

Emphatic segments and emphasis spread in Lebanese Arabic: a Real-time Magnetic Resonance Imaging Study

Assaf Israel¹, Michael Proctor^{1,2}, Louis Goldstein¹
Khalil Iskarous¹, Shrikanth Narayanan^{1,2}

¹Department of Linguistics, University of Southern California, USA

²Viterbi School of Engineering, University of Southern California, USA

assafisr@usc.edu

http://sail.usc.edu/span

Abstract

Production of emphatic consonants by a speaker of Lebanese Arabic was examined using real-time magnetic resonance imaging (rtMRI). Emphatic consonants were found to be articulated with a lowered, more retracted tongue body than their non-emphatic counterparts, with the narrowest emphatic constriction observed in the upper pharynx. Both progressive and regressive emphasis spread was observed; spreading was not blocked by an intervening palatal approximant [j]. Emphaticized segments exhibit similar retraction and depression, with magnitudes that vary depending on the direction of spreading. These data suggest that emphasis spread may operate in a phonetically-complex way, not currently accounted for by phonological theory, and in addition, illustrate the advantage of real-time MRI as a method for studying emphasis in Semitic phonology.

Index Terms: speech production, Arabic, emphatic, emphasis spread, pharyngealization, real-time MRI

1. Introduction

Lebanese Arabic, like all varieties of Levantine Arabic, makes use of a class of emphatic coronal consonants /T, D, S, Z/, which contrast with their non-emphatic or ‘plain’ counterparts /t, d, s, z/.¹ Although there is an extensive body of work examining this opposition in many different Arabic varieties, we still do not have a complete understanding of the phonetic basis of emphasis and the relationships between the phonetic and phonological properties of emphatic consonants.

It has long been established that Arabic emphatics are produced with a secondary, posterior constriction not present in their non-emphatic counterparts, but the exact characterization of the differences in articulation are unclear. Phonetic investigations have revealed that emphatic consonants produced by speakers of different varieties of Arabic are produced with a constriction in the upper pharynx, at or slightly below the uvula [3, 4, 5], a lower pharyngeal/tongue-root constriction and epiglottis retraction [6, 7, 8], global pharyngeal narrowing [9], and lowering of the tongue dorsum in the palatal region [4]. More data is needed to establish with greater clarity what the goals of production of emphatic segments are, and which aspects of articulation are dialect- and speaker-specific.

It is also unclear how best to describe emphatic consonants so as to be able to adequately account for certain aspects of

their phonological behavior. Based on the way that they interact with neighbouring segments in different varieties of Arabic, emphatic consonants have been variously characterized as pharyngealized [10, 11], velarized [12, 13], uvularized [1], or characterized by a dorsal constriction [14].

The range and influence of emphatic consonants on neighbouring segments is also an active area of research in phonological theory: it has been variously argued that the domain of emphasis is the syllable [15], the word [2, 9], the uninflected word [16], only the immediate vowel to the right [16], or unbounded leftward within word [17, 14]. Different classes of segment appear to block spreading of emphasis within these domains [17, 16] – specifically it has been observed that the palatal segments /i-j-f-ʒ/ block left-to-right emphasis spread in some dialects [16, 2] – but once again, it is unclear how best to phonologically characterize emphatic and emphaticized segments so as to account for the full range of behaviors attested in different languages.

1.1. Studying Emphatic Production

A major obstacle to studying consonantal emphasis in Semitic languages is the difficulty of safely obtaining information about the articulation of the pharynx. Electromagnetic articulometry [18] and x-ray microbeam are ill-suited to such studies because sensors cannot be placed at sufficiently posterior locations on the tongue, and these methods provide no information about the pharyngeal wall. Ultrasound [18] cannot consistently or reliably image the epiglottal region because of the obscuring effect of the hyoid bone. Endoscopy [7, 9, 8] is a highly invasive technique, and X-ray [19, 4, 6] and videofluoroscope [3, 5] expose subjects to unacceptable levels of radiation.

Real-time MRI (rtMRI) is a safe, non-invasive technique which provides views of the entire midsagittal plane of the vocal tract at sufficiently high frame rates to examine articulatory coordination of lingual, labial and velic articulators [20, 21]. Most importantly, rtMRI allows unparalleled views of the dynamic state of articulation of the whole of the pharynx, from the glottis into the nasopharynx, and in this respect, it is an ideal modality with which to examine pharyngeal articulation. Structural MRI has been used to examine the articulation of sustained Assiri Arabic gutturals [22], but to our knowledge, no previous study has examined emphasis using real-time MRI.

1.2. Goals

The goal of this study is to examine the dynamic articulation of emphatic and emphaticized segments produced by a speaker of Lebanese Arabic. Specifically, we aim to:

¹Emphatic consonants are represented throughout this paper in capital letters. Although not strict IPA, this system has the advantage that it makes no assumptions about the phonetic realization of emphatics, and is standard practice for romanization of Semitic languages, e.g. [1, 2].

- i. test the suitability of rtMRI for investigating emphatic production;
- ii. characterize differences in global tongue shaping between emphatic consonants and their non-emphatic counterparts;
- iii. examine the extent and articulatory characterization of emphasis spread in this variety of Lebanese Arabic

2. Method

Verbs contrasting plain and emphatic consonants were elicited from a native speaker of Lebanese Arabic, a 28 year old female from Tripoli. Target consonants were elicited in word-initial and word-final position, with a range of intervening consonants. In this study, we limited the prosodic structure of words to CVCVC and CVCCVC, where the emphatic and emphaticized consonants are at the edges (Table 1). Palatal and non-palatal consonants were contrasted word-medially, to examine whether the former block spreading. All stimuli were elicited in the carrier phrase اَزْمَانٌ هُوَ (*huwwa ezzamaan*; ‘he the time’), and repeated by the subject 5 times.

NE L→R	EMPH L→R	NE R→L	EMPH R→L
talaf	Talab	balad	ballaT
tamman	Tamman	taman	zamaT
tanaffos	Sannaf	nafas	naffaD
taraf	Darab	—	—
zayyaf	Sayyaf	bayyat	bayyaD
nayyam	Tayyab	fayyaʕ	fa:yeD
dayyan	Tayyan	—	—
zayyat	tSayyad	—	—

Table 1: *Stimuli used in rtMRI experiment. Non-emphatic (NE) and emphatic (EMPH) consonants juxtaposed in minimally contrastive word pairs in (a) initial position (left two columns) to examine left-to-right emphasis spread; and (b) word-final position (right two columns) to examine right-to-left emphasis spread, through intervening (word-medial) consonants: non-palatal (top four rows) and palatal (bottom four rows). Word-level emphasis spread will be examined on underlined consonants in emphatic contexts and compared to corresponding underlined consonants in non-emphatic contexts.*

2.1. Image Acquisition

Data were acquired using a rtMRI protocol developed specifically for the dynamic study of speech production [20]. The subject’s upper airway was imaged in the midsagittal plane with spatial resolution 68 x 68 pixels over a 200 x 200 mm field of view. New image data were acquired at a rate of 12.5 frames per second, and reconstructed as 23.18 frames/sec. video using a sliding window technique [21]. The subject’s head was fixed throughout the scan to allow for comparison of tongue position between tokens. Audio was recorded inside the scanner, at 20 kHz, simultaneously with the MRI acquisition, and subsequently noise-reduced [23]. The resulting companion video and audio recordings allow for dynamic visualization of the entire midsagittal plane of the subject’s vocal tract during speech, including the trachea, glottis, oropharynx and nasopharynx.

2.2. Articulatory Analysis

MRI data were loaded into a custom graphical user interface especially designed for the synchronization, inspection and anal-

ysis of companion audio and video recordings [24, 25]. The image frame corresponding to the articulatory center of each target consonant was identified by examining video sequences, time-aligned audio and spectra. Because the subject’s head remained stationary throughout the acquisition session, tongue position can be compared across tokens.

For each segment of interest, midsagittal tongue posture at the consonantal or vocalic target was captured by automatically identifying air-tissue boundaries [24], and manually correcting the tongue outline against the MR image where the algorithm failed to locate the edges of lingual tissue with sufficient accuracy. Tongue contours and passive vocal tract structures, including the rear pharyngeal wall, were defined with respect to a semi-polar analysis grid superimposed on the vocal tract [26]; 41 tract-normal gridlines were distributed at 4.8 mm intervals, extending from the glottis (gridline 0) to a point beyond the lips (gridline 41). This grid-based segmentation of the midsagittal lingual surface and pharyngeal structures allows for the parametric analysis of tongue displacement in specific regions of the vocal tract, and comparison of constriction degree at different points in time.

3. Results

3.1. Articulation of Emphatics

Articulation of the plain and emphatic coronal stops contrasted at the beginning of the words *tamman* ‘to set price’ and *Tamman* ‘reassure’ are compared in Fig. 1. Tongue edges extracted from the frames showing maximum coronal constriction during the production of the initial consonant in five utterances of each word have been superimposed.

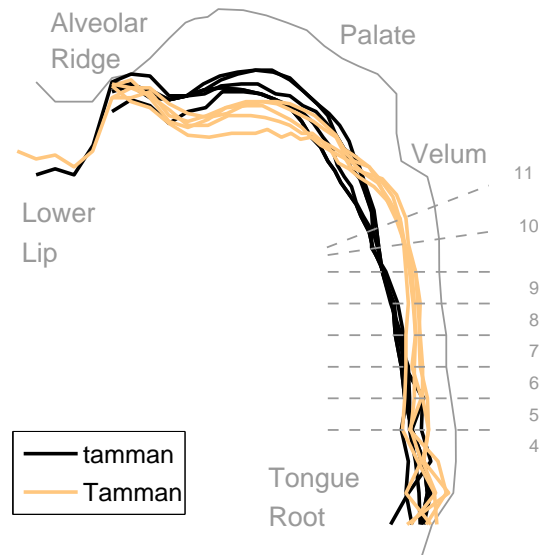


Figure 1: *Plain and emphatic word-initial coronal stops Tongue edges captured during mid-consonantal production of initial stops [t] and [T], in five repetitions of the words ‘tamman’ and ‘Tamman’. Pharyngeal region of semi-polar tract analysis grid (gridlines 4 to 11) superimposed on pharynx.*

The data in Fig. 1 reveal that, for this speaker, emphatic stops are characterized by a lowered tongue dorsum and a

more constricted pharynx, consistent with previous descriptions of Arabic emphatics in Iraqi, Tunisian and Jordanian dialects [3, 4, 5]. The narrowest part of the pharyngeal constriction formed between the back of the tongue dorsum and the upper rear pharyngeal wall in the word-initial emphatic stop /T/ is located approximately 40 mm above the glottis (gridline 8: mean pharyngeal aperture = 3.84 mm). Mean pharyngeal aperture over the entire pharyngeal region (gridlines 4 to 10)², calculated over five repetitions of the word-initial emphatic stop in *Tamman*, is 5.69 mm, compared with a mean pharyngeal aperture of 7.64 mm calculated over five repetitions of the word-initial non-emphatic stop in *tamman*. Similar patterns of articulation were observed in comparisons of other minimally-contrastive emphatic/plain consonant pairs in the experimental dataset.

3.2. Phonetic Characterization of Emphasis Spread

Our data indicate that emphasis spreads in both directions and affects a non-emphatic segment at the other end of the word. Lingual postures compared in Fig. 2 reveal that articulatory characteristics of the word-initial emphatic alveolar /((t)S/ spread rightwards through the word, and affect the articulatory posture of the word-final (non-emphatic) stop in *tSayyad*. The mean lingual posture at the mid-point of word-final consonantal production in '*tSayyad*' is compared to mean postures at the same point in time during the word-final consonantal production in '*bayyat*' and '*bayyaD*'. Each tongue edge was calculated from the mean of five utterances of each of the words. As illustrated in this example, emphaticized segments typically exhibit the same characteristics observed in intrinsically emphatic segments, but these characteristics appear to be less pronounced. For example, in Fig. 2, the word-final emphaticized [d] was produced with retracted and depressed tongue body, although both of these articulatory characteristics are less pronounced than observed in the intrinsically-emphatic final stop of *bayyaD*.

Disagreement among previous studies regarding the spread of emphasis across palatal intervening segments (see [1, 2, 11]) motivated us to investigate this issue further. We first compared word-final emphaticized and plain segments separated from the emphatic by *y* or *yy* (Table 2). *Global* mean pharyngeal apertures calculated for each type of segment (rightmost column) initially suggested that emphaticized consonants do not differ significantly from plain consonants word-finally. However, closer comparison of the two types of segments at different regions in the pharynx revealed significant differences in degree of constriction between word-final plain and emphaticized segments. Specifically, emphaticized segments were produced with a significantly narrower constriction in the upper pharynx, just below the uvula (gridline 10: pharyngeal aperture of 6.1 mm vs. 7.6 mm in plain consonants). Degree of constriction in the lower portion of the pharynx was not found to be significantly different for emphaticized and plain word-final consonants.

To assess the effect of the palatal segment, the articulation of word-final emphaticized segments following palatals and non-palatals was compared. A significant difference in constriction degree in the upper half of the pharynx (gridlines 7-10) was found in this context: emphaticized segments following palatals exhibited greater average pharyngeal aperture in this region (5.39 mm) than segments following palatals (4.45 mm). Finally, we tested whether the spread of emphasis across

²Pharyngeal aperture could not be robustly measured at gridline 11 across all frames of interest because the velum lowers into this region during nasal and nasalized segments.

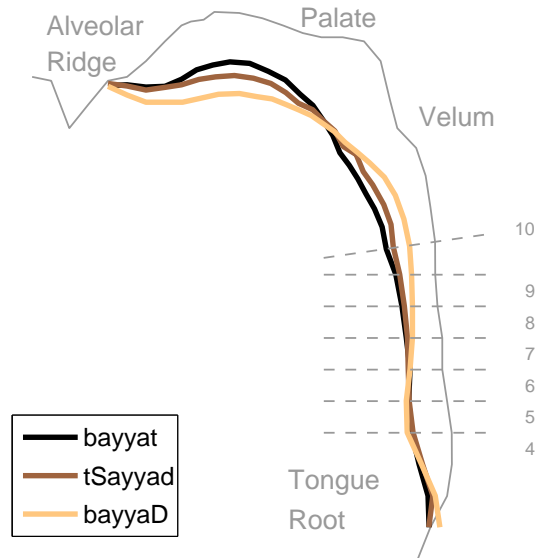


Figure 2: **Emphasis spreads rightwards across an intervening palatal segment.** Mean tongue posture during mid-consonantal production of word-final stops [t], [d] and [D], calculated from five utterances each of the words '*bayyat*', '*tSayyad*' and '*bayyaD*'. Upper-pharyngeal constriction is most pronounced in the intrinsically emphatic stop (*bayyaD*), but greater upper-pharyngeal constriction is evident in the emphaticized word-final stop (*tSayyad*), compared to the plain non-emphaticized word-final stop (*bayyat*).

an intervening palatal segment to the left of the emphatic has a greater emphaticization effect. The analysis (not included in Table 2) indicated that this is indeed the case, and that, compared to word-initial plain segments, word-initial emphaticized segments were produced with a significantly narrower constriction at all but the lowest point in the pharynx (5.6 mm vs. 6.9 mm in plain segments).

3.3. Summary of Results

- i. Emphatics show clear difference in tongue shape
 - tongue body is depressed and retracted
 - narrowest constriction in upper pharynx
- ii. A high segment /j(:)/ does not block emphasis
 - word-finally, emphaticization is weaker but significant

4. Discussion

The current study provides further support for some previous observations about the articulatory correlates of Arabic emphasis, mainly with respect to the differences in tongue shape between emphatics and their plain counterparts. These data lend support to findings of upper-pharyngeal constriction and depressed tongue dorsum as cross-dialectal characteristics of emphasis. In addition, while the robustness of emphasis marks it as a phonological phenomenon, our results suggest that there are robust differences in magnitude between *emphatics* and *emphaticized* segments in some contexts. At face-value, these differences stand in contrast to phonological accounts in which

FINAL SEGMENT	INTERVENING SEGMENT	PHARYNGEAL GRIDLINE (DISTANCE FROM GLOTTIS)								MEAN APERTURE
		4 (19.2 mm)	5 (24.0 mm)	6 (28.8 mm)	7 (33.6 mm)	8 (38.4 mm)	9 (43.2 mm)	10 (48.0 mm)		
plain	non-palatal	7.06	6.48	5.39	5.36	4.85	5.46	6.80	5.91	
	palatal	6.48	5.88	5.24	5.46	5.29	5.95	7.60	5.99	
emphaticized	non-palatal	6.96	5.92	4.88	4.49	4.10	4.02	5.21	5.08	
	palatal	6.96	6.02	5.31	5.32	5.00	5.15	6.10	5.69	
emphatic	non-palatal	6.49	4.99	3.90	3.70	3.13	3.18	4.05	4.21	
	palatal	7.21	6.39	4.76	4.63	3.94	3.47	3.96	4.91	

Table 2: *Pharyngeal Aperture in Plain, Emphaticized and Emphatic Consonants.* Distance (mm) between back of tongue and rear pharyngeal wall, measured at gridlines 4 (19.2 mm above glottis) to 10 (24.0 mm above glottis). Apertures shown for word-final consonants in six different contexts: (a) plain consonants following a non-palatal word-medial consonant (e.g. *tamman*); (b) plain consonants following a palatal word-medial consonant (e.g. *dayyan*); (c) emphaticized consonants following a non-palatal word-medial consonant (e.g. *Tamman*); (d) emphaticized consonants following a palatal word-medial consonant (e.g. *Tayyan*); (e) intrinsically emphatic consonants following a non-palatal word-medial consonant (e.g. *naffaD*); and (f) intrinsically emphatic consonants following a palatal word-medial consonant (e.g. *bayyaD*).

the spread of emphasis results in identical representations. Although further research beyond this pilot study is needed in order to establish the articulatory facts, we believe that any adequate description of the processes underlying emphasis spread should address contrasts found at this level. Finally, while it is likely that prosodic structure interacts significantly with emphasis, to the best of our knowledge, this area has not yet been explored, and remains a topic for further investigation. We believe that the use of rtMRI technology holds great promise for continued investigation into these and other important topics in Semitic phonetics and phonology.

5. Conclusions

The results of this study show that rtMRI is a promising technique for shedding light on the mechanisms of production of emphatic consonants. Alongside agreement with some previous phonetic studies of emphasis, the analysis points to differences not addressed by existing phonological accounts.

6. Acknowledgements

Research supported by NIH Grant R01 DC007124-01.

7. References

- [1] J. J. McCarthy, "The phonetics and phonology of Semitic pharyngeals," in *Laboratory Phonology III: Phonological Structure and Phonetic Form*, P. Keating, Ed. CUP, 1994, pp. 191–233.
- [2] J. C. E. Watson, "The directionality of emphasis spread in Arabic," *Linguistic Inquiry*, vol. 30, no. 2, pp. 289–300, 1999.
- [3] L. H. Ali and R. Daniloff, "A contrastive cinefluorographic investigation of the articulation of emphatic-non emphatic cognate consonants," *Studia Linguistica*, vol. 26, no. 2, pp. 81–105, 1972.
- [4] S. Ghazeli, "Back consonants and backing coarticulation in Arabic," Ph.D. dissertation, University of Texas, Austin, 1977.
- [5] F. Al-Tamimia, F. Alzoubib, and R. Tarawnah, "A videofluoroscopic study of the emphatic consonants in Jordanian Arabic," *F. Phoniatria et Logopaedica*, vol. 61, no. 4, pp. 247–253, 2009.
- [6] A. Giannini and M. Pettorino, "The emphatic consonants in Arabic," Istituto Universitario Orientale, Napoli, Speech lab. report (Vol. 4), 1982.
- [7] A. Laufer and T. Baer, "The emphatic and pharyngeal sounds in Hebrew and in Arabic," *Language and speech*, vol. 31, 1988.
- [8] F. Al-Tamimi and B. Heselwood, "Nasoendoscopic, videofluoroscopic and acoustic study of plain and emphatic coronals in Jordanian Arabic," in *Instrumental Studies in Arabic Phonetics*, Z. Hassan and B. Heselwood, Eds. John Benjamins, 2011, pp. 163–192.
- [9] B. A. Zawaydeh, "The phonetics and phonology of gutturals in Arabic," Ph.D. dissertation, Indiana University, 1999.
- [10] S. Al-Ani, *Arabic Phonology: An Acoustic and Physiological Investigation*. The Hague: Mouton, 1970.
- [11] S. Davis, "Emphasis spread in Arabic and grounded phonology," *Linguistic Inquiry*, vol. 26, pp. 465–498, 1995.
- [12] W. H. T. Gairdner, *The phonetics of Arabic. A phonetic inquiry and practical manual for the pronunciation of Classical Arabic and of one colloquial (Egyptian)*. London: OUP, 1925.
- [13] N. Trubetzkoy, *Anleitung zur phonologischen Beschreibung*. Göttingen: Vandenhoeck und Ruprecht, 1958, (Originally published 1935, Edition du Cercle Linguistique de Prague).
- [14] R. Herzallah, "Aspects of Palestinian Arabic phonology: a nonlinear approach," 1990, note: Distributed as Working Papers of the Cornell Phonetics Laboratory No. 4.
- [15] E. Broselow, "The phonology of Egyptian Arabic," Ph.D. dissertation, University of Massachusetts, Amherst, 1976.
- [16] M. A. Younes, "Emphasis spread in two Arabic dialects," in *Perspectives on Arabic Linguistics V*, M. Eid and C. Hole, Eds. John Benjamins, 1993, pp. 119–145.
- [17] —, "Problems in the segmental phonology of Palestinian Arabic," Ph.D. dissertation, University of Texas, Austin, 1982.
- [18] C. Zeroual, J. H. Esling, and P. Hoole, "EMA, endoscopic, ultrasound and acoustic study of two secondary articulations in Moroccan Arabic: labial-velarization vs. emphasis," in *Instrumental Studies in Arabic Phonetics*, Z. M. Hassan and B. Heselwood, Eds. John Benjamins, 2011, pp. 277–298.
- [19] P. Delattre, "Pharyngeal features in the consonants of Arabic, German, Spanish, French, and American English," *Phonetica*, vol. 23, pp. 129–155, 1971.
- [20] S. Narayanan, K. Nayak, S. Lee, A. Sethy, and D. Byrd, "An approach to real-time magnetic resonance imaging for speech production," *JASA*, vol. 115, no. 4, pp. 1771–1776, 2004.
- [21] E. Bresch, Y.-C. Kim, K. Nayak, D. Byrd, and S. Narayanan, "Seeing speech: Capturing vocal tract shaping using real-time magnetic resonance imaging [Exploratory DSP]," *IEEE Signal Process. Mag.*, vol. 25, no. 3, pp. 123–132, May 2008.
- [22] S. Shar and J. Ingram, "MRI investigation of Arabic gutturals," in *Proc. ICPhS*, Hong Kong, 17-21 Aug 2011, pp. 1802–1805.
- [23] E. Bresch, J. Nielsen, K. Nayak, and S. Narayanan, "Synchronized and noise-robust audio recordings during realtime MRI scans," *JASA*, vol. 120, no. 4, pp. 1791–1794, 2006.
- [24] M. I. Proctor, D. Bone, and S. S. Narayanan, "Rapid semi-automatic segmentation of real-time MRI for parametric vocal tract analysis," in *Proc. Int'l. Conf. on Speech Communication and Technology*, Makuhari, Japan, Sep 2010.
- [25] S. Narayanan, E. Bresch, P. K. Ghosh, L. Goldstein, A. Katsamanis, Y.-C. Kim, A. Lammert, M. I. Proctor, V. Ramanarayanan, and Y. Zhu, "A multimodal real-time MRI articulatory corpus for speech research," in *Proc. Int'l. Conf. on Speech Communication and Technology*, Florence, Italy, Aug 2011.
- [26] S. E. G. Öhman, "Numerical model of coarticulation," *JASA*, vol. 41, no. 2, pp. 310–320, 1967.