

# Does vowel space size depend on language vowel inventories? Evidence from two Arabic dialects and French.

*Jalal-Eddin Al-Tamimi & Emmanuel Ferragne*

Laboratoire Dynamique du Langage (UMR 5596) (CNRS & Lyon 2)  
Institut des Sciences de l'Homme,  
14, avenue Berthelot – 69007 Lyon - France

Jalal-Eddin.Al-Tamimi@univ-lyon2.fr & Emmanuel.Ferragne@univ-lyon2.fr

## Abstract

The aim of this paper is to study the effect of the number of vowels a language has on the size of its vocalic space, in the production of speech from two Arabic dialects, Moroccan and Jordanian Arabic, and from French. 5 speakers per language (or dialect) recorded a list of vowels in 3 conditions: vowels were embedded in 1) Words, 2) Syllables, and 3) produced in Isolation. For each condition, there were 3 consonantal contexts: /b d k/. Our results corroborate the hypothesis that vowel inventory size affects the size of the acoustic vowel spaces: the larger the vowel inventory, the bigger the acoustic vowel space.

## 1. Introduction

Languages differ in the size and organization of their vowel inventories; they vary from 3 to 24 distinct vowels ([1] & [2]). Cross-linguistic investigations revealed that the general organization of vowel inventories is governed by auditory and articulatory constraints. Theoretical studies tried to predict the effect of vowel inventory size on the general organization of vowel systems. Some of them are briefly discussed in the rest of the Introduction.

Given the non-linear relationship between articulatory movements and their acoustic correlates, the Quantal Theory of speech ([3]) states that there are certain regions of stability in the phonetic space, corresponding to the point vowels [i], [a], and [u]. These point vowels (or Hot-Spots) should be in approximately the same location across all languages, independently of vowel inventory size. And since these point vowels are in phonetically stable regions, there should be less intra-category variability than for other vowels.

On the contrary, the Dispersion Theory ([4]) claims that speech sound organization is ruled by an “Adaptive Dispersion” of their elements, following a “Sufficient Perceptual Contrast” principle. According to this theory, the vowels of a given language are organized in the acoustic vowel space in such a way that they be sufficiently distinct on the perceptual level. [4] explains that phonetic values of vowel phonemes should exhibit more variation in small than in large systems.

With different visions about the general organization of vowel inventories, these two theories proposed some common universal principles to account for the cross-linguistic tendencies observed in vowel inventories. Many subsequent studies assessed the validity of these universal principles ([5], [6], [7], [8], etc.).

The studies by [5] & [7], based on a large corpus of vowel inventories (7 languages for [5] and 28 for [7]), do not

provide evidence either for an effect of inventory size or for tighter vowel areas in Hot-Spots.

In the study by [6], the author compared the acoustic vowel spaces of English (11 monophthongs) and Spanish (5 vowels) in CVC and CVCV sequences. She found that: 1) the location of similar vowels in the acoustic vowel spaces was determined, in part, by a language-specific base-of-articulation, 2) the English crowded vowel system occupies a greater space than that of Spanish (she notes that this effect depends on the syllabic structure of English vowels), and 3) there is no difference in the tightness of within-category clusters for large versus small vowel inventories.

In the work by [8], the authors investigated the acoustic dispersion of vowels, in both production and perception, in 3 languages with different vowel inventories: English with 11 vowels, French with 10 vowels, and Spanish with 5 vowels. The corpus was based on realizations of vowels in isolation. Their findings do not support the hypothesis that the density of vowel systems influences the size of the acoustic space in production, but it does influence the vocalic space in perception.

The aim of our study is to investigate the effect of vowel inventory size on the general organization of acoustic vowel spaces, in two languages: Arabic (2 dialects) and French. Based on the theoretical claims mentioned above, we can formulate the following hypotheses:

- H1. Larger vowel inventories result in larger acoustic spaces,
- H2. Point vowels should be in approximately the same location across languages, and should have a lower intra-category variability compared to the non point vowels,
- H3. Vowel areas may be dependent on vowel inventory size: the smaller the vowel inventory, the greater the acoustic dispersion per vowel type, and vice-versa,
- H4. Acoustic vowel spaces in the Word condition will cover a smaller area than in the Syllable or in the Isolation conditions, according to our interpretation of the H&H hypothesis proposed by [9]: when producing vowels in isolation, speakers should give the maximum precise acoustic cues about the vowel identity to minimize confusion with other vowel types (which leads to minimal intra-category variability), and when producing vowels in words, speakers give less information about the vowel identity, because it is already included in the surrounding context, resulting in a vowel undershoot phenomena.

## 2. Method

### 2.1. Speech Material

Arabic (with two dialects: Jordanian Arabic with /i i: e: a: o: u u:/ ([10]) and Moroccan Arabic with /i: ə a: u u:/ ([11])), and French with /i e ε a α ɔ o y ø œ/ (JA, MA and FR henceforth, respectively) were compared. 5 male speakers per language (or dialect; i.e. 15 subjects in all, aged 20 to 30) recorded a list of items (with C<sub>1</sub>VC, C<sub>1</sub>VCV, and C<sub>1</sub>VCVC). C<sub>1</sub> was one of the 3 phonologically common consonants between the two languages: /b d k/ and each vowel. The items were randomly presented with 5 repetitions in an adapted carrier sentence (for Arabic, the Modern Standard Arabic script was used without vocalization). The speakers were asked to pronounce these items as realized in Word, Syllable and in Isolation (e.g. [bo:se ~ bo: ~ o:] = “a kiss”). The recordings were made in a sound-attenuated room, on a PC, with 22050 Hz, 16 bits, mono. We ended up with 1106 vowels for MA, 1645 for JA, and 2390 for FR (all consonantal and realization contexts).

### 2.2. Data Analysis

The data were segmented manually and acoustic measurements of the first 4 center formant frequencies at temporal midpoint were carried out with Praat, using 23 LPC coefficient (autocorrelation), with a 16 ms Gaussian window. Formant values were then converted to Barks using the formula proposed by [12].

## 3. Discussion

### 3.1. Vowel Inventory Size

In order to investigate the effect of vowel inventory size on the size of the acoustic vowel spaces, we calculated the dispersion area (Convex Hull) of: 1) all vowels per language, 2) [i a u] (long vowels in JA & MA) to compare directly between common vowels. Figures 1, 2 & 3 show the mean and standard deviation of [i a u] per language and per condition (Word, Syllable and in Isolation). The graphs below show that in the three conditions, the MA vowel system is more reduced than that of JA or FR, and that JA and FR acoustic vowel spaces seem more similar.

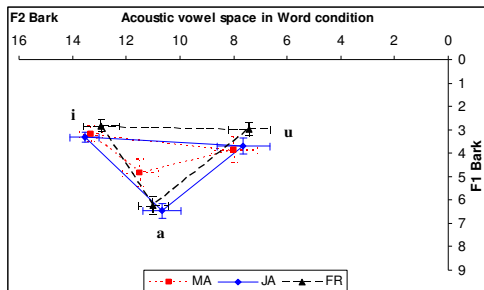


Figure 1: [i a u] acoustic vowel spaces in Word condition.

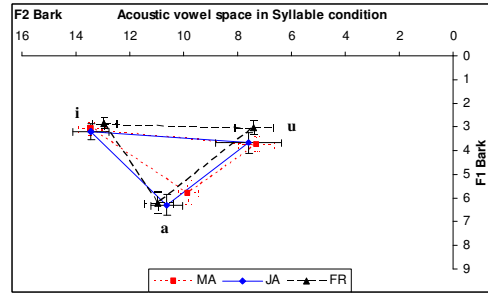


Figure 2: [i a u] acoustic vowel spaces in Syllable condition.

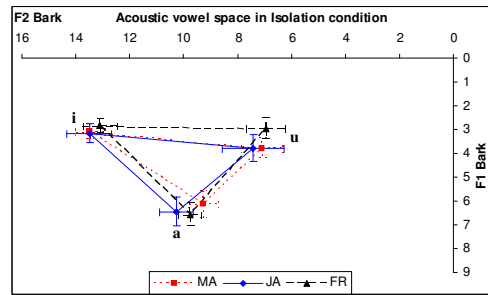


Figure 3: [i a u] acoustic vowel space in Isolation condition.

Table 1 & 2 give the size of dispersion areas with reference to MA for all vowels and [i a u] respectively.

Table 1: Comparison of the size of dispersion areas with reference to MA for all vowels, in the 3 conditions.

	All Vowels		
	MA	JA	FR
Word	100%	246%	261,78%
Syllable	100%	124,25%	141,36%
Isolation	100%	123,44%	165,45%

Table 2: Comparison of the size of dispersion areas with reference to MA for [i a u], in the 3 conditions.

	[i a u]		
	MA	JA	FR
Word	100%	231,04%	229,62%
Syllable	100%	118,48%	115,62%
Isolation	100%	109,31%	125,04%

The comparison between areas in the three conditions show that, the MA vowel space is smaller than that of JA and FR, whether the calculation is based on all vowels or on [i a u], and that vowel spaces in JA and FR are almost the same in both contexts (all vowels and [i a u]) (Figure 4 & 5).

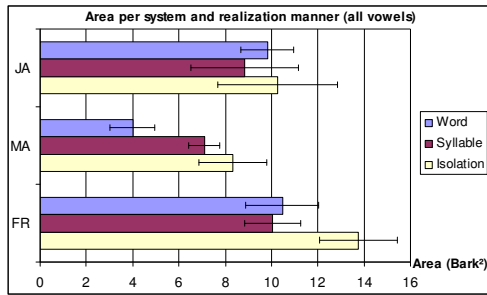


Figure 4: Mean areas and Standard deviations per language and condition, for all vowels.

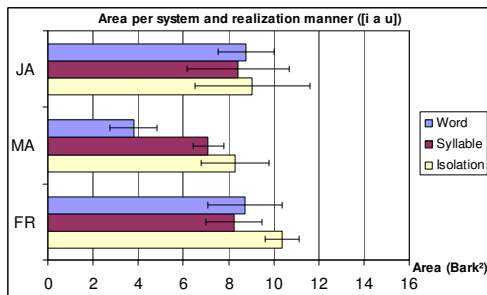


Figure 5: Mean areas and Standard deviations per language and condition, for [i a u].

Here, we can see that there is an effect of vowel inventory size on the size of the acoustic vowel spaces between the 3 languages (on both metrics: all vowels and [i a u]). Therefore, our results confirm H1. Concerning the condition (Word, Syllable and Isolation) effect on acoustic spaces, our results for MA are conclusive; we can see that the areas are smaller for the Word condition than for Syllable or Isolation. For JA and FR, the acoustic vowel space in Isolation is bigger than that in the other 2 conditions, but the vowel space in the Syllable condition is smaller than that in the Word condition. Overall, our results confirm the issue addressed by [9] (and then H4): when a speaker produces vowels in isolation, s/he may *Hyper-Articulate* the vowel in order to eliminate any confusion with other vowel types, and that for vowels produced in words, or in syllables, the speaker will *Hypo-Articulate* because the identity of vowels is determined by the context.

### 3.2. Vowel Areas

In this section, we address the following issues: 1) do the 3 point vowels [i a u] occupy the same position whatever the language? [3] and 2) do vowel inventory size affect vowel space size? We calculated the dispersion area (Convex Hull) per vowel type for each language and condition. Figures 1, 2, and 3 show that the three point vowels are not in the same position in the acoustic vowel spaces (in Figure 1, the MA [a] is more centralized than in JA or FR, but for Figures 2, and 3, there is a slight difference between the point vowels in the three languages).

In order to compare the within-category variability between the three point vowels in the 3 languages, we calculated areas per vowel type, per language, and per

condition (Word, Syllable and Isolation). Figures 6, 7, and 8 show that: 1) generally, [i] across the 3 languages has the smallest dispersion, 2) [u] in MA and FR has little variability, but 3) [a] has greater variability across the 3 languages. So our results can partly confirm the issues addressed by [3] (H2).

Then, in order to investigate the effect of vowel inventory size on the tightness of the vowel areas (H3), we used the same calculation presented in Figures 6, 7, and 8. Our results do not confirm this hypothesis. We can see that the dispersion areas per vowel depends in part on the condition (Word, Syllable, and Isolation), and seem to depend on individual strategies.

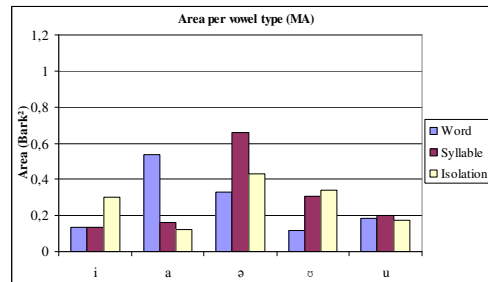


Figure 6: MA vowel dispersion areas.

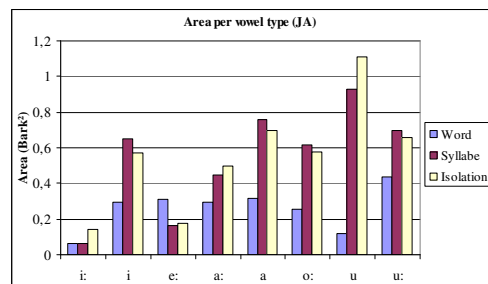


Figure 7: JA vowel dispersion areas.

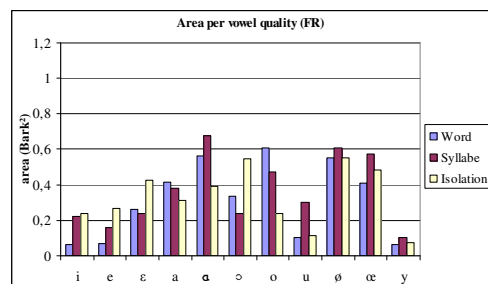


Figure 8: FR vowel dispersion areas.

### 3.3. [i a u] distances

The next step was to assess whether simple metrics related to vowel space size would allow reasonable separability between languages. Centre formant frequencies at vowel temporal midpoint for [i], [a], and [u] were computed for all occurrences of these vowels in all 3 languages. Distances

were measured between all pairs of vowel tokens of the vowel types in a given speaker, yielding the following metrics:  $i\sim uFx$ ,  $i\sim aFx$ ,  $a\sim uFx$  (the first item reads: the distance between one occurrence of vowel [i] and one occurrence of vowel [u] for formant  $x$  [ $x$  takes the values 1 and 2] in Bark). On average, 350 distance vectors of size 6 were obtained per language. Comparisons were then carried out across languages for each condition (Word, Syllable, and in Isolation). We performed principal component analysis (PCA) as a preliminary check for separability.

Figure 9 shows a biplot of each reduced distance vector (after PCA) for the Word condition in a plane defined by the first two components, which account for approximately 73% of the original variance. Note that distances representing vowel space size in MA are clearly distinguished from the rest. The other two conditions (vowel in Isolation and Syllable) showed much more class overlap. From this display, it was hoped that a neural network could be trained for classification. A multilayer perceptron with 4 hidden neurons was trained with data from the first three speakers of each language. Data from 2 speakers per language were held out for the test phase.

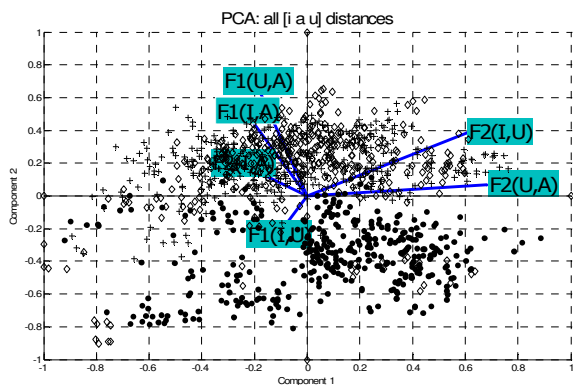


Figure 9: Biplot displaying first two PCA components of all [i a u] distances; labelled plain lines show correlation of original variables with components (e.g. F2(I,U), the distance between [i] and [u] in the F2 dimension, highly contributes to component 1). JA: +; MA: •; FR: ◊.

Table 3 is a confusion matrix that shows the performance of the model. The network achieves on average more than 88% correct classification. Note that all instances of MA are correctly classified; approx. 20% of FR items are incorrectly classified as JA, and most misclassifications of JA items are confused with FR. These results confirm that, in terms of vowel space size, FR is closer to JA than is MA. Besides – although the automatic classification scores can be called "preliminary", given the number of speakers involved – it seems that a rather simple metric that gauges vowel space size could be used for automatic language identification.

Table 3: Confusion matrix for classification.

	JA	MA	FR
JA	119	6	25
MA	0	150	0
FR	29	0	121

## 4. Conclusion

In this paper, we have investigated the effect of vowel inventory size on the size of the acoustic vowel spaces in three languages differing in their vowel systems (MA, JA, and FR). Our results show that: 1) the FR vowel space is larger than that of JA or MA (H1), 2) the point vowels seem to have approximately the same position in the acoustic vowel spaces across the 3 languages in only two conditions (Syllable and in Isolation, but not in Word) (H2), 3) the tightness of vowel type areas cannot be demonstrated by the current data, and more speakers need to be analyzed, and finally 4) a linguistically interpretable metric for vowel space size may be used in automatic language/dialect identification.

## 5. Acknowledgements

We are grateful to François Pellegrino and Egidio Marsico for their help.

## 6. References

- [1] Maddieson, I., *Patterns of Sounds*, Cambridge: Cambridge University Press, 1984.
- [2] Vallée, N., *Systèmes vocaliques : de la typologie aux prédictions*, thèse de Doctorat en Sciences du Langage, Université Stendhal, Grenoble, 1994.
- [3] Stevens K.N., "On the quantal nature of speech", *J. of Phonetics*, Vol. 17, p. 3-46, 1989.
- [4] Lindblom, B., Phonetic universals in vowel systems, *In Experimental Phonology*, Edited by J. Ohala and J. Jaeger, Academic New York, p. 13-44, 1986.
- [5] Engstrand, O. and Krull, D., "Effect of inventory size on the distribution of vowels in the formant space: preliminary data for seven languages", *PERILUS*, Vol. 13, p. 15-18, 1991.
- [6] Bradlow, A., "A comparative acoustic study of English and Spanish vowels", *J. Acoust. Soc. Amer.*, Vol. 97 (3), p. 1916-1924, 1995.
- [7] Livijn, P., "Acoustic distribution of vowels in differently sized inventories - hot spots or adaptive dispersion?" *XIIIth Swedish Phonetics Conference Proc. (FONETIK 2000)*, Skövde, Sweden, May 24-26, p. 93-96, 2000.
- [8] Meunier, C., Frenck-Mester, C., Lelekov-Boissard, T. and, Le Besnerais, M., "Production and perception of vowels: does the density of the system play a role?", *15th ICPHS*, p. 723-726, 2003.
- [9] Lindblom, B., *Explaining phonetic variation: A sketch of the H&H theory*, In *Speech production and speech modeling*, edited by W. Hardcastle & A. Marchal (Kluwer Academic, Dordrecht), p. 403-439, 1989.
- [10] Bani-Yasin, R. & Owens, J., "The Phonology of a Northern Jordanian Arabic Dialect", in *Zeitschrift der Deutschen Morgenlandischen Gesellschaft*, Vol. 137(2), p. 297-331, 1987.
- [11] Hamdi, R., *Etude phonologique et expérimentale de l'emphase en arabe marocain de Casablanca*, thèse de Doctorat en Sciences du Langage à Lyon2. p.172, 1991.
- [12] Schroeder, M.R., Atal, B.S., & Hall, J.L., "Optimizing digital speech coders by exploiting masking properties of the human ear", *J. Acoust. Soc. Amer.*, Vol. 66, p. 1647-1652, 1979.