Journal of Phonetics 39 (2011) 85–95

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/phonetics

Acoustics and perception of emphasis in Urban Jordanian Arabic

Allard Jongman *, Wendy Herd, Mohammad Al-Masri, Joan Sereno, Sonja Combest

Linguistics Department, University of Kansas, Lawrence, KS 66044, USA

ARTICLE INFO

Article history:
Received 7 October 2009
Received in revised form
12 November 2010
Accepted 22 November 2010
Available online 12 January 2011

ABSTRACT

Acoustic and perceptual effects of emphasis, a secondary articulation in the posterior vocal tract, were investigated in Urban Jordanian Arabic. Twelve speakers of Jordanian Arabic recorded both consonants and vowels of monosyllabic minimal CVC pairs containing plain or emphatic consonants in initial and final position to investigate the extent of coarticulatory effects of emphasis. In general, the acoustic correlates of emphasis include a raised F1, lowered F2, and raised F3 in the vowel adjacent to the emphatic consonant, consistent with a narrowing near the uvula. These effects are similar in magnitude for vowels following and preceding emphatic consonants. In addition, the spectral mean of emphatic stops, but not emphatic fricatives, was lower than that of plain consonants. A perception study with cross-spliced natural stimuli explored whether Arabic listeners' recognition of emphasis is based on information in the target consonant or the rest of the word (vowel+non-target consonant). Results show that the rest of the word contributes significantly more to the perception of emphasis than the target consonant itself. Overall, the acoustic data and perceptual results will address the correlates of emphasis, spread of emphasis, and the asymmetry between stops and fricatives.

1. Introduction

Emphasis is a distinctive feature of Semitic languages such as Arabic and Hebrew. The term 'emphasis' refers to consonants produced with a secondary constriction in the posterior vocal tract and a primary constriction typically in the dental/alveolar region. Most, if not all, dialects of Arabic distinguish minimal word pairs that differ only in the presence of a plain versus an emphatic consonant. In the present paper, we focus on Urban Jordanian Arabic, spoken in the Ammani and Irbid regions of Jordan as well as in Zarqa, which distinguishes emphatic (/ḍ, ṭ, ṣ, ð̣/) from plain (/d, t, s, ð/).

Most studies of emphasis are phonological in nature, each analyzing a different dialect of Arabic. These studies highlight the observed spread of emphasis, showing emphatic consonants influencing both later segments (carry-over coarticulation) as well as earlier segments (anticipatory coarticulation). These studies have posited various articulatory correlates of emphasis, including uvularization (e.g., McCarthy, 1994), pharyngealization (e.g., Al-Ani, 1970; Davis, 1995; Younes, 1982), or pharyngealization and dorsalization (Herzallah, 1990).

While many instrumental studies of emphasis have investigated its articulatory correlates (e.g., Al-Ani, 1970; Ali & Daniloff, 1972; Ghazeli, 1977; Laufer & Baer, 1988; Zawaydeh, 1999), consensus on the exact nature of the posterior constriction is still lacking. For example, on the basis of lateral X-ray still pictures from one speaker of Tunisian Arabic, Ghazeli (1977) reports that emphasis involves retraction of the tongue toward the posterior pharynx wall at the level of the second cervical vertebra which is near the uvula. In addition, the palatine dorsum of the tongue is slightly depressed. This is consistent with data from Zawaydeh (1999)'s endoscopic study of one speaker of Ammani Jordan and from Al-Tamimi, Alzoubi, and Tarawnah (2009)'s videofluoroscopic study of four speakers of Jordanian Arabic. Both studies conclude that emphatics involve tongue root retraction in the oropharynx at the level of the second vertebra (uvula). Zawaydeh (1999) also had observed that the pharynx itself is also actively involved in the narrowing of the pharyngeal cavity for emphatics. In contrast, Giannini and Pettorino (1982) and Laufer and Baer (1988) report that emphasis involves a constriction much lower in the pharynx than the uvula. Specifically, in their X-ray study of one speaker of Iraqi Arabic, Giannini and Pettorino (1982) find the greatest constriction in the lower pharynx, at the level of cervical vertebrae 4 and 5. Laufer and Baer (1988) positioned a fiberoptic endoscope in the pharynx near the level of the rest position of the uvula in 7 speakers representing 6 Arabic dialects. They conclude that in emphatic sounds, the epiglottis forms a constriction with the pharynx wall and they observed that the tongue root moves back as well. Articulatorily, both uvularization and pharyngealization have been proposed for emphatics.

When an underlying emphatic consonant occurs in a word, the coarticulatory effects of emphasis typically affect neighboring sounds. However, the pattern of emphasis spread seems to vary across dialects. Phonological analysis has shown that in some dialects, such as Cairene Arabic, emphasis spreads throughout the

* Corresponding author. Tel.: +1 785 864 2384; fax: +1 785 864 5724.
E-mail address: jongman@ku.edu (A. Jongman).

0095-4470/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.wocn.2010.11.007

entire word. In others, emphasis rarely spreads beyond an adjacent vowel (see, for example, Younes (1993), on the Abha dialect of southern Saudi Arabia). In addition, asymmetries between rightward and leftward emphasis spread have been reported. For example, in a rural northern Palestinian dialect, emphasis seems to spread leftward from the emphatic consonant to the beginning of the word while rightward emphasis spread is usually restricted to a following adjacent vowel. Furthermore, it has been claimed that certain phonemes block rightward, but not leftward, emphasis spread (Herzallah, 1990). Finally, in Qatari Arabic, a dialect in which emphasis spreads in both directions throughout the entire word, emphasis may spread leftward across a word boundary into the preceding word if the target-word-initial consonant is emphatic (Bukshaisha 1985 in Watson, 1999). In general, then, phonological analyses suggest that rightward emphasis spread seems much more restricted than leftward spread (Davis, 1995), an interesting fact given that most instrumental research has focused on rightward spread.

Phonetically, the coarticulatory effects of emphasis have received little attention. In their early articulatory study of Iraqi Arabic, Ali and Daniloff (1972) report greater rightward than leftward spread, both in terms of the magnitude of the tongue backing gesture and in terms of the number of segments affected. However, emphasis never spreads throughout the entire word. For example, in a C̣VC word, the initial emphatic consonant was found to affect the vowel but not the final consonant. In a C̣VCVCVC word, emphasis was observed on the initial C̣VCV but not on any segments of the final syllable. In her acoustic study of Jerusalem Arabic, Card (1983) rejects Ali and Daniloff's earlier interpretation and, based on acoustic data, argues for the word as the domain of emphasis spread. Measurements of F2 of consonants and vowels suggested no asymmetry in spread, as emphasis spread both rightward and leftward to the end and beginning of the word, respectively.

Zawaydeh (1999) objected that the words Card used did not really allow for a test of rightward or leftward spread. Investigating emphasis spread in Ammani-Jordanian Arabic using four words, Zawaydeh's F2 measurements of the vowels indicated that the word is the domain of emphasis spread. Rightward and leftward spread were found but not to the same extent. Leftward spread did not display any gradiency: F2 of the vowel was equally low regardless of its distance from the emphatic consonant. In contrast, rightward spread occurred in a gradient manner: F2 lowered for vowels that were closer to the emphatic consonant. Consonant measures were not provided. These studies together provide conflicting data as to the nature of the rightward and leftward spread of emphasis.

The present acoustic and perceptual study is a comprehensive instrumental investigation of emphasis for a single dialect of Arabic. For the acoustic analyses, both stop and fricative emphatics are included and three distinct vowels, both short and long, are examined. Moreover, F1, F2, and F3 are inspected throughout the vowel and the non-target consonant is also examined in rightward and leftward directions. These measurements will provide not only relevant information about the direction and extent of the influence of the emphatic consonant but also critical data on the location of the secondary constriction. The perceptual data will further provide unique data on the relative contribution of consonantal and vocalic information to the emphatic and plain distinction. Emphatics provide an excellent domain in which to investigate anticipatory and carry-over coarticulation. The acoustic and perceptual data of the present study systematically document the coarticulatiory effects of emphasis in Arabic.

While linguists consider consonants as the primary locus of emphasis and speak of emphatic consonants, most previous acoustic analyses of emphasis have focused on properties of the vowels surrounding the emphatic consonant rather than the consonant itself. In all dialects of Arabic that have been instrumentally investigated, emphasis is consistently manifested by a lowering of F2 of the vowel following the emphatic consonant (e.g., Card, 1983; Zawaydeh, 1999). In her study of four Palestinian Arabic speakers, Card (1983) reports a 200–300 Hz lowering of F2 of vowels in an emphatic context as compared to a plain, non-emphatic context. This lowering was shown to be quite systematic, although these results were not subjected to statistical analysis. This F2 lowering has been corroborated for other dialects of Arabic, including Ammani-Jordanian Arabic (Al-Masri & Jongman, 2004; Khattab, Al-Tamimi, & Heselwood, 2006; Zawaydeh, 1999), Cairene Arabic (Kahn, 1975), Egyptian Arabic (Norlin, 1987; Wahba, 1993), Lebanese Arabic (Obrecht, 1968), Moroccan Arabic (Alioua, 1995; Yeou, 1997), and Tunisian Arabic (Ghazeli, 1977). While most of these studies only included a small number of speakers from a sampling of dialects and they often focused on a small set of words, all have shown a consistent lowering of F2 at the midpoint of the vowel that immediately follows an emphatic consonant.

However, while emphasis clearly lowers F2 of the following vowel, its effect seems to be modified by both vowel quality and vowel length. As observed by Card (1983), Alioua (1995), and Yeou (1997), the effect of emphasis (F2 lowering as measured at vowel midpoint) differs for the vowels /æ/, /i/, and /u/. The greatest lowering occurs for the low front vowel /æ/, followed by the vowels /i/ and /u/. In fact, F2 lowering for /æ/ results in a vowel with a distinctly different quality (the low back vowel [ɑ] occurs as an allophone of the low front vowel /æ/ in an emphatic context) which may be why many studies have limited themselves to this vowel. In addition, Yeou (1997) reported greater F2 lowering at vowel midpoint in short vowels relative to long vowels following emphatic consonants. However, a direct comparison of emphatic influence across long and short vowels in these studies is not possible since F2 measures were examined at the midpoint of each vowel.

While vowel F2 has received most attention, the patterning of F1 and F3 may also provide important information about the exact location of the posterior constriction. Specifically, while a uvular or pharyngeal constriction both result in a low F2 and a high F1, vocal tract modeling studies suggest that F2 would be lower for uvulars while F1 would be less high for a uvular constriction relative to a pharyngeal one. On the other hand, F3 would be lower for a pharyngeal constriction (Alwan, 1986; Klatt & Stevens, 1969). However, only a few studies have critically included F1 or F3 measures, with contradictory findings. Of those studies, Alioua (1995), Yeou (1997), and Zawaydeh (1999) show a rise in F1, suggesting a more pharyngeal constriction, while Card (1983) and Norlin (1987) report no consistent effect of emphasis on F1 or F3.

Acoustic properties of emphatic consonants themselves have received scant attention in the literature. For stop consonants, Card (1983) reports F2 values taken from spectrograms, but it is not clear at which point (e.g., at release burst or onset/offset of the formant transition) these measurements were taken. Her data (not subjected to statistical analysis) suggest that F2 in emphatic consonants may be lower than in their plain counterparts. For fricatives, Card focused on the bottom cut-off frequency of the frication noise on spectrograms. Card found no correlation between the cut-off frequency and the presence versus absence of an emphatic fricative and therefore did not report this measure. Kahn (1975) similarly reports no difference between plain /s/ and emphatic /ṣ/ when measuring the bottom cut-off frequency. Norlin (1987) compared the spectral center of gravity in fricatives as measured with a 26.5 ms window taken after the first third of the fricative. Norlin concludes that emphatic /ṣ, ẓ/ in Egyptian Arabic have a concentration of energy at lower frequencies than plain /s, z/. However, *t*-tests performed by the present authors on the means included in the appendix (Norlin, 1987) suggest that these comparisons are not statistically significant. Most recently, Al-Khairy (2005) measured the spectral mean at various locations in Arabic fricatives produced by male Saudi speakers of Modern Standard Arabic and reported no significant differences between plain and emphatic fricatives. It seems that there are few acoustic effects of emphasis in consonants.

While the acoustic analyses of emphasis in Arabic have shown consistent lowering of second formant frequencies following emphatic consonants, very little is known about the perceptual correlates of emphasis. In a pioneering study, Obrecht (1968) investigated the role of the F2 transition from an initial consonant to the following vowel. Using the Haskins pattern playback synthesizer to create synthetic speech modeled after a male native speaker of Lebanese Arabic, Obrecht explored different segments, mostly in monosyllabic words. In general, F2 was systematically varied in 120 Hz steps from 1080 to 1800 Hz. The transition duration of F2 was fixed at 70 ms while overall syllable duration ranged from 300 to 400 ms. F3 was flat and fixed at 3000 Hz, and F1 was flat (for voiceless consonants) and varied as a function of the particular vowel used in a given experiment. Results from a /ṭiː–tiː/ continuum created this way suggested that perception of the emphatic/plain distinction was quite categorical. The continuum endpoints with low and high F2 values were perceived as /ṭiː/ and /tiː/, respectively. The category boundary fell around an F2 value of 1560 Hz. Yeou (1995) confirmed the role of F2 in a synthetic /ṣiː–siː/ continuum and further showed that F1 alone was not a perceptual cue to the emphatic/plain distinction.

Ali and Daniloff (1974), using naturally produced utterances, compared minimal plain/emphatic Iraqi Arabic word pairs in which the target consonant occurred in either word-initial or word-final position. Truncated stimuli, from which the plain and emphatic target segments, as well as their transitions, had been removed, were presented to listeners who indicated on an answer sheet from which member of the word pair the truncated word stem was derived. Results showed that listeners correctly identified the missing consonant with about 68% accuracy (73% for emphatic stems, 62% for plain stems; both significantly better than chance). This finding suggests that the emphatic consonant itself is not required for reasonably accurate perception of the plain/emphatic distinction. In contrast to the acoustic data, Ali and Daniloff (1974) report that neither vowel quality nor consonant type had an effect, nor did position of the target consonant.

The studies summarized above are among the very few that considered the perceptual correlates of emphasis. While some studies suggest that F2 transition is a strong perceptual cue, listeners seem to perceive the distinction between plain and emphatic consonants with reasonable accuracy in the absence of transition information. Moreover, some of the perceptual results are not consistent with the findings from acoustic studies. This mismatch may be partially due to differences across studies in terms of the dialects, consonants, and vowels that were included.

The present study is an acoustic and perceptual examination of emphasis for a single dialect of Arabic (Urban Jordanian Arabic). In Experiment 1, the acoustic correlates of emphasis were explored by measuring a number of acoustic parameters in both emphatic consonants and their surrounding vowels and consonants using six male and six female speakers. For Experiment 1, both emphatic and plain consonants, including both stops and fricatives, were examined in initial position and in final position. The first spectral moment ('spectral mean') was measured for the consonants. Inclusion of measures of the consonants themselves allowed us to compare the effect of emphasis in consonants as well as to assess the extent of emphasis spread. Acoustic correlates of emphasis were also examined in the adjacent vowel. While previous studies have shown that F2 of the following or preceding vowel usually is affected by emphasis, the detailed nature of this effect is less clear. In the present study, F1, F2, and F3 were measured at the onset, midpoint, and offset of all vowels, using three different vowel contexts and both long and short vowels, to detail anticipatory and carry-over effects. Inclusion of F1 and F3 in addition to F2 for the vowels allowed us to infer the nature and place of constriction of the secondary articulation as well as explore effects of vowel quality in more detail. In the present acoustic study, measurements of both stops and fricative consonants and measurements at multiple locations in the vowel as well as measurements of non-target consonants will provide a more fine-grained account of the spread of emphasis and the roles of vowel length and vowel quality in this process.

A perception experiment (Experiment 2) was then conducted using cross-spliced speech. To evaluate the role of the emphatic consonant as well as the preceding or following vowel and consonant as perceptual cues to the emphatic/plain distinction, thirty native Jordanian speakers listened to the cross-spliced natural speech stimuli. The use of conflicting-cue stimuli (e.g., emphatic consonant followed by vowel or consonant taken from non-emphatic environment) allowed for the assessment of the relative contribution of consonantal and vocalic information to the emphatic/plain distinction.

2. Experiment 1: Acoustic study

Jordanian Arabic is typically divided into Urban, Rural, Bedouin, and Ghorani Arabic. The present study focused on a single dialect of Arabic to rule out effects due to dialectal variation. For this purpose, Urban Jordanian Arabic, spoken in the Ammani and Irbid regions of Jordan as well as in Zarqa, was used. This dialect is spoken by people living in the major cities of Jordan—more than two-thirds of the population of Jordan (see Zuraiq and Zhang (2006) for a phonological description of this dialect).

2.1. Method

2.1.1. Stimulus materials

The four emphatics (2 stops and 2 fricatives) of Jordanian Arabic /ḍ, ṭ, ð̣, ṣ/ and their plain (non-emphatic) counterparts /d, t, ð, s/ were recorded in target word pairs in the carrier phrase [ˈʔiħki___kæˈmæːn ˈmɑrːɑh] ("Say ____ once more") to control for context effects. The stimuli were printed on notecards in Arabic script. Stimuli consisted of monosyllabic words and non-words. In Arabic script, emphatics are realized by a different orthographic symbol, so native speakers were able to produce non-words with both emphatic and plain consonants.

The stimuli contained the target consonant in word-initial position and word-final position. Specifically, for word-initial fricatives, the following syllable frames were used: /s_b/, /s_q/, /s_x/, /s_ħ/, and /ð_b/; frames for initial stops were /d_b/ and /t_b/. Frames for word-final target consonants were the mirror images. For all consonant contexts, both long vowels /iː, æː, uː/ and short vowels /i, æ, u/ were used. Each stimulus was repeated three times.

The number of plain-emphatic stimulus pairs per speaker was 252 for a total of 504 stimuli per speaker. For each of these emphatic and plain pairs, 126 occurred in initial position and 126 occurred in final position (7 syllable frames $\times$ 6 vowels $\times$ 3 repetitions).

2.1.2. Participants

Twelve speakers (6 females and 6 males) were recorded. All were native speakers of the Irbid dialect of Jordanian Arabic with no known history of either speech or hearing impairment.

2.1.3. Recordings

Speakers were recorded in a quiet room in the Department of English Language and Literature at the Yarmouk University Speech

& Hearing Laboratory in Irbid, Jordan using a high-quality microphone (ElectroVoice N/D767a) and a digital solid-state recorder (Marantz PMD 671). Sampling rate was 22 kHz with 16-bit quantization.

2.1.4. Measurements

All measurements were made with Praat speech analysis software (Boersma & Weenink, 2007). Segmentation was based on waveforms and wide-band spectrograms. The onset of the initial consonant was taken as the offset of the vowel /i/ of the preceding word in the carrier phrase. Vowel onset was defined as the clear emergence of F1; vowel offset was taken as the point at which F2 substantially weakened or disappeared from the spectrogram. The onset of the final consonant was defined as the offset of the preceding vowel. The offset of the final consonant corresponded to the end of the release burst for stops and the end of frication for fricatives.

Formant frequency measures (F1–F3) were taken from LPC spectra calculated over a 20-ms Hamming window at the beginning, middle, and end of the vowel.

Spectral mean was computed for obstruents following the procedures described in Forrest, Weismer, Milenkovic, and Dougall (1988) and Jongman, Wayland, and Wong (2000). For fricatives, a DFT was calculated using a 20-ms Hamming window in the middle of the frication noise. For stops in initial position, a variable-sized Hamming window included the burst and release portion (i.e., from burst to vowel onset). For final stops, a 20-ms full Hamming window was centered over the burst. In addition, for all obstruents, a second window was centered over the boundary between consonant and vowel. Briefly, each DFT was treated as a random probability distribution from which the first moment was computed. In the statistical analyses, the data for the three repetitions of each token were averaged.

2.2. Results: vowel data

2.2.1. Word-initial target position

Four-way Repeated Measures Analyses of Variance (ANOVA) with Emphasis, Manner of Articulation Context, Vowel Quality, and Vowel Length as independent variables were conducted for each formant frequency at each of the three locations (onset, middle, and offset) in the vowel for target consonants in word-initial position.

2.2.1.1. Emphasis. As shown in Fig. 1 (left panel), vowels adjacent to an initial emphatic consonant had a consistently higher F1, lower F2, and higher F3 than vowels adjacent to plain consonants. F1 at vowel onset was significantly higher following an emphatic consonant [$F(1, 11)=61.49$, $p=.000$]; F1 at vowel midpoint was also significantly higher following an emphatic consonant [$F(1, 11)=40.74$, $p=.000$]. However, at vowel offset, this difference did not reach significance [$F(1, 11)=1.38$, $p>.263$]. F2 was significantly lower following an emphatic consonant throughout the entire

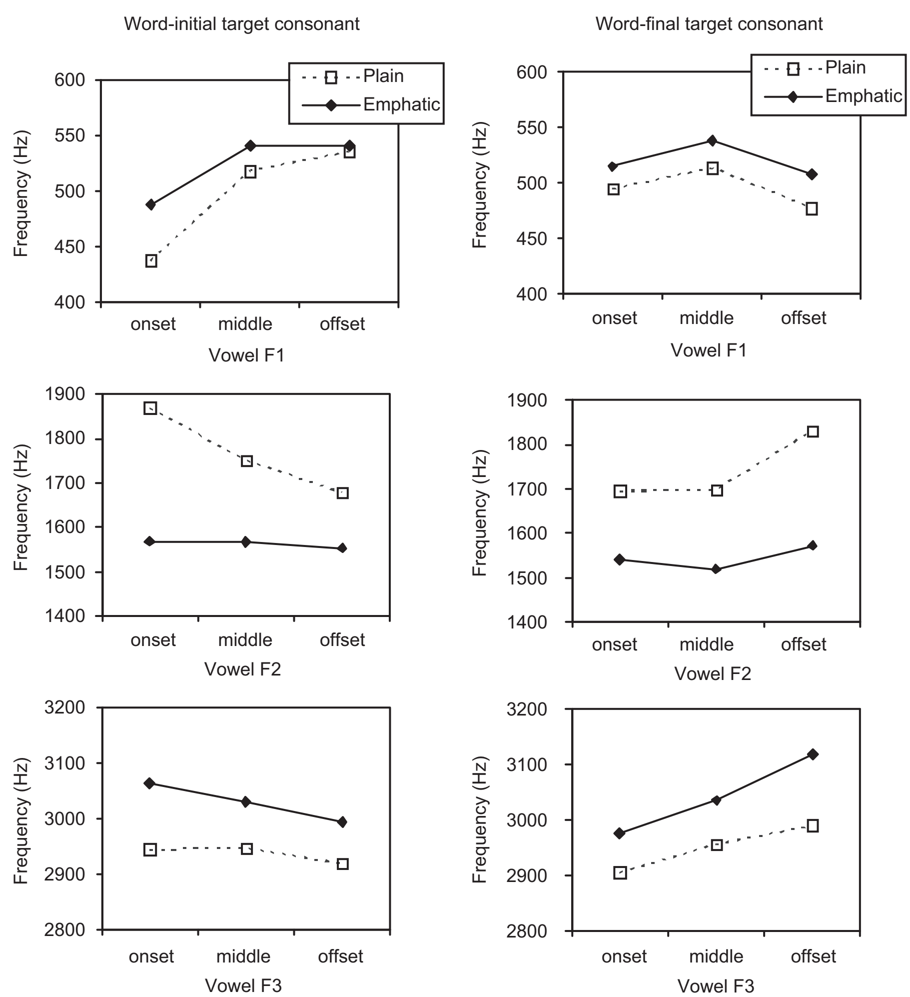

Fig. 1. Formant frequency values (F1 in top panel, F2 in middle panel, F3 in bottom panel) measured at onset, middle, and offset of the vowel in CVC words with the target consonant (plain vs. emphatic) in **word-initial** (left column) and **word-final** (right column) position.

vowel: vowel onset [$F(1, 11)=51.68$, $p=.000$], vowel midpoint [$F(1, 11)=61.24$, $p=.000$], and vowel offset [$F(1, 11)=55.69$, $p=.000$]. Also, F3 was significantly higher following an emphatic consonant throughout the entire vowel: vowel onset [$F(1, 11)=27.80$, $p=.000$], vowel midpoint [$F(1, 11)=46.37$, $p=.000$], and vowel offset [$F(1, 11)=96.24$, $p=.000$].

2.2.1.2. Vowel Quality. As expected, there was also a main effect of Vowel Quality at all three locations in the vowel. F1 was highest for /æ/ and lowest for /i/. For F2, there was an expected main effect of Vowel Quality throughout the vowel, with a significantly higher F2 for /i/ than for /æ/ and /u/ (*F*-values ranging from 82.44 to 679.17, all $p=.000$). F3 showed small but significant differences as a function of vowel quality with the highest value for /u/ and the lowest for /æ/ (*F*-values ranging from 5.59 to 15.55, *p*-values ranging from .023 to .001).

For F1 in the middle of the vowel, there was a significant Vowel Quality by Emphasis interaction [$F(2, 10)=5.31$, $p=.027$], indicating that the effect of emphasis on F1 was more pronounced for the vowel /æ/ than for /i/ and /u/.

For F2, a significant Vowel Quality by Emphasis interaction for all three locations in the vowel indicated that emphasis consistently had a stronger effect for the vowels /æ/ and /i/ than for /u/ (*F*-values ranging from 9.16 to 33.86, *p*-values ranging from .005 to .000). Fig. 2 (top panel) shows the difference between F2 values in plain and emphatic contexts for each vowel quality (averaged across long and short vowels) at three locations in the vowel. It is clear that the effect of emphasis continues throughout the entire vowel for /æ/ but rapidly decreases for /i/ and /u/ as measurements are taken further away from the initial target consonant.

For F3, a significant Vowel Quality by Emphasis interaction revealed that emphasis had a significant effect for the vowels /æ/ and /u/ but not for /i/ throughout the vowel (*F*-values ranging from 14.54 to 20.17, *p*-values ranging from .001 to .000).

2.2.1.3. Vowel Length. Each formant frequency exhibited a main effect for Vowel Length: throughout the vowel, short vowels had a significantly higher F1 (*F*-values ranging from 21.85 to 342.08, *p*-values ranging from .001 to .000) and a lower F2 (*F*-values ranging from 28.00 to 61.73, all $p=.000$) than long vowels. Short vowels generally also had a lower F3; this effect was significant at vowel onset ([$F(1, 11)=10.53$, $p=.008$]) and midpoint ([$F(1, 11)=20.04$, $p=.001$]), and close to significance at vowel offset ([$F(1, 11)=4.16$, $p=.066$]).

Significant Vowel Length by Emphasis interactions for F2 and F3 indicated that the effect of emphasis seemed to be more pronounced in short than long vowels (*F*-values ranging from 4.88 to 19.72, all $p=.000$) throughout the vowel, with the exception of F2 at vowel onset. However, while these statistical analyses compared the effect of emphasis in short and long vowels at the same relative location (e.g., vowel midpoint), the absolute measurement locations could differ because of intrinsic duration differences between short and long vowels. Overall, long vowels (145 ms) were about 1.8 times longer than short vowels (80 ms). Thus, the significantly greater effect of emphasis on short vowels than on long vowels as observed at vowel midpoint and offset could be due to the fact that the temporal interval between the target consonant and measurement location is smaller for short as compared to long vowels. Only measurements taken at vowel onset are not subject to this potential confound. Here, only F3 showed a significant effect of vowel length. A significant Emphasis by Vowel Length interaction ([$F(1, 11)=30.98$, $p=.000$]) indicated that the difference in F3 between emphatic and plain contexts was significantly greater in short vowels (157 Hz) than in long vowels (80 Hz). F1 and F2 showed no such difference.

2.2.1.4. Manner of Articulation Context. In terms of Manner of Articulation Context, no significant main effects were observed in onset position for F1, F2, or F3. Vowel onset values were similar following a fricative as compared to a stop. Nevertheless, significant Manner × Emphasis interactions indicated that, in general, the effect of emphasis seemed to be more pronounced in the environment of stops than fricatives. This was true for F2 throughout the vowel (*F*-values ranging from 7.2 to 54.6, with *p*-values ranging from .000 to .021).

2.2.2. Word-final target position

Four-way Repeated Measures Analyses of Variance (ANOVA) with Emphasis, Manner of Articulation Context, Vowel Quality, and Vowel Length as independent variables were conducted for each formant frequency at each of the three locations (onset, middle, and offset) in the vowel for target consonants in word-final position.

For target consonants in word-final position, all three formant frequencies were significantly affected by the emphatic consonant throughout the entire vowel. As was the case for word-initial target consonants, vowels preceding an emphatic consonant had a consistently higher F1, lower F2, and higher F3 than vowels preceding plain consonants (see Fig. 1, right panel).

2.2.2.1. Emphasis. Four-way Repeated Measures Analyses of Variance (ANOVA) with Emphasis, Manner of Articulation Context, Vowel Quality, and Vowel Length as independent variables were conducted for each formant frequency at each of the three locations (onset, middle, offset) in the vowel for target consonants in word-final position. F1 was significantly higher preceding an emphatic consonant throughout the entire vowel (vowel onset [$F(1, 11)=35.52$, $p=.000$]; vowel midpoint [$F(1, 11)=15.37$, $p=.002$]; vowel offset [$F(1, 11)=7.15$, $p=.022$]). F2 was significantly lower preceding an emphatic consonant throughout the entire vowel (vowel onset [$F(1, 11)=60.98$, $p=.000$]; vowel midpoint [$F(1, 11)=47.35$, $p=.000$]; vowel offset [$F(1, 11)=35.27$, $p=.000$]). Finally, F3 was significantly higher preceding an emphatic consonant throughout the entire vowel (vowel onset [$F(1, 11)=61.08$, $p=.000$]; vowel midpoint [$F(1, 11)=52.05$, $p=.000$]; vowel offset [$F(1, 11)=23.92$, $p=.000$]).

2.2.2.2. Vowel Quality. Similar to word-initial position, there was also a main effect of vowel quality. Vowel quality effects on F1 – F3 patterned similarly to those observed in word-initial position (*F*-values ranging from 18.82 to 314.12, all $p=.000$).

There were also significant Vowel Quality by Emphasis interactions for all three formant frequencies at all locations in the vowel (*F*-values ranging from 3.73 to 84.96, *p*-values ranging from .036 to .000). In general, the effect of emphasis was more pronounced for the vowel /æ/ than for /i/ and /u/. However, at vowel offset, /i/ showed the greatest effect for F2 while /u/ showed the greatest effect for F3. Fig. 2 (bottom panel) shows F2 values for each vowel quality (averaged across long and short vowels) in plain and emphatic contexts at three locations in the vowel. The vowels /æ/ and /i/ show clear effects of emphasis; this effect continues throughout the entire vowel for /æ/ but rapidly decreases for /i/ as measurements are taken further away from the final target consonant. The vowel /u/ is virtually unaffected.

2.2.2.3. Vowel Length. As was the case for word-initial target position, short vowels had a significantly higher F1 and a lower F2 and F3 than long vowels (*F*-values ranging from 11.67 to 183.30, all $p=.000$). There was one exception: there was no significant difference between F3 at vowel offset for short and long vowels ($p=.541$).

Significant Vowel Length × Emphasis interactions for F1 and F3 at vowel onset and for all three formants at vowel midpoint seem to suggest that the effect of emphasis was more pronounced in short than long vowels (*F*-values ranging from 6.90 to 22.10, *p*-values

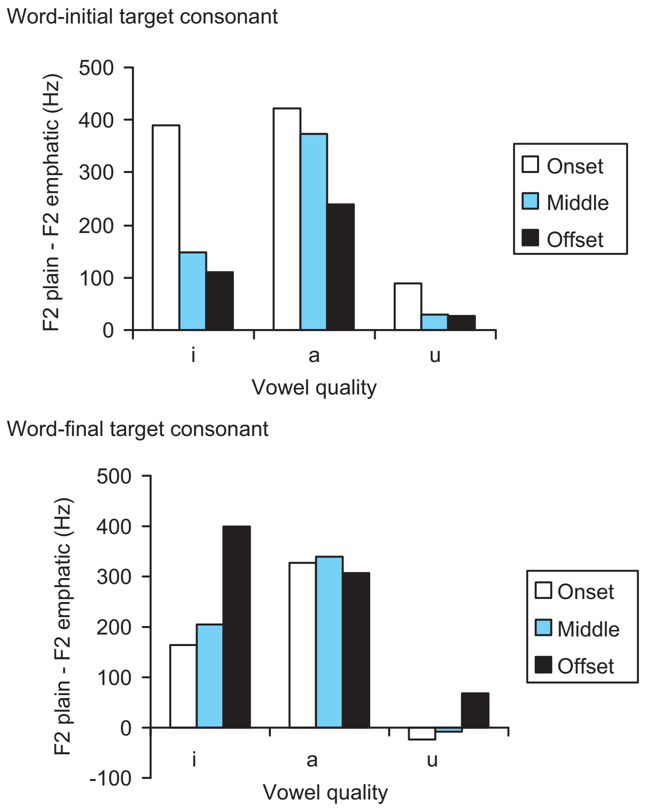

Fig. 2. Differences in F2 for each vowel quality (F2 in plain context-F2 in emphatic context) measured at onset, middle, and offset of the vowel in CVC words with the target consonant (plain vs. emphatic) in **word-initial** (top panel) and **word-final** (bottom panel) position.

ranging from .001 to .024). Again, to avoid a potential confound with distance from the target consonant, the effect of emphasis on vowel length should only be directly compared at one location for the word-final target consonant, namely at vowel offset, immediately adjacent to the target. Here, there were no significant interactions between Emphasis and Vowel Length for any of the three formants, suggesting that emphasis shows similar effects for short and long vowels.

2.2.2.4. Manner of Articulation Context. In terms of Manner of Articulation Context, no significant main effects were observed for F1, F2, and F3 at vowel offset.

However, significant Manner by Emphasis interactions indicated that, in general, the effect of emphasis was more pronounced in the environment of stops than fricatives. This was true for F2 throughout the vowel (F-values ranging from 29.76 to 94.26, all $p=.000$).

2.3. Results: consonant data

2.3.1. Word-initial target consonants

A four-way Repeated Measures Analyses of Variance (ANOVA) with Emphasis, Manner of Articulation, Vowel Quality, and Vowel Length as independent variables and spectral mean of the consonant as the dependent variable was conducted for target consonants in word-initial position. A main effect of Emphasis ([$F(1, 11)=16.36$, $p=.002$]) indicated that the spectral mean of emphatic consonants (6487 Hz) was significantly lower than that of plain consonants (6904 Hz). A main effect of Manner ([$F(1, 11)=34.06$, $p=.000$]) indicated that the spectral mean of stop consonants (5522 Hz) was significantly lower than that of fricatives (7869 Hz). Finally, a main effect of Vowel Quality ([$F(2, 10)=31.53$, $p=.000$]) and subsequent Bonferroni posthoc tests indicated that the spectral mean of consonants was significantly higher in the context of the vowel /i/ (7043 Hz) than the vowel /æ/ (6711 Hz) which was in turn significantly higher than that of consonants in the context of /u/ (6333 Hz).

A significant Emphasis by Manner interaction ([$F(1, 11)=21.09$, $p=.001$]) indicated that the main effect of Emphasis was carried by the stop consonants only. While emphatic stops (5133 Hz) had a significantly lower spectral mean than plain stops (5912 Hz) emphatic fricatives (7842 Hz) did not differ significantly from plain fricatives (7896 Hz). No significant interactions involving vowel length or vowel quality were observed.

The observed difference between stop consonants and fricatives could be due to the location of the analysis window. The spectral mean for fricatives was calculated in the middle of the fricative noise. For stops, the analysis window included the burst to the onset of the following vowel. As a result, measurements for stops included information closer to the vowel. The mean duration of the analyzed portions was 22, 16 ms for emphatic stops (ranging from 4 to 52 ms) and 27 ms for plain stops (ranging from 7 to 106 ms). It is possible that fricatives show a similar effect as stops if their spectral mean were measured closer to the vowel. However, a one-way Repeated Measures ANOVA on spectral means derived over the final 20 ms of the fricative did not reveal a significant effect of Emphasis [$F(1, 11)=.135$, $p=.720$]: the spectral mean of emphatic fricatives (5326 Hz) was not significantly different from that of plain fricatives (5349 Hz) when measured at fricative offset. Thus, the observed difference between stops and fricatives is not due to differences in the location at which the spectral mean was computed.

2.3.2. Word-final target consonants

A four-way Repeated Measures Analyses of Variance (ANOVA) with Emphasis, Manner of Articulation, Vowel Quality, and Vowel Length as independent variables and spectral mean of the consonant as the dependent variable was conducted for target consonants in word-final position. A main effect of Emphasis ([$F(1, 11)=22.91$, $p=.001$]) indicated that the spectral mean of word-final emphatic consonants (7162 Hz) was significantly lower than that of plain consonants (7555 Hz). A main effect of Manner ([$F(1, 11)=128.97$, $p=.000$]) indicated that the spectral mean of stop consonants (6505 Hz) was significantly lower than that of fricatives (8212 Hz). Finally, a main effect of Vowel Quality ([$F(2, 10)=35.73$, $p=.000$]) and subsequent Bonferroni posthoc tests indicated that the spectral mean of consonants was significantly higher in the context of the vowels /æ/ (7526 Hz) and /i/ (7618 Hz) than the vowel /u/ (6933 Hz).

A significant Emphasis by Manner interaction ([$F(1, 11)=14.08$, $p=.003$]) indicated that the main effect of Emphasis was carried by the stop consonants only. Similar to word-initial targets, while emphatic stops (6160 Hz) had a significantly lower spectral mean than plain stops (6850 Hz) emphatic fricatives (8163 Hz) did not differ significantly from plain fricatives (8261 Hz). No significant interactions involving vowel length or vowel quality were observed.

2.4. Results: spread of emphasis

The vowel formant frequency data indicated that the effect of emphasis extended throughout the vowel. To investigate the extent of the spread of emphasis, the spectral means of the non-target consonants were also analyzed. This allowed us to determine if emphasis spreads beyond the vowel into the following non-target consonant (for word-initial targets) or preceding non-target consonant (for word-final targets).

To assess the extent of rightward spread from word-initial target consonants to word-final non-target consonants, a two-way Emphasis by Manner Repeated Measures ANOVA on plain word-final non-target consonants preceded by emphatic or plain word-initial target consonants indicated neither a significant main effect of Emphasis or Manner nor a significant Emphasis by Manner interaction [all F's <1]. There was no significant difference in spectral mean between final obstruents that were preceded by emphatic obstruents (5814 Hz) compared to final obstruents preceded by plain obstruents (5733 Hz).

To assess the extent of leftward spread from word-final target consonants to word-initial non-target consonants, a two-way Emphasis by Manner Repeated Measures ANOVA on plain word-initial consonants followed by emphatic or plain word-final consonants indicated significant main effects of both Emphasis [$F(1, 11)=11.04$, $p=.007$] and Manner [$F(1, 11)=21.86$, $p=.001$]. There was a significant difference in spectral mean between initial non-target obstruents that were followed by emphatic obstruents (4147 Hz) compared to initial non-target obstruents that were followed by plain obstruents (4421 Hz). There also was a significant difference between non-target stops (3774 Hz) and non-target fricatives (4794 Hz). A marginal Emphasis by Manner interaction [$F(1, 11)=3.49$, $p=.089$] suggested that significant spectral lowering occurred only in stops (3550 Hz in emphatic context compared to 3998 Hz in plain context); for fricatives, no significant difference obtained (4745 Hz in emphatic context compared to 4844 Hz in plain context).

2.5. Summary of acoustic data

In general, F1 is raised, F2 is lowered, and F3 is raised in vowels that are either preceded or followed by emphatic obstruents, as compared to plain obstruents. This effect was observed throughout the vowel. Except for F1 at vowel offset, vowels adjacent to an initial emphatic consonant had a higher F1, lower F2, and higher F3 than vowels adjacent to a plain initial consonant.

Vowel quality was seen to be an important factor in that the effect of emphasis continues throughout the vowel /æ/ but rapidly diminishes for /i/ and /u/ as distance from the target consonant increases.

The vowel length distinction manifested itself not only in duration but also in spectral quality. Long vowels were approximately 80% longer than their short counterparts. In addition, long vowels had a lower F1 and higher F2 than short vowels. Significant Emphasis by Vowel Length interactions suggested that emphasis was stronger in short than long vowels. Because of intrinsic duration differences between short and long vowels, there are only two measurement locations, vowel onset for target consonants in initial position and vowel offset for target consonants in final position, at which the effect of emphasis in short and long vowels could be directly compared. Focusing on vowel onset for target consonants in initial position, only F3 showed a significant effect of vowel length. Measurements taken at vowel offset for target consonants in final position showed no effect of vowel length on emphasis. Overall, then, with the exception of F3 at vowel onset, the results suggest that the effect of emphasis is comparable in short and long vowels.

For both word-initial and word-final consonants, the spectral mean of emphatic consonants was significantly lower. However, this lowering was observed for stops but not fricatives. The spectral mean of emphatic stops was significantly lower than that of plain stops and this was true for stops in both word-initial and word-final positions.

Finally, the spread of emphasis from an emphatic target consonant through the vowel to the non-target consonant was detected in stops but not in fricatives, and only for non-target stops preceding emphatic final consonants. The acoustic results thus generally show very similar effects for word-initial and word-final target consonants. However, they also show a significant difference in the direction of emphasis spread, with stronger effects from a word-final emphatic target consonant to an initial non-target consonant.

3. Experiment 2: Perceptual study

A perception experiment was conducted with natural cross-spliced speech to evaluate the perceptual role of the acoustic patterns. Specifically, this experiment allowed us to explore the extent to which consonantal and vocalic information contribute to perception of the emphatic/plain distinction. Stops and fricatives in word-initial and word-final position were manipulated in a variety of vowel contexts.

3.1. Methods

3.1.1. Stimuli

The perception experiment included a subset of the stimuli from the acoustic study. Only real words were used. The 9 plain-emphatic minimal word pairs (5 word-initial, 4 word-final) are shown in Table 1.

These words were selected from two of the speakers from the acoustic study, one female and one male. These particular speakers were selected because native speakers judged them as highly intelligible. The acoustic characteristics of the selected stimuli pattern the same as the overall results reported in Experiment 1.

Table 1
Stimulus words used in the perception experiment. For each word, a plain and emphatic version was created by cross-splicing (see text for details).

Word-initial				**Word-final**			
Plain		Emphatic		Plain		Emphatic	
Word	Gloss	Word	Gloss	Word	Gloss	Word	Gloss
sæb	he cursed	ṣɑb	he poured	bæs	enough!	bɑṣ	he saw
sæːb	he left	ṣɑːb	he scored	bæːs	he kissed	bɑːṣ	bus
siːb	you leave	ṣiːb	you score	biːd	you kill/end	biːḍ	white (pl.)
tæːb	he repented	ṭɑːb	he recovered	bæːt	he slept	bɑːṭ	armpit
tuːb	you repent	ṭuːb	brick				

For the perception experiment, each target word was first extracted from the carrier sentence. Both waveforms and spectrograms were examined in order to determine the boundary between the carrier sentence and the target word. In the case of both word-initial stops and fricatives, the onset of the first consonant was extracted at the zero-crossing nearest the offset of F2 in the preceding vowel when it corresponded to the end of periodicity in the waveform. When periodicity continued in the waveform after the offset of F2, the zero-crossing nearest the offset of F1 was used. The word-final boundary was placed at the end of the visible burst for word-final stops and at the end of visible frication for word-final fricatives. Because these stimuli were followed by the closure and burst of a stop in the carrier sentence, the boundary was always clearly visible in the waveform.

After the target words were isolated as described above, word-initial stops and fricatives were extracted from the word onset to the onset of F1 in the following vowel. The segment was then removed at the nearest zero-crossing according to the waveform. Word-final stops and fricatives were also excised from the zero-crossing nearest the offset of F2 of the preceding vowel when it coincided with the end of periodicity in the waveform, or the offset of F1 when periodicity persisted, to the end of the word.

For each of the 9 word pairs, four versions were created using cross-splicing. Two stimuli were generated starting with the pair [sæb]–[ṣạḅ] (the subscript dot under /s/ indicates that this is an emphatic target consonant; the dots under the following vowel and consonant indicate that these segments were produced in the context of an emphatic consonant). One word was generated by splicing the [ạḅ] portion from [ṣạḅ] onto the [s] portion of [sæb], creating the hybrid [sạḅ]; a second word was formed by splicing the [æb] portion from [sæb] onto the [ṣ] portion of [ṣạḅ], creating the hybrid [ṣæb]. The two remaining stimuli were created by means of 'same-splicing', such that [sæb] and [ṣạḅ] were generated by replacing the [s] and [ṣ] portions, respectively, with those from different productions of the same words by the same speaker. To avoid differences in intensity between the two speakers, all stimuli were normalized following splicing such that the overall intensity of all vowels was 75 dB (SPL). The total number of stimuli was 72 (9 word pairs × 4 permutations × 2 speakers).

3.1.2. Participants

Thirty native speakers of Urban Jordanian Arabic (19 females) participated in the perception experiment. All were students at the University of Jordan and were paid for their participation. None reported any speech or hearing impairment.

3.1.3. Procedure

Participants were tested individually in a quiet room on the campus of the University of Jordan. The perception experiment was designed using Paradigm software (Perception Research Systems) and conducted using a Dell Inspiron laptop and Sony MDR-7506 headphones. Four repetitions of each stimulus were presented in random order. On each trial, orthographic representations of the two members of a minimal pair appeared on the computer screen in Arabic script. The location of each response alternative on the screen was randomized across trials. Participants heard a stimulus and indicated which word they heard (containing an emphatic or plain consonant) by clicking on the corresponding word on the screen. After a 2 s interval, the next trial started.

Participants received instructions in Arabic and started off with 10 practice trials to familiarize themselves with the nature of the stimuli and the presentation rate. Practice stimuli were not used in the main experiment. The perception experiment consisted of four repetitions of the 72 spliced stimuli, for a total of 288 trials.

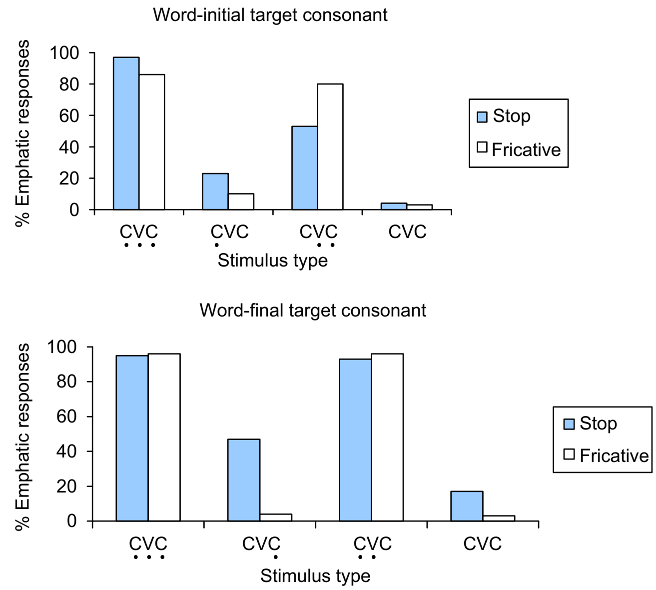

Fig. 3. Percentage of emphatic responses to cross-spliced CVC words with the target consonant (plain vs. emphatic) in **word-initial** (top panel) and **word-final** (bottom panel) position. For both word-initial and word-final consonants, C̣ṾC̣ represents an originally emphatic word and CVC represents an originally plain word. For word-initial consonants, C̣VC represents an originally plain word in which the initial plain consonant has been replaced with its emphatic counterpart; CṾC̣ represents an originally emphatic word in which the initial emphatic consonant has been replaced with its plain counterpart. For word-final consonants, CVC̣ represents an originally plain word in which the final plain consonant has been replaced with its emphatic counterpart; C̣ṾC represents an originally emphatic word in which the final emphatic consonant has been replaced with its plain counterpart.

3.2. Results: perception

3.2.1. Word-initial target consonants

The results are shown in Fig. 3. Percentage emphatic responses are shown for the four types of manipulation.[1] Focusing first on words with the target consonant in initial position, the splicing technique proved effective in that words in which all three segments came from an emphatic word received overwhelmingly emphatic responses (90%) while words in which all three segments came from a plain word received virtually no emphatic responses (3%). When there are conflicting cues to emphasis, the emphatic nature of the context is crucial. Words in which the VC portion came from an emphatic word received a majority of emphatic responses (69%) while words in which only the initial consonant came from an emphatic word received very few emphatic responses (15%). A one-way repeated measures ANOVA with Splice Type as the within-subjects variable revealed a main effect [$F(3,87)=990.31, p=.000$]. Bonferroni posthoc tests indicated that the percentage of emphatic responses to all four types of stimuli was significantly different ($p=.000$ for each comparison).

A two-way repeated measures ANOVA with word-initial Target and VC portion as independent variables revealed main effects of Target [$F(1, 29)=257.63$, $p=.000$] and VC portion [$F(1, 29)=1919.03$, $p=.000$]. Overall, words with emphatic target consonants received significantly more emphatic responses (53%) than words with plain consonants (36%); and words with emphatic ṾC̣ portions (80%) received significantly more emphatic responses than words with plain VC portions (9%). In addition, a significant Target × VC

[1] Percent correct scores were also converted to rationalized arcsine units (RAUs) prior to statistical analysis (Studebaker, 1985). Since analyses based on RAUs were identical to those based on percentages, results are presented in terms of percentages for ease of interpretation.

portion interaction [$F(1, 29)=31.35$, $p=.000$] indicated that an emphatic VC portion as compared to a plain VC portion led to a significantly greater increase in emphatic responses when added to an emphatic initial consonant (75% increase) as compared to a plain initial consonant (62% increase).

Finally, a two-way repeated measures ANOVA with Target and Manner as independent variables was conducted. A significant Target by Manner interaction [$F(1, 29)=101.48$, $p=.000$] showed that while words with initial emphatic (48%) or plain (41%) fricatives received a similar percentage of emphatic responses, words with emphatic stops (60%) received significantly more emphatic responses than words with plain stops (28%).

3.2.2. Word-final target consonants

When the target consonant was in final position, words in which all segments came from an emphatic word received the most emphatic responses (96%) while those with all plain segments received the fewest (10%). When there are conflicting cues to emphasis, the emphatic nature of the context is crucial. Words in which only the final consonant came from an emphatic word received few emphatic responses (25%) while those with the initial C̣Ṿ portion from an emphatic word received overwhelmingly emphatic responses (94%). A one-way repeated measures ANOVA with Splice Type as the within-subjects variable [$F(3, 87)=930.32$, $p=.000$] and subsequent Bonferroni posthoc tests indicated that the percentage of emphatic responses to all types of stimuli was significantly different ($p=.000$) except for the difference between words in which all segments came from an emphatic word and words in which the initial C̣Ṿ portion (but not the final consonant) came from an emphatic word.

A two-way repeated measures ANOVA with Target and CV portion as independent variables revealed main effects of Target [$F(1, 29)=79.79$, $p=.000$] and CV portion [$F(1, 29)=1245.08$, $p=.000$]. Overall, words with emphatic target consonants (61%) received significantly more emphatic responses than words with plain consonants (52%). Words with emphatic C̣Ṿ portions (95%) received significantly more emphatic responses than words with plain CV portions (18%). In addition, a significant Target $\times$ CV portion interaction [$F(1, 29)=56.20$, $p=.000$] indicated that an emphatic C̣Ṿ portion led to a significantly greater increase in emphatic responses when added to a plain final consonant (84% increase) as compared to an emphatic final consonant (70% increase).

Finally, a two-way repeated measures ANOVA with Target and Manner as independent variables was conducted. A significant Target by Manner interaction [$F(1, 29)=101.24$, $p=.000$] showed that while words with final emphatic (50%) or plain (49%) fricatives received a similar percentage of emphatic responses, words with emphatic stops (71%) received significantly more emphatic responses than words with plain stops (55%).

3.2.3. Spread of emphasis

As shown in Fig. 3, the data pattern the same for target consonants in word-initial and word-final positions, with one notable exception: words consisting of a plain target consonant and emphatic ṾC̣ or C̣Ṿ portion receive more emphatic responses when the target consonant occurs in word-final position (94%) as compared to word-initial position (66%). A two-way repeated measures ANOVA with Target Location (CṾC̣ vs. C̣ṾC) and Manner as independent variables was conducted. A main effect of Target Location indicated that this difference is significant [$F(1, 29)=161.61$, $p=.000$], with more emphatic responses when the target consonant occurs in word-final position. A significant main effect of Manner ([$F(1, 29)=44.89$, $p=.000$]) revealed that, overall, words with target fricatives received more emphatic responses (88%) than words with stops (73%). Finally, a significant Target Location by Manner interaction [$F(1, 29)=50.86$, $p=.000$] showed that the increase in emphatic responses from word-initial to word-final position was much greater for words with target stops (40% increase) than with target fricatives (16%). Independent *t*-tests confirmed that both of these increases were significant ([$t(118)=8.17$, $p=.000$] for stops; [$t(148)=4.95$, $p=.000$] for fricatives.

3.3. Summary of perceptual data

Results showed that words consisting entirely of emphatic parts (an emphatic target consonant plus a ṾC̣ or C̣Ṿ portion excised from an emphatic production) were perceived as emphatic 93% of the time. Likewise, words consisting entirely of plain parts were perceived as plain 94% of the time. For conflicting cue stimuli, consisting of emphatic and plain portions, the presence of an emphatic ṾC̣ or C̣Ṿ portion contributed much more to the perception of emphasis than the presence of an emphatic target consonant. Also, words with emphatic and plain fricatives received a similar percentage of emphatic judgments; in contrast, words with emphatic stops received significantly more emphatic judgments compared to words with plain stops. This pattern was observed for target consonants in both positions. Finally, the CV portion derived from a CVC̣ word received significantly more emphatic responses than the VC portion derived from a C̣VC word.

4. Discussion and conclusions

The goal of the present study was to provide a comprehensive account of the acoustics and perception of emphasis for a single dialect of Arabic. For the acoustic study, 12 speakers of Urban Jordanian Arabic (UJA) produced minimal pairs consisting of monosyllabic CVC stimuli with plain and emphatic consonants in either initial or final position. Both stops and fricatives as well as all 6 vowels (3 short and 3 long) of UJA were included.

Acoustic analyses focused on F1–F3 at the onset, middle, and offset of the vowel and on the spectral mean of the flanking consonants. The acoustic results showed that in the context of an emphatic consonant, F2 of the vowel is substantially lowered, replicating previous studies of several dialects (e.g., Card, 1983; Ghazeli, 1977; Khattab et al., 2006; Norlin, 1987; Zawaydeh, 1999). In addition, the present study clearly established that F1 and F3 are also both raised for vowels in emphatic context. These effects of emphasis on F1, F2, and F3 were observed throughout the vowel and when the target consonants were in both initial and final position.

A direct comparison of the degree of emphasis in short and long vowels showed no difference between the two, suggesting that previous reports of less emphasis in long vowels as measured at vowel midpoint (e.g., Yeou, 1997) simply reflected the greater distance from the target consonant rather than any intrinsic differences between short and long vowels.

For vowel quality, emphasis resulted in the greatest and most persistent lowering of F2 in the vowel /æ/. F2 lowering is less for the vowels /i/ and /u/, with a greater lowering for /i/ than /u/. A possible explanation for the difference between /i/ and /u/ is that it results from the contrast in the extent of formant transitions, with the greatest decrease for /i/, due to the vowel's inherently high F2 and the emphatic consonant's inherently low F2. In contrast, because of the inherently low F2 of the vowel /u/, the F2 transition between /u/ and the emphatic is less extensive. Consequently, there would be less lowering of F2 for /u/ as compared to /i/ as the present data show. The absolute F2 location of the vowel, however, does not explain the sizeable F2 lowering for /æ/, whose F2 is intermediate between those of /i/ and /u/. A possible explanation is that this patterning may be perceptually driven (cf. Flemming, 1995), resulting from the lack of contrast in backness for the low vowel. That is, while a substantial lowering of F2 for /i/ could endanger the contrast between /i/ and /u/, such a lowering for /æ/ will not cause any acoustic overlap or perceptual confusion because of the absence of a phonemic low back vowel. Interestingly, F2 lowering in /æ/ and /i/ diminished as distance from the target consonant increased; /u/ showed little to no change in F2 as a function of emphasis, suggesting both extent of transition and perceptual constraints may play a role in the gradient spread of emphasis.

The effect of emphasis was also observed in the target consonants themselves. Emphatic stops showed a substantially lower spectral mean than their plain counterparts. However, the spectral mean of fricatives was not affected by emphasis, consistent with Al-Khairy's (2005) findings for fricatives in Modern Standard Arabic. The present study clearly shows the differential effect of emphasis between stops and fricatives in both word-initial and word-final positions. It should be noted that while emphasis did not seem to affect fricatives, the adjacent vowels did show clear and significant effects of emphasis, although not as strong as those adjacent to stops. Differences between stops and fricatives may explain the sometimes conflicting pattern of results reported in the literature.

Finally, effects of emphasis were detected in the non-target consonants: non-target stops had a lower spectral mean when followed by emphatic target consonants. No such effects were observed in non-target fricatives.

Combining the present vocalic and consonantal measures, a clear picture of the nature and extent of emphasis spread emerges. Inclusion of measurements at multiple locations in the vowel as well as in the flanking consonants allows us to establish that emphasis not only affects the entire vowel, it also affects both the target and non-target stop consonants. Thus, emphasis is seen to spread from the target consonant throughout the entire CVC string. This spread occurs in both directions, rightward from word-initial consonants and leftward from word-final consonants. In this study, then, CVC is the domain of spread. Using these measures, future study with polysyllabic stimuli will provide a detailed account of spread beyond this domain (see Zawaydeh and de Jong (in press) for a report of spread across syllables based on F2 measures in the middle of the vowels).

Some accounts of emphasis also report that certain vowels and consonants block the spread of emphasis. Blocking (also known as 'opaque') segments seem to vary from dialect to dialect. For example, Davis (1995) describes two dialects of Palestinian Arabic, a southern one in which /i, j, ʃ, ʤ/ are blockers and a northern one in which /i, u, j, w, ʃ, ʤ/ block emphasis spread. Adding to this complexity is the observation that blocking segments may be opaque to emphasis spread in one direction but transparent to spread in the opposite direction. Based on impressionistic transcriptions of one speaker of Southern Palestinian Arabic, Davis (1995) reports that /i, j, ʃ, ʤ/ block only rightward emphasis spread while leftward emphasis spread has no blockers. Similarly, Herzallah (1990) reports that in her dialect of Northern Palestinian, rightward emphasis spread is blocked by /iː, j, ʃ/ while leftward emphasis has no blockers. In contrast to these data, Zawaydeh (1999) considered the vowels /i, u/ as potential blockers and found that neither vowel actually blocked the rightward spread of emphasis. For leftward emphasis spread, Zawaydeh (1999) considered the vowel /i/ as a potential blocker and concluded that it also did not block the leftward spread of emphasis. The present results do not provide any evidence for the existence of opaque vowels in UJA, thus confirming Zawaydeh's (1999) results. The present data do clearly show that the degree of emphasis is affected by vowel quality, with all 6 vowels showing significant differences due to the presence of an emphatic versus plain consonant. This difference is smallest and most restricted for the vowel /uː/. Under the view that emphasis involves a raising of the back of the tongue, the smallest effect for /u/ is consistent with the notion that high back vowels are less compatible with such a posterior secondary articulation. It should be noted, however, that blocking in its strongest form (i.e., absolutely no difference between vowels in plain and emphatic contexts) was not observed in the present study.

Consensus on the exact location of the secondary posterior constriction involved in emphasis is still lacking. Some researchers have described emphasis as velarization (e.g., Obrecht, 1968), others as uvularization (e.g., McCarthy, 1994; Zawaydeh, 1999), and still others as pharyngealization (e.g., Al-Ani, 1970; Davis, 1995; Younes, 1982) or pharyngealization and dorsalization (Herzallah, 1990). Acoustically, vocal tract modeling studies (e.g., Chiba & Kajiyama, 1946; Klatt & Stevens, 1969; Stevens, 1998) show that a low F2 results from a constriction for a considerable range of constriction locations, ranging from approximately 3–7 cm from the glottis. By including F1 and F3 measurements, the present acoustic results allow us to infer the location of the secondary constriction involved in emphasis in more detail. A high F1 also results from constriction locations at 3–7 cm from the glottis. Thus, a high F1 and low F2 are consistent with a constriction in the region engaged in the production of pharyngeal consonants. However, this still leaves the question about the exact location of the secondary posterior constriction for emphatics within this relatively large range. According to Klatt and Stevens (1969), the position of F3 can distinguish a more posterior from a more anterior constriction within this region. The present finding of a raised F3 is consistent with a more anterior constriction, similar to that for a uvular consonant. Taken together, our acoustic results suggest that emphasis involves a secondary constriction in the upper pharynx near the uvula. This is consistent with several articulatory studies (e.g., Ghazeli, 1977; Zawaydeh, 1999) and provides additional support for the call to replace the term 'emphasis' with 'uvularization' (e.g., McCarthy, 1994; Zawaydeh, 1999).

The perception of emphasis was investigated by using cross-spliced CVC syllables in which emphatic or plain consonants were spliced onto plain or emphatic syllable portions, respectively. Both stops and fricatives occurred as target consonants in both word-initial and word-final positions. Overall, words with emphatic target consonants received significantly more emphatic responses than words with plain consonants. Words with emphatic C̣Ṿ or ṾC̣ portions received significantly more emphatic responses than words with plain CV or VC portions. Moreover, the presence of an emphatic ṾC̣ or C̣Ṿ portion resulted in a dramatically greater percentage of emphatic responses than the presence of an emphatic consonant. There was, however, a marked difference depending on whether the target consonant was a stop or fricative. An emphatic stop elicited significantly more emphatic judgments than an emphatic fricative. In sum, the combination of a vowel and non-target consonant contributed more to the perception of the plain/emphatic distinction than the target consonant.

In general, the perceptual results closely match our acoustic findings. Acoustically, emphasis showed up clearly throughout the vowel. Perceptually, listeners based their judgment of emphasis largely on the presence of an emphatic ṾC̣ or C̣Ṿ portion and less on the emphatic consonant itself. The spectral mean of emphatic stops but not fricatives was significantly lower compared to their plain counterparts. In perception, emphatic stops were much more likely to be identified as emphatic than emphatic fricatives. Overall, the perceptual results for word-initial and word-final target consonants are similar and match our acoustic findings. Regarding spread, perceptually, in the context of a plain target consonant, the emphatic character of the initial C̣Ṿ portion affected listeners' emphatic judgments more than the emphatic character of the final ṾC̣ portion. Considering that the initial emphatic C̣Ṿ portion came from a CVC̣ word, this finding would be consistent with the observation that the leftward spread of emphasis is stronger than the rightward spread. Our acoustic results show a similar pattern for both directions of spread with differences observed in F1 which support a stronger effect of the final target consonant. F1 is consistently significantly lower at the onset of vowels preceding a word-final emphatic consonant while for word-initial emphatic consonants, there is no significant lowering in F1 at the offset of vowels. Thus, in terms of F1, the effect of the word-final emphatic consonant can be said to be stronger than that of the word-initial emphatic consonant, consistent with our perceptual data.

Altogether, the current study presents three findings that all indicate stronger anticipatory than carry-over coarticulation: vowel formant data (particularly F1), spread to initial non-target stops, and the greater number of emphatic judgments for CV portions derived from a CVC̣ word. Stronger anticipatory coarticulation may be due to the fact that leftward spread from a word-final target consonant to the beginning of the word would maximize perceptual cues available to the listener. The speaker would initiate the secondary articulation early in the word to signal the emphatic nature of the upcoming final consonant, thus facilitating lexical access. Our perceptual results support this view.

Finally, we also find a consistent difference between stops and fricatives in terms of emphasis. The present study is the first to report spectral means for emphatic stop consonants and to show that the spectral mean of emphatic stops is substantially lower than that of plain stops. In contrast, we find no difference in spectral means between emphatic and plain fricatives. This difference between stops and fricatives is consistent with the claim that laminal fricatives are more resistant to coarticulation than apical stops (e.g., Farnetani, 1997; Fowler, 1994; Keating, 1990; Recasens, 1999). The greater coarticulatory resistance of fricatives is hypothesized to stem from the fact that their production is governed by strict aerodynamic constraints which don't tolerate much variability in the shape of the tongue body behind the constriction. This is supported by Ghazeli's (1977) observation that /ṭ/ shows more tongue backing than /ṣ/ and /ð̣/. The greater tongue backing for emphatic stops as compared to emphatic fricatives means that the length of the front cavity increases much more in the production of emphatic stops and this is consistent with the present finding of a significant lowering of the spectral mean for emphatic stops but not for emphatic fricatives.

In conclusion, the present study examined acoustic measures of both consonants and vowels as well as perceptual measures to provide insight into the nature and extent of emphasis and its spread. Future studies will extend the scope of this research to polysyllabic words, providing a detailed account of emphasis.

Acknowledgements

Thanks to Bertrand Kotewall, Ariane Rhone, and Emma Frosini for their help and feedback. This research was partially supported by the National Science Foundation (# 0518969) and the University of Kansas General Research Fund.

References

Al-Ani, S. (1970). *Arabic phonology*. The Hague: Mouton.
Ali, L., & Daniloff, R. G. (1972). A contrastive cinefluorographic investigation of the articulation of emphatic–nonemphatic cognate consonants. *Studia Linguistica*, *26*, 81–105.
Ali, L., & Daniloff, R. G. (1974). The perception of coarticulated emphaticness. *Phonetica*, *29*, 225–231.
Alioua, A. (1995). *L'effet des consonnes d'arriere et des emphatiques sur la nature acoustique des voyelles longues de l'Arabe litteral moderne*. Unpublished doctoral dissertation, Universite Laval, Montreal, Canada.
Al-Khairy, M. A. (2005). *Acoustic characteristics of Arabic fricatives*. Unpublished doctoral dissertation, University of Florida.
Al-Masri, M., & Jongman, A. (2004). Acoustic correlates of emphasis in Jordanian Arabic: Preliminary results. In A. Agwuele, W. Warren, & S.-H. Park (Eds.), *Proceedings of the 2003 Texas Linguistics Society conference* (pp. 96–106). Somerville, MA: Cascadilla Proceedings Project.
Al-Tamimi, F., Alzoubi, F., & Tarawnah, R. (2009). A videoscopic study of the emphatic consonants in Jordanian Arabic. *Folia Phoniatrica et Logopaedica*, *61*, 247–253.
Alwan, A. (1986). *Acoustic and perceptual correlates of uvular and pharyngeal consonants*. Unpublished MA thesis, MIT.
Boersma, P., & Weenink, D. (2007). Praat: Doing phonetics by computer (Version 4.6.09) [Computer program]. Retrieved July 19, 2007, from 〈http://www.praat.org/〉.
Card, E. (1983). *A phonetic and phonological study of Arabic emphasis*. Unpublished doctoral dissertation, Cornell University.
Chiba, T., & Kajiyama, M. (1946). *The vowel: Its nature and structure*. Tokyo: Phonetic Society of Japan.
Davis, S. (1995). Emphasis spread in Arabic and grounded phonology. *Linguistic Inquiry*, *26*, 465–498.
Farnetani, E. (1997). Coarticulation and connected speech processes. In W. Hardcastle, & J. Laver (Eds.), *Handbook of the phonetic sciences* (pp. 371–404). Oxford: Blackwell.
Flemming, E. (1995). *Auditory representations in phonology*. Unpublished doctoral dissertation, UCLA.
Forrest, K., Weismer, G., Milenkovic, P., & Dougall, R. N. (1988). Statistical analysis of word initial voiceless obstruents: Preliminary data. *Journal of the Acoustical Society of America*, *84*, 115–124.
Fowler, C. A. (1994). Invariants, specifiers, cues: An investigation of locus equations as information for place of articulation. *Perception and Psychophysics*, *55*, 597–610.
Ghazeli, S. (1977). *Back consonants and backing coarticulation in Arabic*. Unpublished doctoral dissertation, University of Texas at Austin.
Giannini, A., & Pettorino, M. (1982). The emphatic consonants in Arabic. *Speech laboratory report* (Vol. 4). Istituto Universitario Orientale, Napoli.
Herzallah, R. (1990) *Aspects of Palestinian Arabic phonology: A non-linear approach*. Unpublished doctoral dissertation, Cornell University.
Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America*, *108*, 1252–1263.
Kahn, M. (1975). Arabic emphatics: The evidence for cultural determinants of phonetic sex-typing. *Phonetica*, *31*, 38–50.
Keating, P. (1990). The window model of coarticulation: Articulatory evidence. In M. Beckman, & J. Kingston (Eds.), *Papers in laboratory phonology, I: Between the grammar and physics of speech* (pp. 451–470). Cambridge: Cambridge University Press.
Khattab, G., Al-Tamimi, F., & Heselwood, B. (2006). Acoustic and auditory differences in the /t/–/ṭ/ opposition in male and female speakers of Jordanian Arabic. In S. Boudelaa (Ed.), *Perspectives on Arabic linguistics, XVI: Papers from the sixteenth annual symposium on Arabic linguistics*. Cambridge: John Benjamins.
Klatt, D. H., & Stevens, K. N. (1969). Pharyngeal consonants. *MIT Research Laboratory of Electronics quarterly progress report* (Vol. 93, pp. 208–216).
Laufer, A., & Baer, T. (1988). The emphatic and pharyngeal sounds in Hebrew and in Arabic. *Language and Speech*, *31*, 181–205.
McCarthy, J. (1994). The phonetics and phonology of Semitic pharyngeals. In P. Keating (Ed.), *Phonological structure and phonetic form: Papers in laboratory phonology III* (pp. 191–234). Cambridge: Cambridge University Press.
Norlin, K. (1987). A phonetic study of emphasis and vowels in Egyptian Arabic. *Working papers of the Lund University, Department of Linguistics*, *30*, 1–119.
Obrecht, D. (1968). *Effects of the second formant on the perception of velarization consonants in Arabic*. The Hague: Mouton.
Recasens, D. (1999). Lingual coarticulation. In W. J. Hardcastle, & N. Hewlett (Eds.), *Coarticulation: Theory, data, and techniques* (pp. 80–105). Cambridge: Cambridge University Press.
Stevens, K. N. (1998). *Acoustic phonetics*. Cambridge: MIT Press.
Studebaker, G. A. (1985). A "rationalized" arcsine transform. *Journal of Speech and Hearing Research*, *28*, 455–462.
Wahba, K. M. S. (1993). *A sociolinguistic treatment of the feature of emphasis in Egypt.* Unpublished doctoral dissertation, University of Texas at Austin.
Watson, J. C. E. (1999). The directionality of emphasis spread in Arabic. *Linguistic Inquiry*, *30*, 289–300.
Yeou, M. (1995) Trading relations between cues for the pharyngealized/non-pharyngealized contrast. In *Proceedings of the 13th international congress of phonetic sciences* (pp. 464–467), Stockholm.
Yeou, M. (1997). Locus equations and the degree of coarticulation of Arabic consonants. *Phonetica*, *54*, 187–202.
Younes, M. (1982). *Problems in the segmental phonology of Palestinian Arabic.* Unpublished doctoral dissertation, University of Texas at Austin.
Younes, M. (1993). Emphasis spread in two Arabic dialects. In M. Eid, & C. Holes (Eds.), *Perspectives on Arabic linguistics V* (pp. 119–148). Amsterdam & Philadelphia: John Benjamins.
Zawaydeh, B. A. (1999). *The phonetics and phonology of gutturals in Arabic.* Unpublished doctoral dissertation, Indiana University.
Zawaydeh, B. A., & de Jong, K. J. (in press). The phonetics of localising uvularization in Arabic. In B. Heselwood, & Z. Hassan (Eds.), *Instrumental studies in Arabic phonetics*. John Benjamins.
Zuraiq, W., & Zhang, J. (2006). Phonological assimilation in Urban Jordanian Arabic. *Kansas Working Papers in Linguistics*, *28*, 33–64.