Static and Dynamic Cues in Vowel Production: A Cross Dialectal Study in Jordanian and Moroccan Arabic

Jalaleddin Al-Tamimi

Laboratoire Dynamique du Langage (DDL) – UMR 5596 CNRS & Université Lyon 2 – France.

Jalal-Eddin.Al-Tamimi@univ-lyon2.fr -- http://www.ddl.ish-lyon.cnrs.fr/

Abstract

The aim of this paper is to examine the role of dynamic cues (i.e. formant slopes obtained from a linear regression analysis) in comparison with static one (i.e. vowel targets) in the classification of Jordanian and Moroccan vowels, using Discriminant Analysis. 10 speakers per dialect produced a list of vowels in C1VC2, C1VC2V, or C1VC2VC words, where C1 and C2 were either /b/, /d/, /dˤ/ or /k/, and V, each vowel. Results show the possibility of vowel separation between both dialects for a specific consonantal environment. Using dynamic cues improves the correct classification rates of about 5% for Moroccan Arabic and 13% for Jordanian Arabic.

Keywords: Arabic dialects, vowel production, formant slopes, vowel targets, classification.

Introduction

Vowel targets, produced in isolation, are considered as the canonical form of vowels (Joos (1948), among others). However, they must be considered as a *"Laboratory Artefact"*, (Liberman *et al.* (1967)), because: 1) vowels are mostly produced in coarticulation with consonants according to various syllabic structures, and 2) vowel formants are highly instable due to intra- & inter-individual variability. Some researchers (Strange (1989), among others) have described vowels produced in isolation as different from those produced in context, concluding that listeners use different cues to identify vowels in isolation or in context. Thus, they have considered these isolated vowels as "useless" for the identification~discrimination experiments and that dynamic information (formant movements and transitions) are more useful in speech perception.

The aim of this paper is to evaluate the role of static and dynamic cues in the classification of Arabic vowels by Discriminant Analysis. One of the motivations of this work is that the morphological structure of Arabic (a non-concatenative language with a triconsonantal root that exhibits direct consonant~consonant relations (McCarthy (1979); Pierrehumbert (1992))) implies that vowels never occur in isolation. We have shown that Arabic speakers have difficulties to produce and perceive vowels in isolation. Preliminary results show that dynamic cues (formant transitions) improve the perception of Arabic vowels, (Al-Tamimi (2007)).

We propose to compare the vowel systems of two Arabic dialects: Jordanian and Moroccan Arabic in terms of their static and dynamic representations. The static one is a description of vowel targets at the temporal mid-point; the dynamic one is a representation of vowels by their formant slopes, calculated from onset to temporal mid-point, and obtained from a linear regression analysis. The evaluation of dynamic cues role will be conducted in the basis of vowel classification by Discriminant Analysis. The next step of this research will be to examine the role of these dynamic cues in perception (Al-Tamimi (2007)).

Speech material

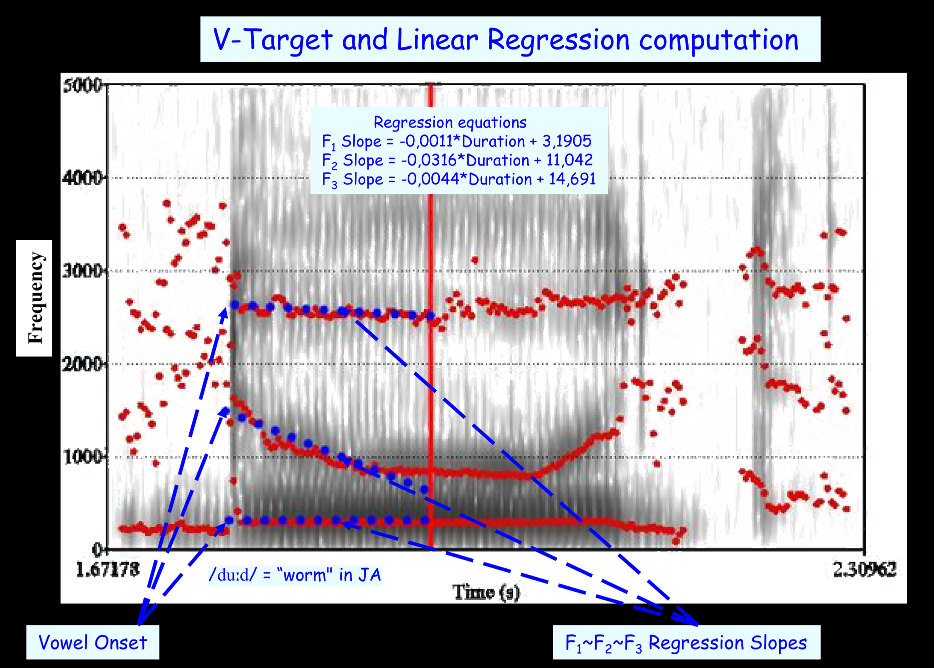
- Moroccan Arabic (MA) from Casablanca, with a 5 vowel system: /iː ə aː ʊ uː/ (Hamdi 1991),
- Jordanian Arabic (JA) from Irbid, with a 8 vowel system: /i iː eː a aː oː u uː/ (Bani-Yassin & Owens 1987),
- 10 male speakers per system: age ➔ 20 to 30, without articulatory impairment, audiometry ok.
- List of items in C_1V, C_1VC_2V ou C_1VC_2VC structures, where C_1 and/or C_2 were /b d dˤ k/,
- Items presented randomly with 5 repetitions per speaker in an adapted carrier sentence (the Modern Standard Arabic script was used without vocalization).
- Recordings were made in a sound-attenuated room, on a PC, with 22050 Hz, 16 bits, mono. We ended up with 986 vowels for MA, and 1432 for JA

Data analysis

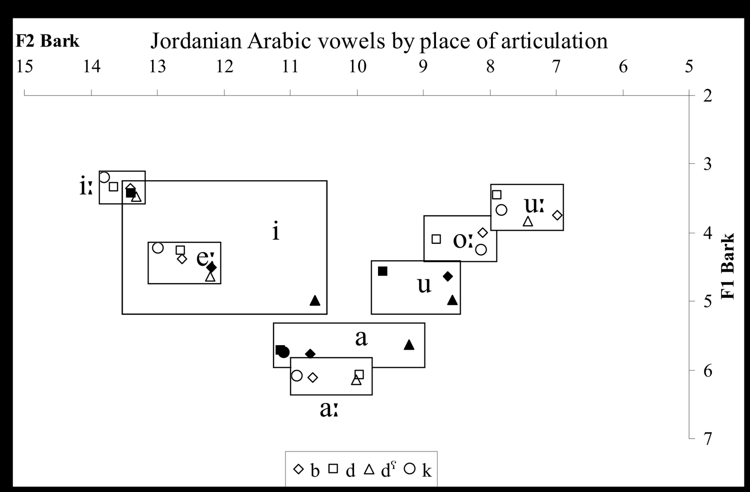
- F_1, F_2 & F_3 were computed with Praat, using "Burg" algorithm with a 12.5ms Gaussian window, and a 5ms step. Formant values extracted every 5 ms were verified manually to prevent automatic error extraction values, and then converted to Barks (Schroeder *et al.* 1979),
- Vowel onset = values obtained 5 msec after the transition release (Al-Tamimi (2004)),
- Vowel Target = values at 50% of vowel duration
- Formant Slopes = linear regression computation from Onset to Target using the formula:

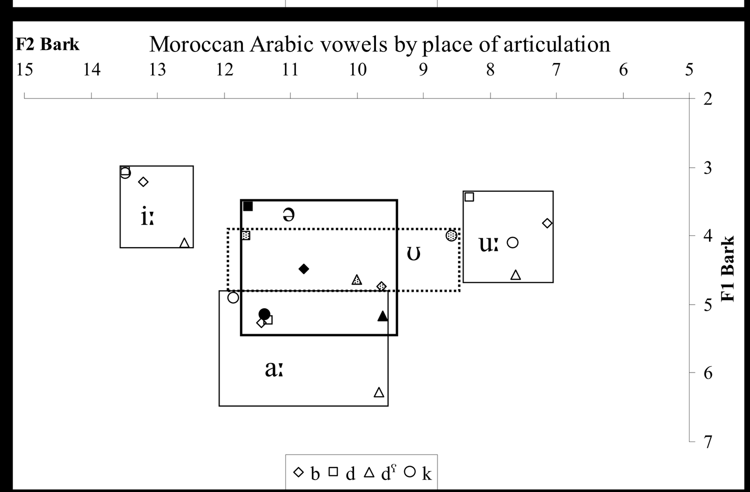
$$Formant_{Slope} = m * Vowel_{Duration} + b$$

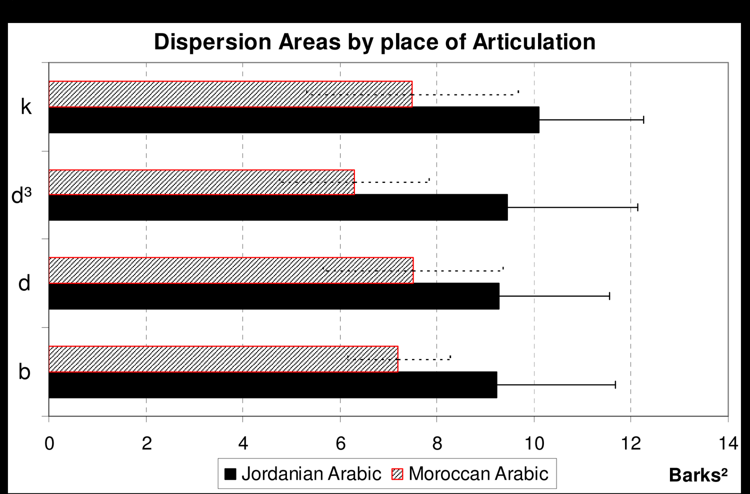
- m = transition slope value, b = intercept
- MANOVA & Discriminant Analysis.



Static cues







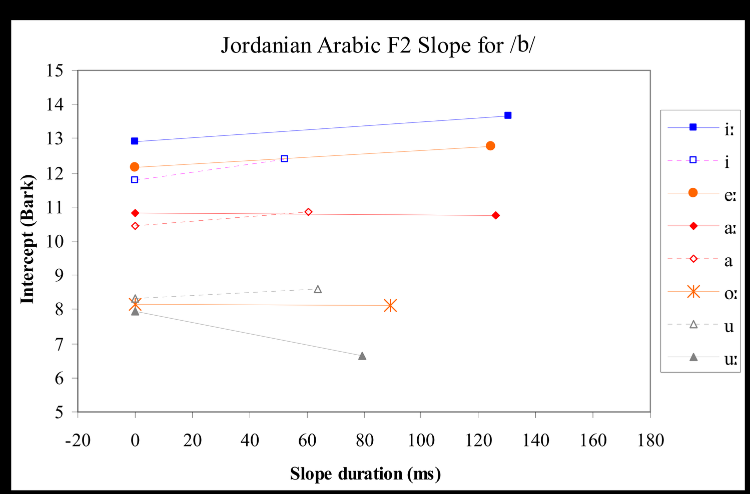
	MA					JA							
	iː	aː	ə	u	uː	iː	i	eː	aː	a	oː	u	uː
b	0,65	1,39	0,42	4,23	0,94	0,53	1,17	0,77	1,94	1,05	1,28	0,73	1,51
d	0,58	1,10	1,16	0,62	0,59	0,75	0,73	1,19	1,79	0,93	1,50	0,52	0,83
dˤ	0,82	0,76	0,77	2,60	1,62	0,64	0,78	1,06	0,69	2,75		0,65	1,02
k	0,45	1,03	2,61	1,31	1,20	0,49		1,29	1,57	0,67	0,77		2,17

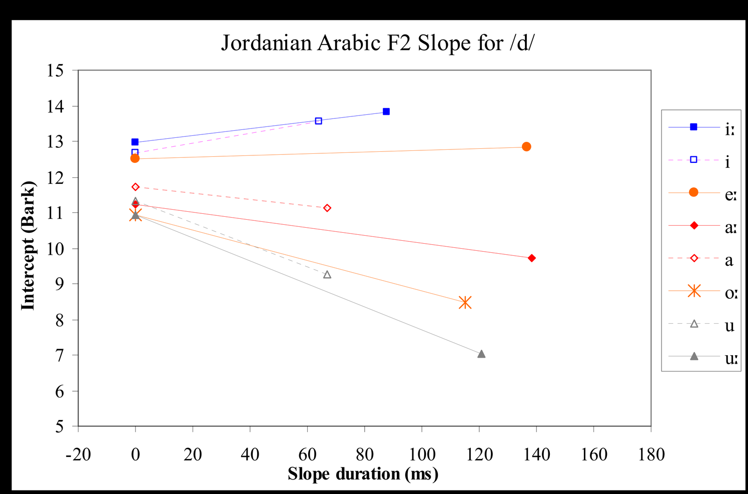
Discriminant Analysis

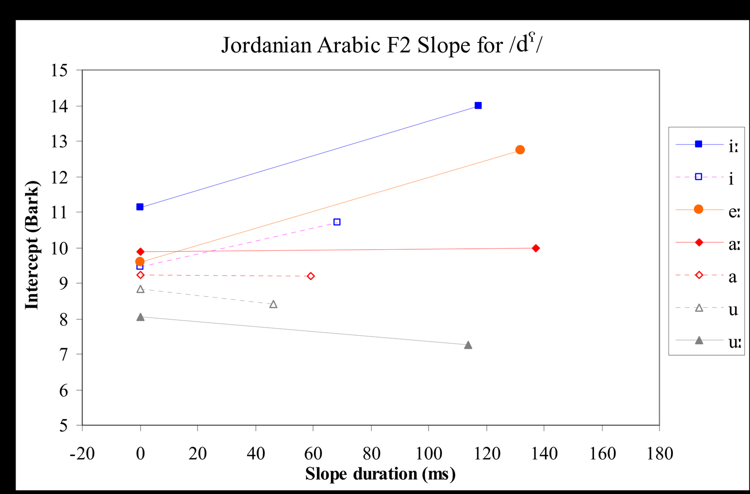
44.2% for MA vowels
32.9% for JA vowels
54.9% between JA & MA vowels.
56.1% between JA & MA vowels in /b/
62.5% between JA & MA vowels in /d/
49.6% betweens JA & MA in /dˤ/
56.3% between JA & MA in /k/
Confusions: merging of MA's /ə/ & /ʊ/, and proximity of JA's /i u/ to /eː oː/,

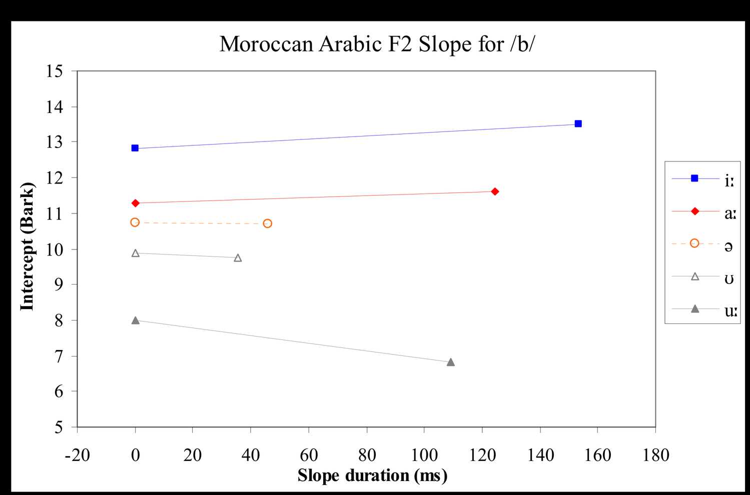
	/b/	/d/	/dˤ/	/k/
MA	82,70%	83,50%	80,40%	75,00%
JA	68,10%	69,70%	83,20%	78,40%

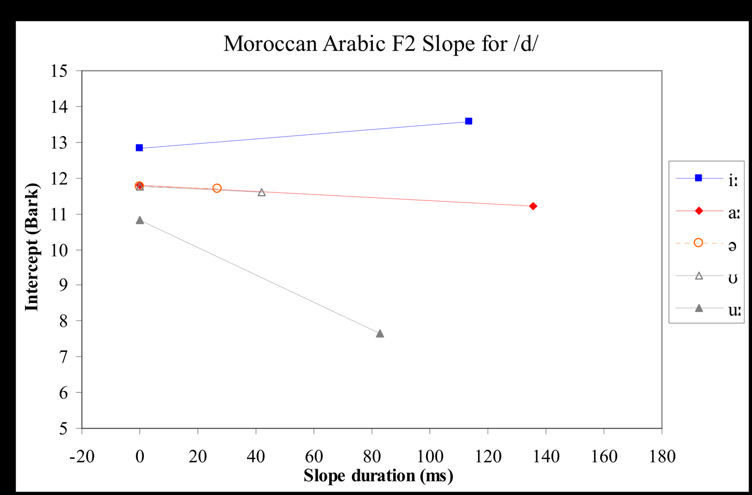
Dynamic cues

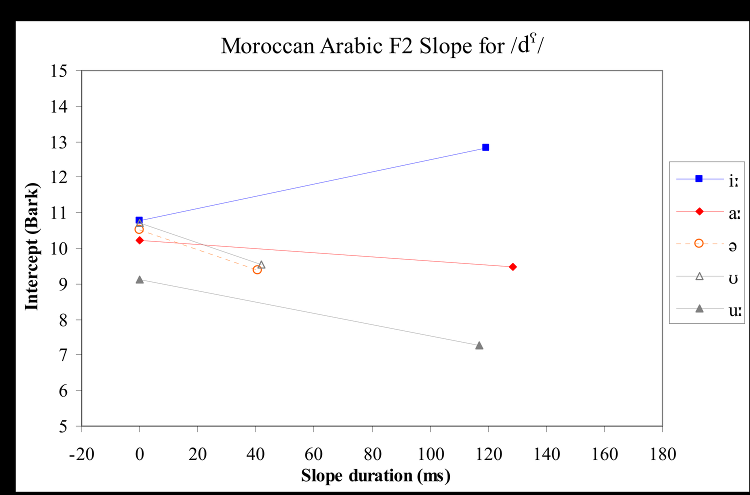


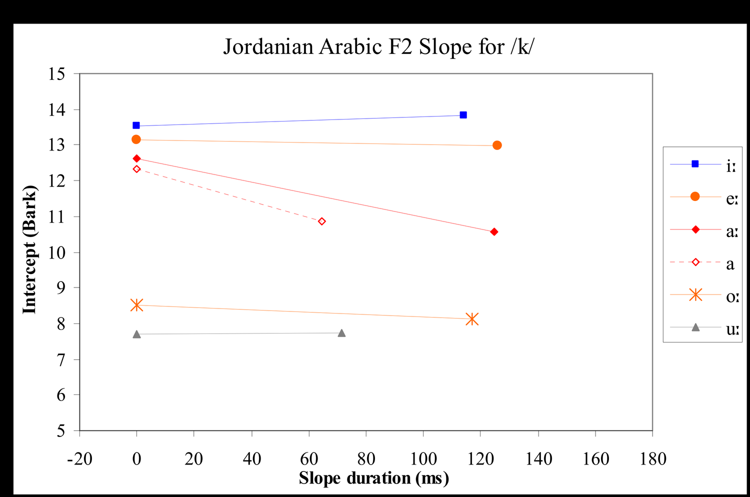


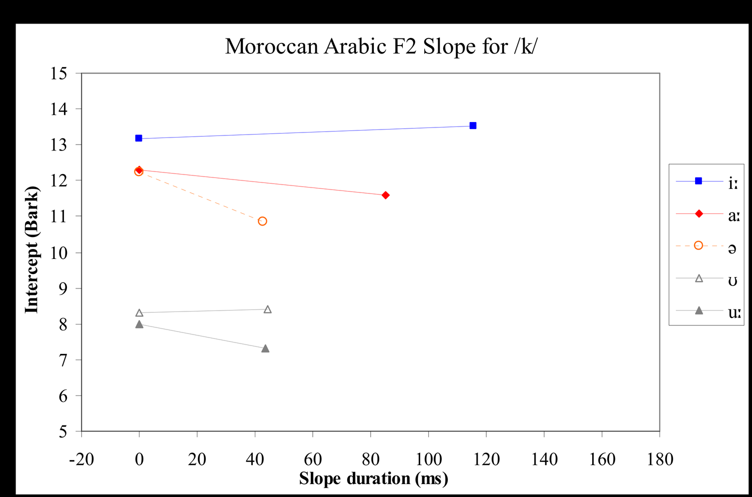












Formant shifts obtained from linear regression coefficient:
JA /iː/ in /dˤ/ environment with a slope duration of 117ms has the following formant shifts: -0.41, 2.85 & 0.22 Barks for F_1, F_2 & F_3 respectively.

	F1	F2	F3
Slope	-0,003	0,024	0,002
Intercept	3,803	11,126	14,898
Formant shift	-0,410	2,852	0,215

Discriminant Analysis

52.7% for MA vowels
54.3% for JA vowels
58.5% between JA & MA vowels.
58.5% between JA & MA vowels in /b/
63.5% between JA & MA vowels in /d/
78.0% betweens JA & MA in /dˤ/
62.5% between JA & MA in /k/
Confusions: merging of MA's /ə/ & /ʊ/, and proximity of JA's /i u/ to /eː oː/,

	/b/	/d/	/dˤ/	/k/
MA	91,20%	88,30%	76,00%	87,20%
JA	87,10%	86,10%	89,00%	92,20%

References

Al-Tamimi, J. 2004. L'équation du locus comme mesure de la coarticulation VC et CV : Étude préliminaire en Arabe Dialectal Jordanien. *Proc. 25ème Journée d'Études sur la Parole*, 9-12., **Al-Tamimi, J. & Ferragne, E.** 2005. Does vowel space size depend on language vowel inventories? Evidence from two Arabic dialects and French. In *Proc. 9th EUROSPEECH*, 2465-2468., **Al-Tamimi, J.** (2007, to appear). *Indices dynamiques et perception des voyelles : étude translinguistique en arabe dialectal et en français.* Unpublished PhD Dissertation, University Lyon 2., **Bani-Yasin, R., Owens, J.** 1987. The Phonology of a Northern Jordanian Arabic Dialect. *Zeitschrift der Deutschen Morgenlandischen Gesellschaft,* 137(2), 297-331., **Boersma, P., Weenink, D.** 2006. Praat. Doing Phonetics by Computer. version 4.4.34., **Hamdi**, R. 1991. *Étude phonologique et expérimentale de l'emphase en arabe marocain de Casablanca.* PhD. Dissertation, Université Lyon 2, 172. **Joos, M.A.** 1948. Acoustic phonetics. *Language* 24 (Suppl.), 1-136., **Liberman, A. M., Cooper, F. S., Shankweiler, D. P., Studdert-Kennedy, M.** 1967. Perception of the speech code. *Psychology Revue*, 74, 431-461., **Johnson, K., Flemming, E., Wright, R.** 1993. The hyperspace effect: Phonetic targets are hyperarticulated, *Language,* 69, 505-528., **McCarthy**, J. 1979. *Formal Problems in Semitic Phonology and Morphology.* Ph.D. dissertation, MIT., **Pierrehumbert, J.** 1992. Dissimilarity in the Arabic Verbal Roots. *Proc. 23rd Meeting of the Northeastern Linguistic Society, Graduate Student Association, U. Mass. Amherst.* 367-381., **Schroeder, M.R., Atal, B.S., Hall, J.L.** 1979. Optimizing digital speech coders by exploiting masking properties of the human ear. *J. Acoust. Soc. Am.* 66, 1647-1652., Strange, W. 1989. Evolving Theories of Vowel Perception", *J. Acoust. Soc. Am.* 85(5), 2081-2087.