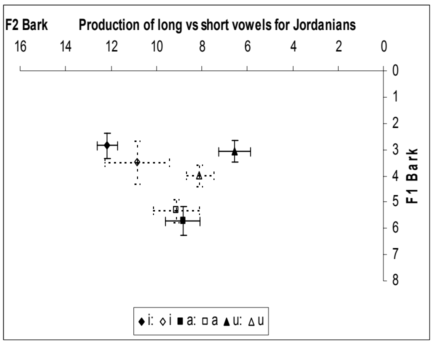

Illustration 12: Long vs. short vowels distribution for JA in production.

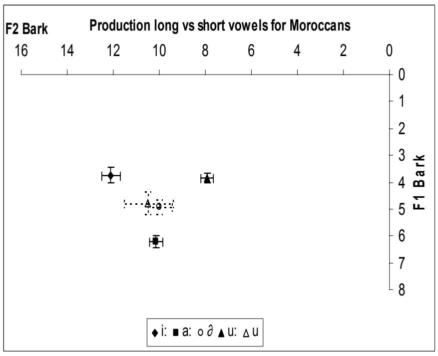

Illustration 13: Long vs. short vowels distribution for MA in production.

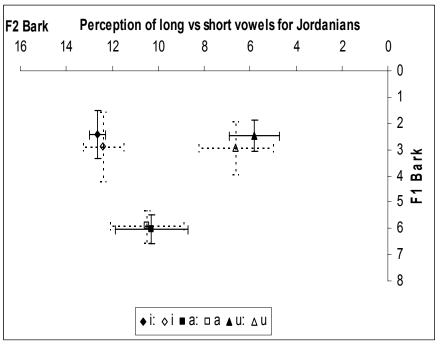

Illustration 14: Long vs. short vowels distribution for JA in perception.

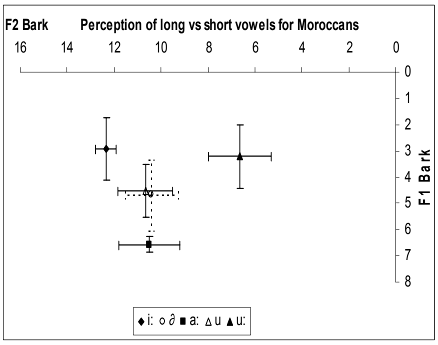

Illustration 15: Long vs. short vowels distribution for MA in perception.

We found the same pattern of distribution for JA and MA in perception ($p<0.0001$) (cf. illustrations 14 & 15).

5.4. Effect of Pharyngealization

Ghazali 1977 and Ghazali 1979, Elgendi 2001 and Al-Tamimi 2002 studied the effect of pharyngealization on adjacent vowels. They found that adjacent vowels are posteriorized and more opened in a pharyngealized context. In the present work, we will study the effect of pharyngealization on adjacent vowels both in production and perception of speech.

5.4.1. In Production

Statistical analysis on data obtained in production in JA and MA showed vowels are indeed posteriorized and more open in pharyngealized context than in neutral one (cf. illustrations 16 & 17).

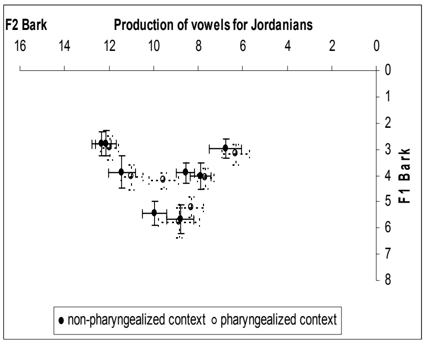

Illustration 16: Effect of pharyngealization on JAs' vowels in production.

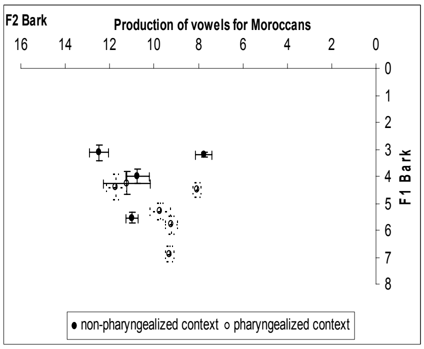

Illustration 17: Effect of pharyngealization on MAs' vowels in production.

5.4.2. In Perception

JA vowels in pharyngealized context attest a significant difference on F1 axis only (i.e. pharyngealized vowels are more open) ($p<0.0001$). We did not find any effect of pharyngealization on F2 (i.e. front ~back axis). On the other hand, we observed an effect of pharyngealization both on F1 and F2 axis in MA ($p<0.0001$) (cf. illustrations 18 & 19).

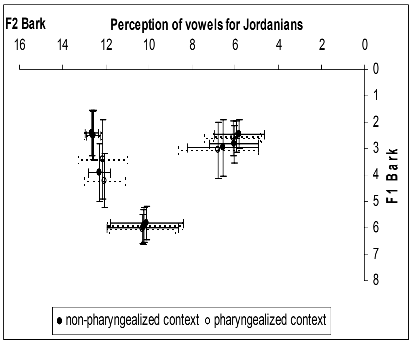

Illustration 18: Effect of pharyngealization on JAs' vowels in perception.

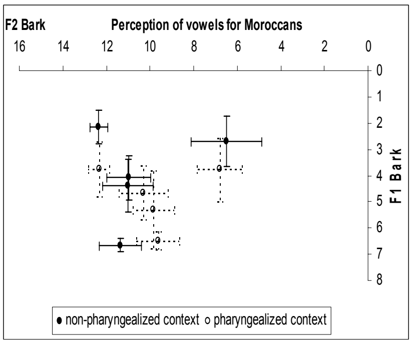

Illustration 19: Effect of pharyngealization on MAs' vowels in perception.

5.5. Perceptual vocalic Space is "Hyper-Articulated"

As done in Johnson (1993), we compared production and perception vocalic spaces. If the perceptual vocalic space is larger than the one observed in production then, we could confirm that the perceptual vocalic space corresponds as a matter of fact to a hyper-articulated space. We applied different statistical analysis in order to compare: the global vocalic distribution in production vs. Perception then for each vowel separately.

Both in JA and MA, the perceptual vocalic space is larger that it is in production ($p<0.0001$). We noticed as well that back vowels are posteriorized in perception whereas front vowels are anteriorized. As for F1 axis, it appeared that closed vowels are more closed in perception whereas open vowels are more open ($p<0.0001$) (cf. illustrations 20 & 21).

5.6. Cross-Dialectal Comparison

The idea in this section is to compare – both in production and perception – the organization of vocalic space in two dialects attesting a different number of phonological segments (8 vowels for JA, 5 for MA) so as to see if the size of vocalic inventory affects its organization in terms of distribution. *A priori*, our claim was that the fewer segments are attested in a language~dialect, the larger the distribution for each vowel could be observed. In other words, a great amount of variability for the realization of one segment could be observed without leading to categorization mistakes.

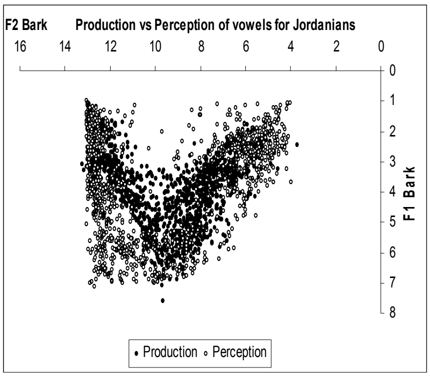

Illustration 20: Vocalic distribution in Production vs. Perception in JA.

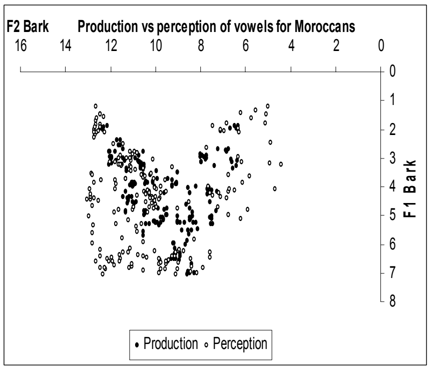

Illustration 21: Vocalic distribution in Production vs. Perception in MA.

5.6.1. In Production

Graphical representations and statistical analysis show us that there are differences in the dispersion of vowels in the Arabic dialects. MAs' vocalic space, for long and short vowels, is more centralized than JAs' ($p<0.0001$). /a:/ JA is more posteriorized than in MA ($p<0.001$). Long vowels distribution in JA is larger than in MA ($p<0.001$). In non-pharyngealized dental context, MA's /u/ and /«/ merge ($p<0.001$) (cf. illustrations 22 & 23).

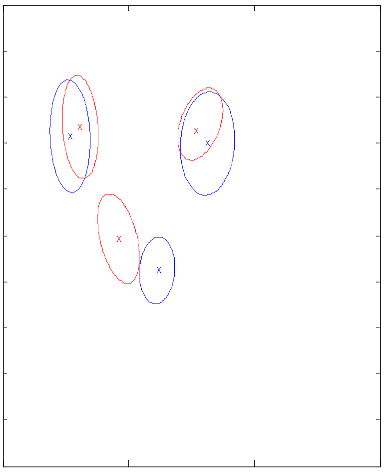

Illustration 22: JA and MA long vowels

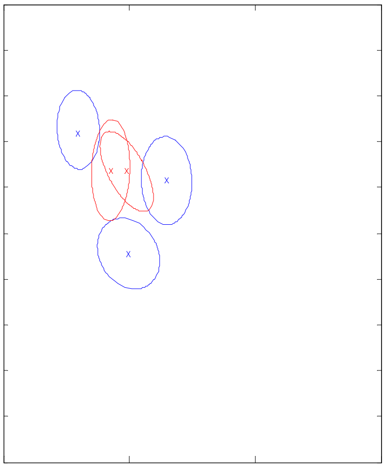

Illustration 23: JA and MA short vowels

5.6.2. In Perception

We applied the same statistical analysis to perceptual data both for JA and MA and found the same patterns of distribution: MA vowels are more centralized than JA ones ($p<0.0001$). /u/ and /«/ merge in non-pharyngealized context in MA ($p<0.002$). Though the dispersion for /u:/ is larger in MA than it is in JA ($p<0.001$) the distribution of all vocalic segments is larger in JA than in MA ($p<0.001$) (cf. illustrations 24 & 25).

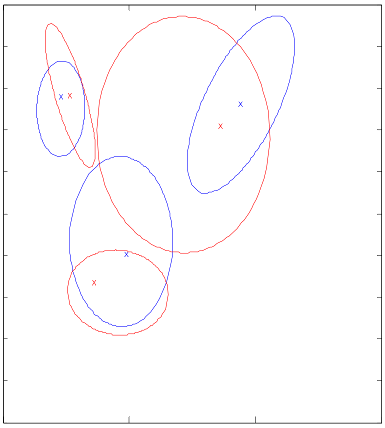

Illustration 24: JA and MA long dispersion

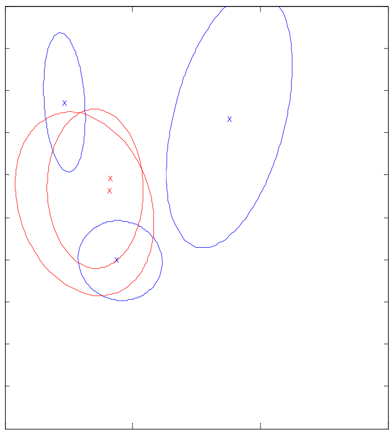

Illustration 25: JA and MA short dispersion

6. CONCLUSION

In this preliminary work, we have tested some theoretical outcomes by an original experimental method in production and perception, such as cross-gender differences, long vs. short vowels distribution, effects of pharyngealization on adjacent segments, cross-dialectal comparison and relation between production and perception.

Results revealed that for some of the issues mentioned, one can observe the same pattern of realization both in speech production and perception modality (i.e., a larger vocalic distribution both in production and perception for female subjects that can be interpreted in the light of mother's role for language transmission which may be helped by over-articulation (Henton 1995); reduction of short vowels as compared to their long counterparts both in JA and MA as already shown for Swedish by Lindblom (1963); centralized vocalic distribution in Moroccan Arabic as opposed to a peripheral dispersion in Jordanian Arabic; backing co-articulation both in JA and MA in pharyngealized context due to the inertia of articulators and the lowering of velum for the articulation of pharyngealized consonants and last but not least, our study confirms Johnson's work since we were able to observe perceptual vocalic space corresponds to a hyper-articulated vocalic triangle so as to integrate speech variability without any risks for speech comprehension.

7. REFERENCES

AL-TAMIMI, J. (2002) *Variabilité phonétique en production et en perception de la parole : le cas de l'arabe jordano-palestinien.*, Mémoire de DEA en Sciences du Langage, Université Louis Lumière, Lyon 2.

AL-TAMIMI, J., GIRARD, M. & MARSICO, E. (2002)"Variabilité inter-langue et inter-individuelle en production et en perception : étude préliminaire en arabe dialectal et en français", *24ème Journée d'Études sur la Parole*, Nancy (France), pp.169-172.

BANI-YASIN, R. & OWENS, J. (1987) "The Phonology of a Northern Jordanian Arabic Dialect", in *Zeitschrift der Deutschen Morgenlandischen Gesellschaft*, Vol. 137(2), pp. 297-331.

BARKAT, M. (2000) *Détermination d'indices acoustiques robustes pour l'identification automatique des parlers arabes*, Thèse de doctorat nouveau régime, Université Lumière Lyon 2, 300p, publiée aux Editions du Septentrion.

BARKAT-DEFRADAS, M., (2201) "Vers l'identification automatique des parlers arabes", *Langue & Linguistique*, n° 7, pp. 47-95

CALLIOPE, (1989) "La parole et son traitement automatique", *Masson* (collection ENET et ENST). 718 p.

DELATTRE, P. (1958). "Les indices acoustiques de la parole. Premier rapport" *Phonetica* 2 : 108-118, 210-251.

ELGENDY, A.M. (2001) *Aspects of pharyngeal coarticulation*, Ph.D. thesis, University of Amesterdam, LOT 44 : 312 pp.

FRY, D.B. (1966). "Mode de perception des sons du langage" dans MOLES, A. et VALLANCIEN, B. *Phonétique et phonation*, 191-206.

GHAZALI, S. (1977), *Back consonants and backing co-articulation in Arabic*, PhD dissertation, University of Texas, Austin.

GHAZALI, S. (1979), "Du statut des voyelles en arabe", *Analyses & Théorie*, 2/3, pp. 199-219.

GHAZALI, S., HAMDI, R., BARKAT, M. (2002) "Speech Rhythm Variation in Arabic Dialects" Paper presented at the *Speech Prosody 2002*, Aix-en-provence (11-13 avril 2002), pages 331-334.

HAMDI, R. (1991) *Étude phonologique et expérimentale de l'emphase en arabe marocain de Casablanca*, Thèse de Doctorat : Sciences du Langage : Lyon2. p.172.

HENTON, C. (1995) "Cross-language variation in the vowels of female and male speakers", *ICPhS*, Stockolm 4 : 421-423.

HOCKETT, C.F. (1955) "A manual of phonology", *International Journal of American Linguistics,* 21.

HOMBERT, J. M., & PUECH, G. (1984) "Espace vocalique et structuration perceptuelle : application au Swahili", *Pholia* 1 :199-208.

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

JOHNSON, K., FLEMMING, E. & WRIGHT, R. (1993) "The hyperspace effect: Phonetic targets are hyperarticulated", *Language*, 69 : 505-528.

KHOULOUGLI, D.E., (1996) "Sur quelques approches de la réalité sociolinguistique arabe", in *Langue en Egypte*, *Egypte / Monde Arabe*, n° 27-28 (3ème et 4ème trimestre), pp. 287-299.

KLUENDER, K.R., GIEHL, R. & KILLEEN, O. (1987) "Japanese quail can learn phonetic categories", *Science*, 237 : 1195-1197.

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

LIBERMAN, A. & MATTINGLY, I. (1985) "The motor theory of speech perception revised", *Cognition* , 21 : 1-36.

LINDBLOM, B. (1963) "Spechographic study of vowel reduction", *Journal of the Acoustical Society of America*, 35 : 1773-1781.

MARÇAIS, Ph. (1977) *Esquisse grammaticale de l'arabe maghrébin*, Paris : Librairie Adrien-Maisonneuve.

PETERSON, G.E. & BARNEY, H.L. (1952). "Control methods used in the study of vowels", *Journal of the Acoustical Society of America*, 24 : 175-184.

RJAIBI-SABHI, N., (1993) *Approches Historique, Phonologique et Acoustique de la Variabilité Dialectale Arabe : Caractérisation de l'origine Géographique en Arabe Standard*, Thèse de doctorat (N.R.) : Univ. de Franche-Comté.

TWADELL, W.F. (1952). "Phonemes and allophones in speech analysis" *Journal of the Acoustical Society of America*, 24 : 607-611.