SPEECH CLARITY AND COARTICULATION IN MODERN STANDARD ARABIC AND DIALECTAL ARABIC

Mohamed Embarki[a], *Slim Ouni*[b] *& Fathi Salam*[a]

[a]Laseldi EA 2281, University of Franche-Comté, Besançon, France;
[b]LORIA, University of Nancy II, France
mohamed.embarki@univ-fcomte.fr; slim.ouni@loria.fr; fathi.salam@univ-fcomte.fr

ABSTRACT

This study deals with the co-variation of speech clarity and coarticulatory patterns. Two experiments were conducted to investigate the influence of two parameters on coarticulation, the speech style (formal *vs.* non formal) and the prosodic position (stressed *vs.* unstressed syllable). The speech material was composed of four word lists varying CV syllable contexts with pharyngealized /tˤ dˤ sˤ ðˤ/ *vs.* non-pharyngealized consonants /t d s ð/ in Modern standard Arabic and dialectal Arabic. Acoustic and EMA materials were analyzed. The results revealed evident relationship between speech clarity and coarticulation: more coarticulation in formal speech (MSA) and under stressed syllable.

Keywords: Arabic, coarticulation, speech clarity, locus equation, pharyngealization

1. COARTICULATORY PATTERNS

Literature on speech production reveals systematic differences in the spatiotemporal characteristics of coarticulation (*Coa*) [22]. These differences are related to the phonological inventory [14, 19, 20, 23], more particularly to the language-specific contrasts [27, 29]. Comparison of VCV sequences in three languages showed differences in the patterns of *Coa* that were attributed to language-specific requirements on the tongue body [27]. Literature [28, 29] reported similar differences in coarticulatory patterns (*CoaP*) between clear *vs.* velarized /l/ in 3 languages, these patterns were found to be the result of specific lingual constraints imposed by each kind of lateral. *Coa* was shown to be more reduced in high density vowel systems [24]. However, other studies showed conflicting data with free vowel variation in small inventory-size systems [4, 5], another aspect, *i.e.* prosodic properties, was described as responsible of the magnitude of *Coa* [4]. *CoaP* were found in a large scale to be language-specific influenced by the prosodic features, like stress [10, 11]. This provides evidence that stressed syllables, produced with minimal gestural overlap, are characterized by stronger coarticualtory effects (*CoaE*) on the neighboring syllables [21], while unstressed ones, produced with greater gestural overlap [6], offer less resistance to *CoaE* [11].

Little is known about variation in speech clarity related to *CoaP*. Nevertheless, the literature showed that the decrease of speech clarity, by a higher speech rate, results in more overlap between successive articulatory gestures [16]. However, lesser *CoaE* were found with clear speech [25]. In this style of speech, the overlap is yet minimal between successive articulatory gestures.

2. COARTICULATION IN ARABIC

Modern Standard Arabic (MSA) and Dialectal Arabic (DA) with a contrast of Pharyngealization (*Pha*) and a low density vowel system can offer an interesting case-study for *Coa*. In MSA, there are three cardinal-like vowels (short /i a u / and long /iː aː uː/), some of regional DA has mid-closed vowels [7]. *Pha*, as a secondary articulation, exists in all varieties of Arabic with two sets of consonants (with some regional variation): pharyngealized consonants (*Ph*) /tˤ dˤ sˤ ðˤ/ and non-pharyngealized cognates (*n.Ph*) /t d s ð/. *Pha* has been associated with a set of articulatory adjustments [1, 2, 3, 13] and acoustic cues [8, 13, 15]. The most articulatory features are tongue retraction and concavity, its back raising, lip rounding and spreading [2, 9]. The most salient acoustic cues are *F1* raising and *F2* lowering [2, 8, 13, 15].

The low density vowel space and the requirements imposed on the tongue body for the *Ph* articulation will produce specific patterns of *Coa* in Arabic. These patterns will be investigated in Arabic (MSA and DA) in relation with speech clarity. Two main hypotheses relating to speech clarity are checked: 1) the alternation MSA (described as formal speech) *vs.* DA (described as non formal speech) will produce specific *CoaP* for each style; 2) different prosodic positions (stressed *vs.* unstressed syllable) will be conveyed by different patterns of coarticulation. Each hypothesis will be investigated by a separate experiment.

3. MATERIALS AND MEASUREMENTS

Two types of data were collected, acoustic and Electromagnetic articulography (EMA) data. For the acoustic data, locus equations (*LE*) were performed. *LE*, suggested by [18; *cf.* 31 for a review], are linear regression functions derived by plotting onsets of *F2* transitions (*F2onset*) [y=axis] of different vowels to their *F2* steady states (*F2mid*) – *F2onset* =k**F2mid*+c (where *k* and *c* are slope and intercept, respectively). A relatively flat slope indicates minimal vowel *CoaE*, in which case *F2onset* is not sensitive to the nature of the following vowel (*i.e.* maximal coarticulatory resistance of the consonant articulation to vowel effects). On the other hand, a relatively steep slope indicates maximal *Coa* of the consonant with the vowel as *F2onset* and *F2mid* tend to have the same frequency (minimal coarticulatory resistance of the consonant articulation).

EMA is a current technique to record articulatory data. A sampling rate of 200Hz gives very good temporal resolution. This allows the capturing of very fine speech movements. The 3D articulograph (AG500, Carstens Medizinelektronik) was used to track the movement of tongue, jaw, lips and head. For this study, we used four sensors to cover the tongue deformation glued on the midsagittal plane.

As our main goal is to study *Pha*, covering the back of the tongue is important. For this reason, we placed one sensor as far as possible to the back of the tongue.

4. EXPERIMENT 1

Two word lists of 24 words per speech style (MSA and DA) were produced by 16 male speakers from Jordan, Kuwait, Morocco and Yemen (4 speakers per country). The words were $C_1V_1C_2V_2C_3V_3$ syllable-typed where C_2 was either /tˤ dˤ sˤ ðˤ/ or /t d s ð/ and V_2 was either /i a u/. A total of 1152 tokens per speech style were labelled with PRAAT. Manual measurement of the frequency of *F2* was taken at the vowel onset (*onset*) and midpoint (*mid*). A total of 4608 formant measurements were carried out.

Table 1: mean values of y intercepts (int-y), slopes and regression coefficients (R^2) in MSA for 16 speakers.

MSA	**NON-PHARYNGEALISED**				**PHARYNGEALISED**			
	t	d	s	ð	tˤ	dˤ	sˤ	ðˤ
Int-y	423	515	335	385	473	434	262	420
slope	**0.77**	**0.71**	**0.81**	**0.77**	**0.54**	**0.57**	**0.77**	**0.56**
R^2	0.91	0.82	0.90	0.92	0.76	0.77	0.85	0.79

LE were accurate to indicate CV *Coa* [12, 17, 26, 31]. Stylistic variation (MSA *vs.* DA) was reflected by specific *LE* parameters (Table n ° 1 & 2). Speech style was accurate in producing different CV *Coa*. Slope values of the same consonant are different in MSA and DA. A two-way ANOVA (*Pha* x speech style) showed significant effects [*F*(3, 63) = 2.17, p < .001]).

In MSA, *Ph* consonants showed flatter slopes compared to their *n.Ph* cognates. This result is in accord with other studies [8, 32, 33]. A one-way ANOVA showed significant effects of the *Ph* consonant nature on slope values [*F*(3, 63)=4.86, p < .01]; whereas the effects of the *n.Ph* consonant were not significant [*F*(3, 63)=1.23, p=304]. A one-way ANOVA indicated significant effects across *Pha* for 3 pairs of consonants (/t-tˤ/ [*F*(1, 15)=0.25, p=.006], /s-sˤ/ [*F*(1, 15)=0.27, p=.008], /ð-ðˤ/ [*F*(1, 15) = 0.27, p= .008], /d-dˤ/ [*F*(1, 15)=0.86, p=.392]).

Table 2: mean values of y intercepts (int-y), slopes and regression coefficients (R^2) in DA for 16 speakers.

DA	**NON-PHARYNGEALISED**				**PHARYNGEALISED**			
	t	d	s	ð	tˤ	dˤ	sˤ	ðˤ
Int-y	598	636	385	436	350	437	518	510
slope	**0.67**	**0.65**	**0.79**	**0.66**	**0.67**	**0.60**	**0.69**	**0.537**
R^2	0.83	0.85	0.70	0.74	0.80	0.86	0.72	0.70

Consonant alternation, plain *vs. Ph*, showed regular influence on *LE* patterns in DA. The consonant nature did not show any significant effect on slope values, neither for the *Ph* [*F*(3, 63)=1.45, p=.237], nor for the plain cognates [*F*(3, 63)=2.73, p=.051]. Except /dˤ/ and /tˤ/ which have steeper values, *Ph* consonants show flatter slopes compared to those of their plain cognates. A one-way ANOVA did not indicate significant effects for *Pha* across 3 pairs of consonants (/t-tˤ/ [*F*(1, 15)=0.53, p=.120], /s-sˤ/ [*F*(1, 15)=0.42, p=.055], /ð-ðˤ/ [*F*(1, 15)=1.49, p=.223], /d-dˤ/ [*F*(1, 15)=0.36, p=.029]. Dialectal origin showed some differences on the extent of *CoaE*. A Wilks test three-way ANOVA (*Pha* x speech style x geographical origin) showed higher significant effects [*F*(5, 15)=2.63, p=.003].

5. EXPERIMENT 2

5.1. Libyan Arabic

A word list of 18 words in Libyan Arabic (LA) was produced by 10 speakers, 5 males and 5 females. The words were $C_1V_1C_2V_2C_3V_3$ syllable-typed where C was either *Ph* /tˤ dˤ sˤ/ or *n.Ph* /t d s/ and V was either /i a u/. For the stressed syllable (C_1V_1), 1080 tokens, and for the unstressed syllables (C_2V_2-C_3V_3), 2160 tokens were collected. All the data was measured in the same conditions than the experiment 1 (*cf.* supra), but no statistical design was applied.

LE parameters were flatter for *Ph* consonants, except for /dˤ/, and steeper for *n.Ph* cognates under stress. The model shaped by the focal position (S1) shifts slightly in S2, and clearly in S3. So, the tongue rearward, necessary for the *Pha* articulation, and its expansion over the flanking vowel clearly weaken from S1 to S3 (table 3, 4, 5). The tongue's rearward in *Ph* articulation triggered weaker effects on *F2onset* when the *Ph* consonant is in S2 and S3.

Table 3: mean values of y intercepts (int-y), slopes and regression coefficients (R^2) in stressed syllable for 10 speakers in LA.

S1	NON-PHARYNGEALIZED			PHARYNGEALIZED		
	t	d	s	tˤ	dˤ	sˤ
Int-y	684	1089	965	436	762	731
slope	**0.64**	**0.46**	**0.50**	**0.61**	**0.43**	**0.50**
R^2	0.69	0.53	0.54	0.65	0.36	0.44

Table 4: mean values of y intercepts (int-y), slopes and regression coefficients (R^2) in unstressed syllable for 10 speakers in LA.

S2	NON-PHARYNGEALIZED			PHARYNGEALIZED		
	t	d	s	tˤ	dˤ	sˤ
Int-y	1142	1135	918	153	821	638
slope	**0.44**	**0.43**	**0.53**	**0.86**	**0.39**	**0.58**
R^2	0.44	0.49	0.59	0.63	0.44	0.58

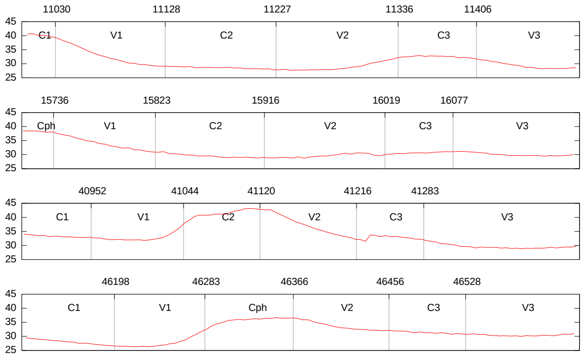
Table 5: mean values of y intercepts (int-y), slopes and regression coefficients (R^2) in unstressed syllable for 10 speakers in LA.

S3	NON-PHARYNGEALIZED			PHARYNGEALIZED		
	t	d	s	tˤ	dˤ	sˤ
Int-y	1204	1352	1207	111	798	322
slope	**0.43**	**0.32**	**0.40**	**0.90**	**0.35**	**0.79**
R^2	0.58	0.27	0.49	0.59	0.41	0.80

5.2. EMA data

We recorded 3 male speakers from Jordan, Sudan and Tunisia. The words in MSA were $C_1V_1C_2V_2C_3V_3$ syllable-typed where C was either *Ph* or *n.Ph* and V was either /i a u/. We examined the initial stressed syllable (C_1V_1) and the medial syllable (C_2V_2). The results presented in this section should be considered as observations, as currently we did not perform a quantitative study but rather an examination of several sequences. The purpose is to show the potential of using EMA data even in *Ph* context. A typical observation across the 3 speakers showed some consistency regarding the influence of the *Pha* on the neighbouring syllables, as can be illustrated in Figure 1. The two top figures present the case of *Ph* *vs.* *n.Ph* consonants under the stressed syllable (C_1V_1). The two bottom figures present the same consonants in the unstressed medial syllable (C_2V_2). The y-axis represents the horizontal displacement of the tongue from the back toward the lips. For sake of clarity, we show only the sensor farthest to back of the tongue. Thus, when the values of a given trajectory decrease, that means the tongue is moving to the back of the vocal tract. When *Ph vs. n.Ph* contrast is in the stressed syllable (S1), the rearward of the tongue is not marked, and the *Coa* does not seem to exceed the flanking vowel. When the contrast is in the medial syllable, the *Ph* consonant is retracted to the back, compared to *n.Ph* cognate. In addition, *Pha* affects the surrounding syllables (C_1V_1 and C_3V_3), with an important anticipatory rearward of the tongue on S1. The clarity effects imposed on S1 does not seem to resist the rearward requirement of the tongue imposed by the *Ph* consonant. This was observed for the three speakers.

Figure 1: Articulatory trajectories of one sensor glued on the back of the tongue of a Sudanese speaker. In the two top trajectories *Ph* is Cph and *n.Ph* is C1. The two bottom trajectories show the case where the Ph is in the middle syllable (*Ph* is Cph and *n.Ph* is C_2). The x-axis is time and the y-axis is the horizontal displacement of the tongue. In this example, the *Ph* is /tˤ/ and the *n.Ph* is /t/.



6. CONCLUSION

This paper showed the co-variation of speech clarity and the patterns of *Coa* between C & V. First, the variation in two speech styles, formal MSA *vs.* non formal DA, was clearly conveyed by different *LE* parameters between *Ph vs n.Ph*. The slope values were flatter for *Ph* consonants and steeper for *n.Ph* ones in MSA compared to their counterparts in DA. Second, the prosodic variation was reflected by different *CoaP*. On the one hand, *LE* parameters were found to be relevant for stress distinction in LA, with a shaped coarticulatory model of C & V under stress and progressively modified in unstressed syllables. On the second hand, articulatory observations showed relative backing of the tongue under stress, and clear backing outside stress. When C_2 was *Ph.*, the rearward of the tongue went on during V_2, which was responsible of *F2* lowering. However, the tongue's rearward anticipatory effects of the medial syllable are stronger than the clarity constraint imposed by S1.

The study showed the effects of a higher speech clarity (formal MSA and stressed syllable) on the articulation of *Ph* consonants, produced with a clear velo-pharyngeal constriction expanding over the flanking vowel, which is in accord with the literature [21], whereas in weaker speech clarity (non formal DA and unstressed syllable) the constriction was lighter. Thus, variation in speech style and prosodic position was accurate in producing different CV *CoaP*. Speakers seem to coarticulate more in higher speech clarity than in weaker speech clarity. These findings are not in accord neither with [25] and [5] who did not find any significant effect according to speech clarity, nor with those of [17] and [30] who found higher *LE* parameters in spontaneous speech, compared to a more formal one.

7. REFERENCES

[1] Abramson, A., Ferguson, C. 1962. *X-ray sound motion picture film in Arabic*. Columbia University: Columbia-Presbyterian Medical Center & Haskins Laboratories.

[2] Al-Ani, S.H. 1970. *Arabic Phonology*. The Hague: Mouton.

[3] Ali, L.H., Daniloff R.G. 1972. A contrastive cinefluographic inverstigation of the articulation of emphatic *vs.* non-emphatic cognate consonants. *Studia Linguistica* 26(2), 81-105.

[4] Beddor, P.S., Harnsberger, J.D., Lindemann, S. 2002. Language-specific patterns of vowel-to-vowel coarticulation: acoustic structures and their perceptual correlates. *J. of Phonet.* 30, 591-627.

[5] Bradlow, A.R. 1995. A comparative acoustic study of English and Spanish vowels, *JASA* 97, 1916-1924.

[6] Edwards, J., Beckman, M.E. & Fletcher, J. 1991. The articulatory kinematics of final lengthening. *JASA* 89, 369-382.

[7] Embarki, M. 2008. Les dialectes arabes modernes : état et nouvelles perspectives pour la classification géo-sociologique. *Arabica* 55, 5(6), 583-604.

[8] Embarki, M., Yeou, M., Guilleminot, C., Al Maqtari, S. 2007. An acoustic study of coarticulation in Modern Standard Arabic and Dialectal Arabic: Pharyngealized *vs.* non-pharyngealized articulation. *Proc. 16th ICPhS*, Saarbrücken, 141-146.

[9] Embarki, M., Ouni, S., Yeou, M., Guilleminot, C., Al Maqtari, S. 2011. Acoustic and EMA study of pharyngealization: Coarticulatory effects as index of stylistic and regional distinction, In Hassan, Z.M., Heselwood, B. (eds.), *Instrumental Studies in Arabic Phonetics*. Amsterdam: J. Benjamins (in press).

[10] Farnetani, E. 1990. V-C-V lingual coarticulation and its spatiotemporal domain. In: Hardcastle, W.J., Marchal, A. (eds.), *Speech Production and Speech Modeling*. Kluwer Academic: Dordrecht, The Netherlands, 93-130.

[11] Fowler, C.A. 1981. A relationship between coarticulation and compensatory shortening. *Phonetica* 38, 35-50.

[12] Fowler, C.A. 1994. Invariants, specifiers, cues: An investigation of locus equations as information for place of articulation. *Perception and Psychophysics* 55, 597-610.

[13] Ghazali, S. 1977. *Back Consonants and Backing Coarticulation in Arabic*. Ph.D. Dissertation, University of Texas at Austin, Austin.

[14] Gick, B., Campbell, F., Oh, S., Tamburri-Watt, L. 2006. Toward universals in the gestural organization of syllables: A cross-linguistic study of liquids. *J. of Phonet.* 34, 49-72.

[15] Hassan, Z.M. 2005. Acoustic evidence of the prevalence of the emphatic feature over the word in Arabic. *Proceedings FONETIK*, 127-130.

[16] Krakow, R. 1993. Nonsegmental influences on velum movement patterns: syllables, sentences, stress, and speaking rate. In Huffman, M., Krakow, R. (eds.), *Phonetics and Phonology*: *Nasals*, *Nasalization*, *and the Velum*, Vol. 5. New York: Academic Press, 87-116.

[17] Krull, D. 1989. Second formant locus patterns and consonant-vowel coarticulation in spontaneous speech. *Perilus* 10, 87-108.

[18] Lindblom, B. 1963. On vowel reduction. Report 29. The Royal Institute of Technology, Speech Transmission Laboratory, Stockholm.

[19] Lindblom, B. 1990. Explaining phonetic variation: A sketch of the H&H theory. W. Hardcastle, & A. Marchal (eds.), *Speech Production and Speech Modelling*, Kluwer: The Netherlands, 403-439.

[20] Lindblom, B., Guion, S., Hura, S., Moon, S.-J.; Willerman, R. 1995. Is sound change adaptive? *Rivista di Linguistica* 7, 5-36.

[21] Magen, H.S. 1997. The extent of vowel-to-vowel coarticulation in English. *J. of Phonet.* 25, 187- 205.

[22] Manuel, S.Y. 1990. The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *JASA* 88, 1286-1298.

[23] Manuel, S.Y. 1999. Cross-language studies: relating language-particular coarticulation patterns to other language-particular facts. In Hardcastle, W., Hewlett, N. (eds.), *Coarticulation: Theory, Data and Techniques*, Cambridge: CUP, 179-198.

[24] Manuel, S.Y., Krakow, R.A. 1984. Universal and language particular aspects of vowel-to-vowel coarticulation. *Haskins Laboratories Status Reports on Speech Research*, SR-77/78, 69-78.

[25] Matthies, M., Perrier, P., Perkell, J.S. Zandipour, M. 2001. Variation in anticipatory coarticulation with changes in clarity and rate. *J. Sp. Lang. Hear. Res*. 44, 340-353.

[26] Modarresi, G., Sussman, H.M., Lindblom, B., Burlingame, E. 2005. Locus equation encoding of stop place: revisiting the voicing/VOT issue. *J. of Phonet.* 33, 101-113.

[27] Öhman, S.E.G. 1966. Coarticulation in VCV utterances: spectrographic measurements. *JASA* 39, 151-168.

[28] Recasens, D. 1987. An acoustic analysis of V-to-C and V-to-V coarticulatory effects in Catalan and Spanish VCV sequences. *J. of Phonet.* 15, 299-312.

[29] Recasens, D., Pallarès, M.D. & Fontdevila, J. 1998. An electropalatographic and acoustic study of temporal coarticulation for Catalan dark /l/ and German clear /l/. *Phonetica* 55, 53-79.

[30] Sussman, H.M., Dalston, E., Gumbert, S. 1998. The effect of speaking style on a locus equation characterization of stop place of articulation. *Phonetica 55*, 204-225.

[31] Sussman, H.M., Fruchter, D., Hilbert, J., Sirosh, J. 1998. Linear correlates in the speech signal: The orderly output constraint. *Beh. & Brain Sci.* 21, 241-299.

[32] Sussman, H.M., Hoemeke, K., Ahmed, F. 1993. A cross-linguistic investigation of locus equations as a relationally invariant descriptor of place of articulation. *JASA* 94, 1256-1268.

[33] Yeou, M. 1996. Locus equations and the degree of coarticulation of Arabic consonants. *Phonetica* 54, 187-202.