ISCA Archive
http://www.isca-speech.org/archive

INTERSPEECH 2012
ISCA's 13th Annual Conference
Portland, OR, USA
September 9-13, 2012

Physiological and acoustic study of word initial post-lexical gemination in Moroccan Arabic

[1,2]Chakir Zeroual, [3]Diamantis Gafos, [4]Phil Hoole and [5]John H. Esling

[1]Faculté Polydisciplinaire de Taza, BP. 1223, Taza, Morocco [2]Laboratoire de Phonétique et Phonologie, CNRS-UMR7018 & Sorbonne-Nouvelle, Paris. [3]University of Postdam. [4]Institut fuer Phonetik und Sprachverarbeitung, University of Munchen. [5]University of Victoria, Canada.

chakirzeroual@yahoo.fr aigafos@gmail hoole@phonetik.uni-muenchen.de.com esling@uvic.ca

Abstract

With the EMA technique we showed that derived word initial [dd] of Moroccan Arabic is produced by our three speakers with a longer acoustic duration and articulatory gesture compared to its initial [d] cognate. The tongue tip gesture during initial [dd] and [d] generally have statistically similar height and amplitude. For two speakers, no significant velocity differences were recorded between [dd] and [d]. Our articulatory observations also show no C-center effect between [d(d)] and [u] in the [d(d)uda] items, which is in accord with the study of Shaw et al. (2009) on word initial [Ci], [CjCi] and [CkCjCi] clusters in MA which they use as an argument against complex onset in this language. Our results seem to show that initial hetero-morphemic geminate [dd] is a cluster of two hetero-syllabic consonants.

Index Terms: EMA, Moroccan Arabic, germination, kinematic, temporal coordination, C-center effect.

10.21437/Interspeech.2012-273

1. Introduction

Compared with their singleton cognates, non-continuant (homo or hetero-morphemic) geminate consonants are mainly characterized by longer acoustic closure phase (Lahiri, et al., 1988; Ladefoged and Maddieson, 1996; Zeroual et al. 2008) and articulatory closure contact (Kraehenmann and Lahiri, 2008; Ridouane, 2007). The articulatory adjustments responsible for this quantity contrast are complex and not well understood (Löfqvist, 2005). Articulatory investigations of geminate will help us not only to characterize the articulatory mechanisms involved in this quantity contrast, but will also improve our understanding of other types of time variations.

For the "mass spring model" (Kelso et al. 1986; Turvey et al., 1986; Saltzman and Kelso, 1987), duration variations are a passive consequence of stiffness adjustment: i.e, geminate involves movements with a lower degree of stiffness than its singleton correspondent. This prediction has been validated for geminate labials in Italian (Gili Fivela et al. 2005, 2 speakers) and Japanese (Löfqvist 2005, 2 speakers), but not for coronal /tt, dd/ in Moroccan Arabic (MA) (Zeroual et al., 2008, 2 speakers) and labials produced by one Japanese speaker (Löfqvist 2005). These latter results show that other articulatory adjustments may distinguish between singleton/geminate consonants.

Löfqvist (2005) suggested that the longer constriction during a geminate is due to its higher virtual target compared with its singleton cognate. According to this hypothesis, the geminate is associated with a higher vertical position to maintain a longer contact with its active articulator. This expectation was confirmed for Japanese /mm/ (3 speakers), but not for MA /tt dd/ (2 speakers). Furthermore, a more recent MRI study demonstrated that geminate labial and dorsal consonants, but not coronals, have a more distributed and compressed contact (Hagedorn et al. 2011). The virtual target hypothesis seems to work well with constriction involving passive and active articulators with soft tissues. It also predicts that geminates would have not only higher amplitude, but also higher peak velocity due to the strong positive correlation generally observed between these parameters (Parush et al., 1983).

Segment duration variations can also be due to a difference in degree of overlap between gestures. Notice that almost all the available data on singleton/geminate contrast come from intervocalic position where, in many languages, geminates induce a preceding vowel shortening. This latter is a regular and more important pattern in Italian (Zmarich et al. 2011), irregular and very slight in MA (Zeroual et al., 2008, 2012) and absent in Japanese (Smith, 1995). This shortening is generally due to the anticipation of the geminate closing movement during the preceding vowel (Zmarich et al. 2011; Zeroual, 2012). Articulatory studies show no difference between the degree of overlap of intervocalic singleton/geminate consonant and the following vowel (Zmarich et al. 2011, Zeroual, 2012).

As mentioned before, almost all the existing data on geminates are from intervocalic position, where the lengthening of their closure duration is their major acoustic correlate (Lahiri, et al., 1988). Few studies have been devoted to word (or phrase) initial position where the consonant quantity contrast is cross-linguistically rare (Ladefoged and Maddieson, 1996), and closure duration differences neutralized between oral voiceless plosives particularly in the absolute initial position. According to Abramson (1987, 1991, 1999), in this position, the RMS amplitude of the first syllable as well as the VOT and F0 are more important after the release of a geminate. However, these results have not been obtained by other studies: non-significant differences between the VOT duration (Kraehenmann and Lahiri, 2008; Ridouane, 2007) and RMS amplitude (Ridouane, 2007) after the release of the singleton and geminate consonants.

For theoretical view of the syllabification, a geminate is a sequence of two identical consonants and only its second half is parsed with the post-geminate vowel into one syllable. To test this hypothesis, we investigated the potential absence of C-center effect between word initial geminate consonant and the following vowel (see section 2.4).

In MA, all the consonants contrast with their geminate cognates, but only in the intervocalic position. In word initial position, several types of hetero-morphemic gemination are attested. We propose an exhaustive physiological description of the spatio-temporal properties of initial [d] vs. [dd], where [dd] is a very productive derived geminate. It results from a total assimilation between the definitive article, whose underlying form is /l/, and the first consonant [d] of a noun. Our present physiological data will be discussed while taking into account several articulatory measures and hypotheses outlined above.

2. EMA experiment

2.1. Speakers and materials

Four MA native speakers participated in an experiment by EMA (AG500 Carstens Medizinelektronik, Hoole et al. 2011). Data from three subjects (S1 S2 S3) will be presented here.

Several items of the form [#C_iV-], [#$C_j$$C_i$V-], [#$C_k$$C_j$$C_i$V-] and [#$C_i$$C_i$V-] were pronounced at least 14 times in the carrier phrase [ʒibi____hnaja] ('bring here'). This corpus was recorded to study a potential C-center effect in MA. Only [duda] ('an earthworm') vs. [dduda] ('the earthworm') items will be discussed here. [dd] is a derived geminate consequence of a total assimilation in the underlying form /l+duda/ (/l/: definite article).

2.2. Method

We recorded (200Hz) horizontal and vertical movements of the tongue, lips and jaw with sensors placed near the tip (TTIP), the mid part (TMID) the back of the tongue (TDOR), on the external extremity of the lower lip (LLIP) and behind the lower incisors (JAW). Here, only TTIPy vertical movements were analyzed.

With a Matlab Mview program developed by M. Tiede (Haskins Laboratories), we identified automatically gestural onset (Gons), Target or achievement (T), maximal (Max), plateau release (R) and gestural offset (Off) positions for TTIPy. The plateau midpoint (Mid) has been identified manually. Gons and target positions correspond to the temporal positions where the instantaneous velocity reaches 20% of its maximum value during the closing movement. The same threshold was used for Release and Goff during the opening movement (Fig. 1). Max represents the point of minimum velocity during closure.

2.3. Acoustic measurements

We measured the duration of the closure, burst and the total duration of [#d] and [#dd] and [u] in [#duda] vs. [#dduda] (Table 1). The closure onset was easily segmented, since almost all of our items were pronounced without a pause with the preceding word [ʒibi] of the carrier phrase. When a pause was present, the onset of voicing was selected as the onset of [d(d)] closure.

2.4. Articulatory measures

During initial [#d(d)] in [d(d)uda], several duration, spatio-temporal and kinematic measurements have been done on TTIPy gesture traces (Fig. 1). From temporal positions extracted automatically, we calculate the duration of TTIPy closing, plateau and opening phases. Amplitude and peak velocity, as well as the y-values at Gons, Target, Midpoint, Release and Goff positions were extracted automatically.

In order to quantify the degree of overlap and a potential C-center effect between initial [#d(d)] and the following [u] in [d(d)uda] items, and partially based on Shaw et al (2009) Browman and Goldstein (1988) and Honorof and Browman (1995) studies, we measured three temporal intervals: TTIPy target during medial [d] to its Target, Midpoint and Release during initial [d(d)] (Fig. 1v). These C to C temporal intervals, with TTIPy Target during median [d] as the common anchor point, are used to provide information about the temporal coordination between [u] and the preceding consonants.

Notice that C-center hypothesis is generally based on relative variability of intervals, rather than invariance of any specific interval. However, in this study we will compare between absolute duration measures since we have a minimal pair where all the segments (i.e. [-uda]) after [#d] and [#dd] are similar.

3. Results and discussion

To limit the variability generally observed with the articulatory measures, our statistical analyses (t-tests) were done separately for S1, S2 and S3 (13, 11 and 14 tokens, respectively).

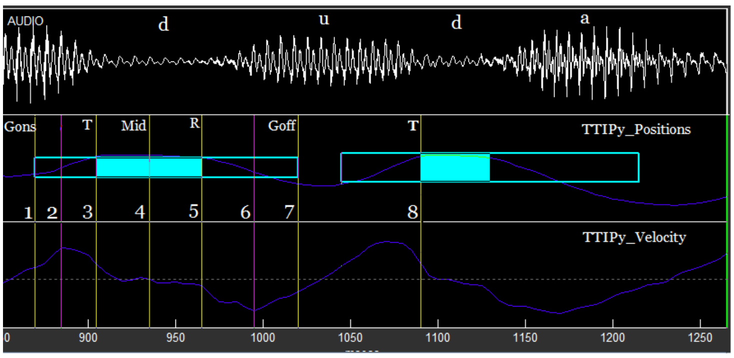

Figure 1: Audio, TTIPy vertical position (TTIPy_Positions) and velocity (TTIYPy_Velocities) traces, with articulatory landmarks used to extract the following duration, spatio-temporal and kinematic measurements:
(i)-Duration of the closing (3-1), plateau (5-3) and opening movement (7-5) of TTIPy during initial [d].
(ii) Amplitude of its closing and opening movements (euclian distance from 1 to 3 and from 5 to 7, respectively).
(iii) Peak velocity of its opening (2) and closing (6) movements.
(iv)-Spatial: height values of TTIPy at 1, 3, 4, 5, and 7 positions.
(v)- C to C temporal intervals: temporal interval between TTIPy target of medial [d] (position 8) and its Target (8-3) Midpoint (8-4) and Release (8-5) positions during initial [d].

3.1. Acoustic durations

As expected, our three speakers produced [dd] with substantial longer total duration than [d] (Table 1), mainly due to the lengthening of its closure phase (% of difference $\geq$ 100%). [u] is slightly shorter in [dduda] than [duda], this difference is significant only for S3, but the highest difference (-15%) is smaller than the JND of duration (20%: Rossi, 1972), and can explain why the shortening of post-geminate vowel was previously not proposed as a secondary correlate for the geminate perception. This result also suggests a similar degree of overlap between the initial [d] and [dd] with [u] in our [duda] and [dduda] items.

		Burst	CL	TL	[u]
S1	[d]	25	76	101	125
	[dd]	24	164	188	121
	% Δ	-4^{ns}	116***	87***	-4^{ns}
S2	[d]	18	70	89	102
	[dd]	24	140	164	98
	% Δ	30**	99***	85***	-4^{ns}
S3	[d]	17	59	75	74
	[dd]	22	119	141	63
	% Δ	31***	103***	87***	-15***

Table 1. Mean acoustic burst, closure and total duration (ms) of initial [d] vs. [dd] and [u] in [duda] vs. [dduda], % Δ: percentage differences. Paired significances: *** = p<0.001, ** = p<0.01, * = p<0.05, ns = not significant.

3.2. Articulatory durations

For our three speakers, TTIPy total duration is significantly longer during initial [dd] than [d] (Table 2). This difference is also largely attributed to the lengthening of plateau phase during [dd]. The closing and opening phases stay generally unaffected (Table 2), except for the TTIPy opening movement produced by S1. Substantially longer constriction duration during geminate, compared to the closing and/or the opening movement, was also reported by published physiological data generally in intervocalic position (Zeroual et al. 2008; Zmarich et al. 2011; Hagedorn, 2011; Zeroual, 2012) and only by few studies for word initial geminate (Kraehenmann and Lahiri, 2008).

		Closing	Plateau	Opening	Total
S1	[d]	60	45	84	188
	[dd]	52	138	74	264
	% Δ	-13^{ns}	209***	-12***	40***
S2	[d]	35	55	66	157
	[dd]	43	120	72	235
	% Δ	22^{ns}	116***	9^{ns}	50***
S3	[d]	38	54	35	128
	[dd]	35	110	38	183
	% Δ	-8^{ns}	103***	8^{ns}	44***

Table 2. Mean aticulatory duration (ms) of the closing, plateau and opening phases of TTIPy gesture and its total duration during initial [d] vs. [dd] in [duda] vs. [dduda], % Δ: percentage differences. Paired significance: *** = p<0.001, ** = p<0.01, * = p<0.05, ns = not significant.

3.3. Velocity, amplitude and height

For our three speakers, the difference between the amplitude of closing and opening movements of TTIPy during initial [d] vs. [dd] is not significant or slightly significant (Table 3). In fact, in only two cases was this difference slightly significant (opening movement for S1 and closing movement for S3, Table, 3), and showed a lower amplitude for the geminate instead of an expected higher one. We can therefore conclude that [dd] is produced by our three speakers without significantly higher amplitude than [d].

		Velocity		Amplitude	
		Closing	Opening	Closing	Opening
S1	[d]	12,9	15,6	5,3	9,7
	[dd]	14,6	18,2	5,2	9,0
	p	**	***	Ns	*
S2	[d]	8,0	10,4	2,2	4,9
	[dd]	7,8	10,8	2,3	5,4
	p	Ns	Ns	Ns	Ns
S3	[d]	19,2	20,2	4,7	4,7
	[dd]	18,7	19,3	4,2	4,8
	p	Ns	Ns	*	Ns

Table 3. Mean values of the velocity (cm/s) and amplitude (mm) of the TTIPy closing and opening movements during initial [d] vs. [dd] in [duda] vs. [dduda] with their paired significant differences: *** = p<0.001, ** = p<0.01, * = p<0.05, ns = not significant.

For S2 and S3, the peak velocity of TTIPy opening and closing movements during [dd] is not significantly higher than during [d] (Table 3). This pattern is in accord with their non-significant or slightly significant amplitude differences. However, for S1, peak velocities of TTIPy closing and opening movements are significantly higher during initial [dd] than its [d] cognate. This result is not expected since, during [dd], the amplitude of the closing movement is not significantly different and the amplitude of its opening movements slightly lower than during [d].

		Target	Mid	Release
S1	[d]	7,40	7,64	7,48
	[dd]	7,04	7,55	7,13
	p	ns	ns	ns
S2	[d]	4,37	4,40	4,03
	[dd]	4,62	4,73	4,47
	P	ns	ns	ns
S3	[d]	3,04	3,46	3,41
	[dd]	3,30	3,73	3,46
	p	ns	Ns	ns

Table 4. Mean height (mm) of TTIPy at its target, midpoint and release positions (see also Fig. 1) during initial [d] vs. [dd] in [duda] vs. [dduda], with their paired significant differences: *** = p<0.001, ** = p<0.01, * = p<0.05, ns = not significant.

Notice that, for our three speakers, the TTIP height at target, midpoint and the release of TTIPy gesture during [dd] is not significantly different compared with [d] (Table, 4 and Fig. 1). This agrees with Zeroual et al. (2008) observations on MA intervocalic [t d tt dd], and probably not with Löfqvist's hypothesis (2005) which predicts a higher virtual target during a geminate. It seems that the more rigid surfaces of the alveolar and teeth surfaces prevent a substantial raising and fronting of the tongue during the geminate coronal plosives.

3.4. [d(d)] and [u] degree of overlap & C-center effect

		To Target	To Mid	To Release
S1	[d]	225	203	180
	[dd]	312	241	174
	p	***	***	ns
S2	[d]	193	165	138
	[dd]	262	201	142
	P	***	***	ns
S3	[d]	149	122	95
	[dd]	202	146	92
	p	***	***	ns

Table 5. Mean durations (ms) of temporal intervals between target of medial [d] (the common anchor point) to Target, Midpoint and Release of initial [d(d)] in [d(d)uda] items (Fig. 1v), with their paired significant differences: *** = p<0.001, ** = p<0.01, * = p<0.05, ns = not significant.

The measures presented in Table 5 show that, for our three speakers, the time interval between TTIPy target during median [d] (common anchor point) and its midpoint position during initial [dd] is statistically longer than that recorded between median and initial [d]. A similar pattern is obtained between [dduda] and [duda] for the temporal interval between TTIPy target during [d(d)] and the common anchor point. These results, and mainly the first one, clearly show that initial [d(d)] induces no C-center effect in [d(d)uda] items. By contrast, in [dduda] and [duda] the duration of the interval between TTIPy release during [d(d)] and its target during median [d] remains constant. This latter result suggests that the temporal coordination between initial [dd] and [u] and their degree of closure overlap seem similar to the ones recorded between initial [d] and [u].

4. Conclusions

With the EMA technique, we tried to determine the articulatory characteristics of word initial derived geminates in Moroccan Arabic.

Indeed, MA [dduda] vs. [duda] items, pronounced by 3 speakers without a pause after the preceding word ([ʒibi]), show that the derived geminate [dd] has a substantially longer acoustic closure than initial [d].

Our articulatory measures show that at the word initial position; [dd] has a longer constriction than [d], and similar durations for their closing and opening movements.

Initial [dd] is produced by our three speakers without significantly higher amplitude than [d]. The height at the midpoint of tongue tip gesture during [dd] is not significantly different compared with [d]. For two speakers, the peak velocity of the closing and the opening phases of tongue tip gesture during initial [dd] and [d] are also not significantly different.

Our articulatory observations seem to show no C-center effect between [d(d)] and [u] in the [d(d)uda] item. This result is in accord with the study of Shaw et al. (2009) which has also reported no C-center effect in [#C_iV-], [#C_jC_iV-] and [#$C_kC_jC_i$V] sequences in Moroccan Arabic. They used this observation as an argument against complex onset in MA.

Since C-center effect seems also absent in [#dduda] and that several kinematic properties of the closing and opening movements of TTIPy during [#d] and [#dd] are generally similar, we can deduce that our results suggest that [#dd] is a sequence of two overlapped hetero-syllabic consonants.

5. References

[1] Abramson, A. S. Word-initial consonant length in Pattani Malay. *Proc. 11th ICPHS*, Tallinn, 68–70, 1987.

[2] Abramson, A. S. Amplitude as cue to word-initial consonant length: Pattani Malay. *Proc. 12th ICPHS*, Aix-en-Provence, 98–101, 1991.

[3] Abramson, A. S. Fundamental frequency as cue to word-initial consonant length: Pattani Malay. *Proc. 14th ICPHS*, San Francisco, 591–594, 1999.

[4] Browman, C. P. and Goldstein, L. Some notes on syllable structure in articulatory phonology. *Phonetica 45*: 140–155, 1988.

[5] Gili Fivela, B. and Zmarich, C. Italian Geminates under Speech Rate and Focalization Changes: Kinematics, Acoustic, and Perception Data. *Proc. InterSpeech*, Lisbon, 2897-2900, 2005.

[6] Hagedorn, C., Proctor, M. and Goldstein, L. Automatic analysis of geminate consonant articulation using real-time MRI. In *Proc. InterSpeech*, Florence, 409-412, 2011.

[7] Honorof, D. and Browman, C.P. The centre or edge: how are consonant clusters organised with respect to the vowel? *Proc. 13th ICPHS*, Stockholm, 552–555, 1995.

[8] Hoole, P. and Zierdt, A. Five-dimensional articulography. In: *Speech Motor Control: New developments in basic and applied research*, Editors: Ben Maassen, Pascal H.H.M. van Lieshout. OUP, 331-349, 2010.

[9] Kraehenmann, A. and Lahiri, A. Duration differences in the articulation and acoustics of Swiss German word-initial geminate and singleton stops. *J. Acoust. Soc. Am*. 123: 4446-4454, 2008.

[10] Ladefoged, P. and Maddieson, I. The Sounds of the World's Languages. Blackwell: Cambridge USA & Oxford UK, 1996.

[11] Lahiri, A., and Hankamer, J. The timing of geminate consonants. *J. Phonetics* 16: 327–338, 1988.

[12] Löfqvist, A. Lip kinematics in long and short stop and fricative consonants. *J. Acoust. Soc. Am*. 117, pages 858-878, 2005.

[13] Parush, A., Ostry, D.J., and Munhall, K.G. A kinematic study of lingual coarticulation in VCV sequences. *J. Acoust. Soc. Am*. 74 (4): 1115-1125, 1983.

[14] Ridouane, R. Gemination in Tashlhyt Berber: An acoustic and articulatory study. *J. Int. Phonetic Assoc*. 37, 119–142, 2007.

[15] Rossi, M. Le seuil différentiel de durée. In (A. Valdman, 1972): 435-450, 1972.

[16] Shaw, J., Gafos, A.., Hoole, P., and Zeroual, C. Temporal evidence for syllabic structure in Moroccan Arabic: data and model. *Phonology* 26(1), 187-215, 2009.

[17] Smith, C.L. Prosodic patterns in the coordination of vowel and consonant gestures. In: B. Connell & A. Arvaniti (eds) *Papers in Laboratory Phonology IV, Phonology and phonetic evidence*. CUP, 205-222, 1995.

[18] Zeroual, C., Hoole, P., and Gafos, A. Spatio-temporal and kinematic study of Moroccan Arabic coronal geminate plosives. *Proc. 8th ISSP*, Strasbourg, 135-138, 2008.

[19] Zeroual, C. Hoole, P., and Gafos, A. and Esling, J.H. Coordination spatio-temporelle dans les suites ab(b)i en arabe marocain. *Proc. JEP-TALN-RECITAL 2012*, Grenoble, 225-232, 2012.

[20] Zmarich, C., Gili Fivela, B., Perrier, P., Savariaux C., and Tisato, G. Speech Timing Organization for the Phonological Length Contrast in Italian Consonants. *Proc. Interspeech*, Florence, 401-404, 2011.