Al-Tamimi J, Khattab G.

Acoustic cue weighting in the singleton vs geminate contrast in Lebanese Arabic: The case of fricative consonants.

Journal of the Acoustical Society of America **2015, 138(1), 344-360.**

DOI link to article:

http://dx.doi.org/10.1121/1.4922514

Date deposited:

17/07/2015

Acoustic cue weighting in the singleton vs geminate contrast in Lebanese Arabic: The case of fricative consonants

Jalal Al-Tamimi[a] and Ghada Khattab
Speech and Language Sciences, Newcastle University, Newcastle Upon Tyne, NE1 7RU, United Kingdom

(Received 30 November 2012; revised 24 October 2014; accepted 28 May 2015; published online 16 July 2015)

This paper is the first reported investigation of the role of non-temporal acoustic cues in the singleton-geminate contrast in Lebanese Arabic, alongside the more frequently reported temporal cues. The aim is to explore the extent to which singleton and geminate consonants show qualitative differences in a language where phonological length is prominent and where moraic structure governs segment timing and syllable weight. Twenty speakers (ten male, ten female) were recorded producing trochaic disyllables with medial singleton and geminate fricatives preceded by phonologically short and long vowels. The following acoustic measures were applied on the medial fricative and surrounding vowels: absolute duration; intensity; fundamental frequency; spectral peak and shape, dynamic amplitude, and voicing patterns of medial fricatives; and vowel quality and voice quality correlates of surrounding vowels. Discriminant analysis and receiver operating characteristics (ROC) curves were used to assess each acoustic cue's contribution to the singleton-geminate contrast. Classification rates of 89% and ROC curves with an area under the curve rate of 96% confirmed the major role played by temporal cues, with non-temporal cues contributing to the contrast but to a much lesser extent. These results confirm that the underlying contrast for gemination in Arabic is temporal, but highlight [+tense] (fortis) as a secondary feature.
© *2015 Acoustical Society of America*. [http://dx.doi.org/10.1121/1.4922514]

[JFL] Pages: 344–360

I. INTRODUCTION

The phonetic and phonological aspects of gemination have been the subject of investigation in various languages, and different approaches to the representation and implementation of the singleton vs geminate contrast have been proposed. From a phonological point of view, gemination typically refers to a consonantal length contrast which can be employed for lexical, morphological, and/or pragmatic purposes (e.g., Broselow *et al.*, 1997; Cohn, 2003; Davis, 2011). From a phonetic point of view, the contrast has a variety of temporal and non-temporal manifestations which vary in their magnitude and domain depending on the language in question.

A. Acoustic and articulatory characteristics of gemination

The majority of studies have shown that consonant duration is a major acoustic cue to the singleton vs geminate contrast [e.g., Al-Tamimi and Khattab (2011) and Khattab and Al-Tamimi (2014) on Lebanese Arabic; Arvaniti and Tserdanelis (2000) and Tserdanelis and Arvaniti (2001) on Cypriot Greek; Esposito and di Benedetto (1999) on Italian; Ham (2001) on Bernese, Levantine Arabic, Hungarian, and Madurese; Hansen (2004) on Persian; Hassan (2003) on Iraqi Arabic; Idemaru and Guion (2008) on Japanese; Lahiri and Hankamer (1988) on Turkish; Ridouane (2007) on Berber; among others].

Studies of the preceding vowel's duration have yielded conflicting results. Some studies show the vowel preceding the geminate consonant to be shorter (Esposito and di Benedetto, 1999; Ham, 2001; Ridouane, 2007), while others have found the reverse pattern (Hansen, 2004; Hassan, 2003; Idemaru and Guion, 2008; Lahiri and Hankamer, 1988; Tserdanelis and Arvaniti, 2001). Various explanations relating to language-specific rules for weight, stress patterns, and syllable structure have been proposed in these studies.

In addition to quantity differences between singleton and geminate consonants, researchers have found qualitative differences in the articulation of each category. For example, in articulatory work geminate stops and fricatives have been shown to involve more contact than singletons in electropalatography traces [Payne (2006) on Italian], while singletons are often lenited/fricated (Ridouane, 2007). In the case of fricatives, the increased area of contact in geminates creates a narrower constriction, which is presumed to increase the noise frequency due to higher air pressure (Payne, 2006).

Acoustically, geminate stops have been observed to have higher burst amplitude, a higher number of bursts and stronger bursts [Abramson (1999) on Pattani Malay; McKay (1980) on Rembarrnga; Ridouane (2007)]. The syllable in a post-geminate position has been shown to have higher intensity, root mean square amplitude and $f0$ than in post-singletons (Abramson, 1999; Idemaru and Guion, 2008; Ridouane, 2007). Lateral geminates have been shown to have a more palatalized configuration with lower F_1 and higher F_2 and F_3 [Local and Simpson (1999) on Malayalam; Payne (2006)]. Voice quality differences have also been associated with the singleton vs geminate contrast, though

---

[a] Electronic mail: Jalal.Al-Tamimi@newcastle.ac.uk

the picture here is mixed. Idemaru and Guion (2008) and Local and Simpson (1988) suggest that geminate consonants are associated with phonatory tightness, creak, and tense articulation, while Arvaniti and Tserdanelis (2000) suggest that this only applies to fricative geminates, whereas stops and affricates are associated with breathy voice. Preceding vowels have been shown to present more fronted and/or closer articulation (higher F_2 and/or lower F_1) in geminate laterals (Local and Simpson, 1988, 1999), while no differences in the formant structure of the preceding vowel have been found between singleton and geminate Italian stops (Esposito and di Benedetto, 1999) or stops, fricatives, nasals, and liquids in Cypriot Greek (Arvaniti and Tserdanelis, 2000).

The prevalence of the non-temporal manifestation reported above has led some researchers to suggest that the singleton vs geminate contrast in some languages is fundamentally based on a lenis/fortis or lax/tense contrast, with secondary [e.g., Kohler (1984); McKay (1980); Nellis and Hollenbach (1980) on Cajonos Zapotec] and sometimes primary temporal consequences [e.g., DiCanio (2012) on Itunyoso Trique]. This is not surprising given that long duration and articulatory strength often go hand in hand. For instance, long segments require articulatory strength to maintain the constriction, since the act of moving and holding the articulators for longer requires higher articulatory effort (Catford, 1977). On the other hand, fortis consonants are produced with higher pulmonic strength, leading to high pressure behind the place of articulation and greater time spent in an extreme articulation (Jaeger, 1983; Jakobson *et al.*, 1976); this in turn leads to longer duration, which contrasts with the shorter duration of lenis consonants, alongside reduced voicing and many of the non-temporal manifestations reported above. Jessen (2001) provides an account of how phonetic lengthening as a secondary consequence of tense articulations can be phonologized in some languages like Swiss German and become the primary cue for the contrast in question. DiCanio (2012) further shows how a fortis/lenis contrast in Zapotec languages can be primarily based on duration and glottal width, which varies depending on stress position. In relation to the role of glottal states, Jessen (2001) and Nellis and Hollenbach (1980) note a correlation between a contrast based on tenseness or fortis articulation and languages which exhibit a certain profile of glottal timing in their stop contrast, e.g., the presence of aspirated stops, the lack of voiced geminates, and/or compensatory shortening/lengthening of preceding vowels. It is therefore important to consider language-specific prosodic constraints which govern phonetic and phonological timing, syllable structure, and (non-)contrastive vowel length (see, e.g., Ham, 2001, pp. 6–14).

B. Gemination in Lebanese Arabic

Vowel and consonant length play a major role in Arabic phonology and morphology [e.g., /ˈkatab/ "*(he) wrote*" vs /ˈkatːab/ "*(he) made someone write*"; /ˈdaːm/ "*(he) lasted (verb)*" vs /ˈdamː/ "*blood (noun)*," Ham (2001); Nasr (1960)]. All consonants in Lebanese Arabic (LA) can be geminated and vowel length is also contrastive. Word medial fricatives in LA can occur in different trochaic syllable structures with short and long vowels preceding singleton and geminate consonants:

ˈCVCVC ⇒ /ˈʔasˤam/ (*he divided*),
ˈCVC:VC ⇒ /ˈʔasˤːam/ (*he partitioned*),
ˈCV:CVC ⇒ /ˈʔaːsˤam/ (*he shared*),
ˈCV:C:V(C)⇒ /ˈʔaːsˤːa/ (*having cut (fem.)*).

Medial geminate consonants are also found in iambic structures (e.g., /baˈsˤːaːr/, "*fortune teller*") but these are not examined here. Post-lexical geminates are also common, arising from assimilation of the definite article /ʔal/ with a following coronal consonant, e.g., Standard Arabic /ʔal/+/suuq/ "*the market*" > /ʔassuuq/ > [ssuuʔ] in LA. Studies on gemination in Arabic are relatively scarce and tend to include very few subjects (e.g., Ham, 2001; Hassan, 2003; Nasr, 1960). In these and our own previous studies (e.g., Al-Tamimi and Khattab, 2011; Khattab, 2007; Khattab and Al-Tamimi, 2014) durational differences have been reported to significantly distinguish between singleton and geminate consonants in Arabic, however, we are not aware of any study that has researched non-temporal patterns in the implementation of this contrast in Arabic and their potential contribution to the acoustic basis of the contrast.

This is interesting given that the Arabic term for gemination, /taʃdiːd/, literally means "strengthening," "intensification," or "reinforcement." There are various motivations for an investigation of this kind: first, a qualitative distinction would parallel relatively recent findings on qualitative differences in contrastive vowel length in Arabic (Alghamdi, 1998; Al-Tamimi, 2007); these have only been particularly noted since experimental work on Arabic vowels started to emerge, with previous small-scale studies suggesting that the contrast is purely durational (e.g., Al-Ani, 1970, among others). Second, while the phonology of Arabic is heavily oriented towards phonological contrasts in vowels and consonants that are based on length and moraic timing (e.g., Broselow *et al.*, 1997; Davis, 2011; Watson, 2007), it is important to examine the phonetic basis of phonological length in order to test whether articulatory strength still plays a role in a contrast that is heavily based on phonetic timing. This would highlight the correlation between the two and enable the study of perceptual cues that might enhance this contrast. Third, and in relation to this last point, an exploration of primary and secondary cues in the implementation of gemination can help interpret developmental patterns in the acquisition of gemination in Arabic, where children might initially latch on to a secondary cue and use it instead of a primary one in their production (Khattab and Al-Tamimi, 2013).

C. Acoustic characteristics of fricatives

Acoustic characteristics of fricatives have been described in various studies, and most of the research has attempted to classify fricatives in terms of place of articulation and/or voicing differences (for a comprehensive review of the literature, see Maniwa *et al.*, 2009; Shadle, 2012, among others). Several acoustic cues for distinguishing fricatives have been investigated, including the peak frequency, spectral moments, formant transitions, overall and dynamic amplitude, duration, to name a few (Forrest *et al.*, 1988; Jesus and Shadle, 2002; Jongman *et al.*, 2000; Li *et al.*, 2009; Maniwa *et al.*, 2009; Shadle, 2012). The peak location and/or spectral moments are mostly used to describe fricatives and the literature suggests that the peak frequency, the centroid (M1), and the skewness (L3) are negatively correlated with the length of the front resonating cavity; with more front articulations showing higher centroid (M1) and peak frequencies and positive skewness (L3). The standard deviation (M2) and the kurtosis (L4) can distinguish flat-diffuse from peaked-compact spectra with higher values for the former, and can also distinguish the tongue posture between apical and laminal areas with the former posture showing more peaked spectrum with lower standard deviation (M2) and higher kurtosis (L4) as between Swedish and American /t/ (Forrest *et al.*, 1988; Jongman *et al.*, 2000; Li *et al.*, 2009; Maniwa *et al.*, 2009; Shadle, 2012, among others). In investigating acoustic characteristics of fricatives in clear vs conversation speech (or high vs low effort levels, respectively), some studies report that fricatives produced in clear speech have longer duration, higher peak location and centroid and higher F_2 transitions (Maniwa *et al.*, 2009), and higher dynamic amplitude (A_d) reflecting higher effort levels associated with clear speech (Jesus and Shadle, 2002). We chose to borrow these measures for the investigation of the singleton-geminate contrast in fricatives given their potential in detecting place of articulation differences and/or fortis articulation in geminates.

II. METHOD

A. Speakers and data recording

Twenty Lebanese speakers (ten male, ten female) with no reported history of speech disorders and aged between 18 and 40 were recruited from Beirut. All speakers were university-educated and were born and raised in Lebanon. Half of the speakers lived in Beirut for the majority of their life and the remaining speakers studied and lived there for at least 2 years. No other criteria were used to control for their dialectal background. They were all familiar with Standard Arabic through education, and they were all exposed to English and French due to the multilingual nature of Lebanon. The speakers were audio-recorded while reading a word-list with randomized target short and long vowels preceding singleton and geminate medial fricatives in four trochaic disyllabic structures: ˈCVCVC, ˈCVːCVC, ˈCVCːVC, and ˈCVːCːVC (see examples in Sec. I B). These structures represent the four-way durational contrast that can occur in LA, whereby both long and short vowels can precede singleton and geminate consonants. While the first three structures are very common in LA, the fourth (with phonological length in both the medial fricative and preceding vowel) is relatively rare and in fact a small number of target words with the ˈCVːCːV were rejected by some of our participants despite surviving our piloting phase.

The corpus consisted of production of all singleton and geminate consonants (C/Cː hereafter) in LA, including stops, fricatives, nasals, liquids, and approximants. The total number of words produced by the speakers was 5171. A subset of the corpus dealing with fricative consonants is presented here, and the remaining results are presented elsewhere (e.g., Al-Tamimi and Khattab, 2011; Khattab and Al-Tamimi, 2014). The decision to focus on fricatives was made because they constitute the largest category in the LA consonant inventory (10 out of the 24 consonants, covering most places of articulation. Moreover, the non-durational acoustic cues that are relevant for analysis in fricatives are quite different from those relevant for other manners of articulation.

Near minimal-pair sets were used with the medial C/Cː being one of all possible fricatives in LA: /f s sˤ z ʃ ʒ x ɣ ħ h/, and target vowels preceding and following the medial C/Cː were either /a/ or /aː/, though /aː/ was frequently realized as [eː] or [ɛː] due to Imāla (a process that involves raising long /aː/; Nasr, 1960). Up to three words per consonant and syllable structure were selected as representative of all possible words containing the target singleton vs geminate environments and fricative phonemes (see Table IV in the Appendix). Due to the large number of words in the total corpus (in which we looked at all manners of articulation), no repetitions were recorded and no carrier sentence was used in order for the task not to be too tedious for the participants, who were recorded for a total of an hour each. Instead, the tokens were randomized and fillers were added before presentation in order to distract the attention of the speakers from the aim of the study. While the tokens were presented to the participants in the written form using Arabic script, which in Arabic can run the risk of eliciting a Standard Arabic pronunciation, many of the lexical items were specific to the Lebanese variety and the subjects were instructed to produce all the words as if they were speaking in their own variety in an informal style. This method worked well for all but one subject, who could not refrain from switching to the Standard Arabic variety on seeing the written script; this subject was subsequently replaced. For the remaining subjects, we obtained speech material that is representative of the LA variety. Recordings were made in a quiet room, using an R9 solid-state recorder with a SONY MS957 Uni-directional Stereo Electret Condenser microphone (frequency response 50–18000 Hz), and digitized at 44.1 kHz, in mono channel and 16-bit quantization.

B. Data processing and acoustic analyses

1. Data segmentation

Due to technical errors or words being rejected by speakers, 1726 different words out of 1880 target words were produced by all the speakers. Acoustic and auditory analyses were done using PRAAT (Boersma and Weenink, 2009). The data were labeled semi-automatically using the package STK (Farinas *et al.*, 2005); whereby several C and V

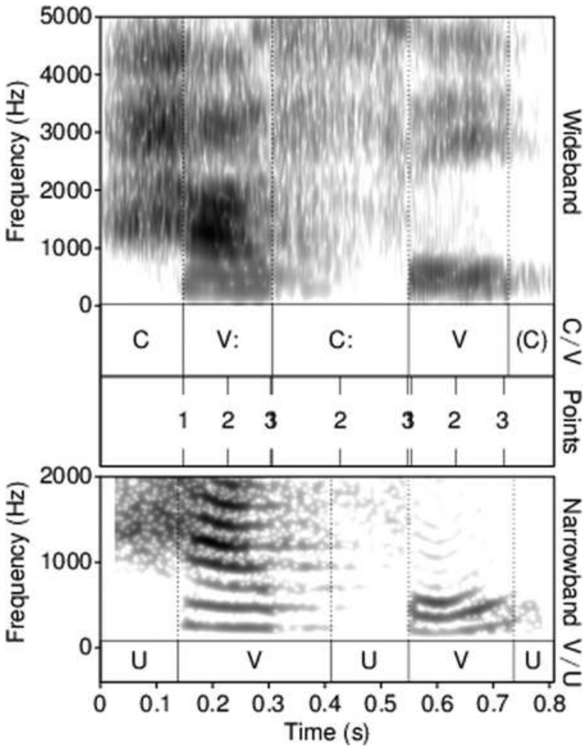


FIG. 1. Wideband spectrogram of the word /ħaːsːe/ “*I feel*” produced by a female speaker, and segmentation into C and V (consonant and vowel, respectively, see tier C/V), points of measurements (tier points), followed by a narrowband spectrogram and boundaries for the voiced and unvoiced frames (V and U, respectively, see tier V/U).

intervals were automatically added based on fundamental frequency *f*0 and intensity computation. These intervals were then transferred into TextGrids and manually inspected to correct for potential errors in boundary positioning, using the following segmentation criteria (see Fig. 1).

(a) The starting point of vowels was determined in accordance with the rise in amplitude from the previous consonant and appearance of a homogeneous formant structure, and the end point in accordance with a drop in amplitude and disappearance of or abrupt change in formants. Any voiceless portions after the following vowel [see (C) on tier “C/V,” Fig. 1] were not included in the analysis as this portion seems not to contribute to the perception of vowel duration (see Nakai *et al.*, 2009).

(b) Boundaries of medial fricatives were determined according to the onset and offset of visible and/or audible friction, including any period of silence which sometimes preceded/followed the fricative.

2. Acoustic analyses

A PRAAT script was designed by the first author to perform all the acoustic measurements. To obtain accurate measurements at the different positions of a vowel or a fricative, measurement frame positions were estimated (following Al-Tamimi, 2007). The acoustic periodicity of voiced frames was first estimated through a PointProcess (cross-correlation) analysis following an *f*0 estimation (see below for more details on *f*0 estimation). Then for each speaker, the average length of a complete glottal cycle was computed, which ranged over 8–10 ms for male and 4–6 ms for female speakers. The initial estimates of measurement times were obtained from the TextGrids following the segmentation as described above. They were then adjusted to match the time of maximum intensity occurring within the length of an average glottal cycle, left-aligned to the original onset estimate, right-aligned to the original offset estimate, and centered at the original midpoint estimate. The intensity values, computed every 5 ms, were interpolated before computing the maximum; the adjustments that resulted from this process were up to 2–3 ms around the original positions (see positions 1, 2, and 3 for onset, midpoint and offset in Fig. 1, tier “points”). All the reported measurements are obtained at the estimated positions. The following acoustic measurements were taken from the data.

a. Absolute duration. Obtained from the start to the end point of each fricative and vowel (Fig. 1, tier 3).

b. Intensity. Obtained at the onset, midpoint and offset of each fricative and vowel using PRAAT’s default settings, with a time step of 5 ms and interpolated.

c. Fundamental frequency. Obtained at the onset, midpoint and offset of each fricative and vowel using the two-pass method to accurately estimate *f*0 for each speaker (Hirst, 2011). PRAAT’s default settings were used for the first pass (5 ms step, 40 ms Gaussian window), while for the second pass, the estimated pitch ceiling and floor were adapted to each speaker, by obtaining the first and third quartiles and multiplying each by a coefficient; 0.75 and 1.5, respectively, with a 5 ms step and an effective Gaussian window length of 30 and 20 ms for male and female speakers, respectively. The new estimated values were in the range of 100–300 Hz for a male and of 150–400 Hz for a female speaker. *f*0 curves were not smoothed to prevent overestimation of *f*0 tracks; therefore, only true voiced frames were measured yielding fewer analyzed tokens depending on the position (see Sec. III C). This measure was taken to explore potential differences in *f*0 frequencies in both the medial fricatives and the surrounding vowels that are due to gemination.

d. Spectral moments in medial fricatives. The four spectral moments [i.e., centroid (M1), standard deviation (M2), skewness (L3), and kurtosis (L4)], the peak location, and the dynamic amplitude (A_d) were used to evaluate potential increase/decrease in these coefficients which may be linked to differences in place of articulation between singleton and geminate consonants, or differences in effort level. To the best of our knowledge, this is the first time spectral moments have been used to investigate potential differences between singleton and geminate fricative consonants. To correctly estimate the four spectral moments, the peak, and the dynamic amplitude, time-averaged power spectra were used. The sound files were low-pass filtered with an anti-aliasing filter with cutoff frequency of 18 kHz (the frequency response of the microphone), down sampled to 36 kHz and pre-emphasized by a factor of 0.98 (following Forrest *et al.*, 1988; Jongman *et al.*, 2000; Maniwa *et al.*, 2009, among others, to allow for direct comparison of results). Time-averaging of the power spectra was used over ensemble-averaging as the target words were repeated once (see Jesus and Shadle, 2002; Shadle, 2012). We considered the long domain time-averaged power spectra to account for dynamic properties of each fricative obtained at the most stable region of the fricative excluding the transitions; 80% of the total duration of the fricative was used for the computations (duration$_{steady\text{–}state}$ hereafter). A minimum duration of 63 ms was considered optimal to estimate the power spectra to allow for up to 50% overlap between the windows; imposing this minimum duration meant that only 49 fricatives out of 1726 were excluded due to having a short duration. Nine 10 ms Kaiser-2 windowed intervals were evenly spaced within the duration$_{steady\text{–}state}$ of the fricative, with a maximum of 50% overlap between the windows for durations of fricatives from 63 ms up to no overlap for durations of 120 ms and above. The first and last windows were left- and right-aligned to the edges of the steady-state region of the fricative, one window was centered in the middle of the fricative and the remaining six windows were evenly spaced from the midpoint. For each windowed interval, a 256-point zero-padded discrete Fourier transform (DFT) spectrum was computed. The complex valued spectrum was averaged prior to log transforming the results in dB. In each frequency bin, the real and imaginary parts were squared and summed to form a magnitude-squared value, then the magnitude-squared values at each frequency bin across the nine spectra were averaged to form a single spectrum. Then the log values of each magnitude-squared value are found to form the log of the averaged power spectrum in dB. The four spectral moments of the averaged power spectrum were obtained using PRAAT's default settings, with the centroid (M1) representing the first spectral moment of the averaged power spectrum, the standard deviation (M2) representing the square-root of the second centralized moment of the averaged power spectrum; the skewness (L3) and kurtosis (L4) represent the normalized third and fourth centralized moments of the averaged power spectrum; these were normalized by the second central moment (the third centralized moment was divided by 1.5 power of the second central moment, and the fourth was divided by the square of the second central moment minus 3) to enable direct comparison of these dimensionless moments (see Forrest *et al.*, 1988; Jongman *et al.*, 2000; Maniwa *et al.*, 2009, among others). The peak was considered to be the peak with the highest amplitude in the whole frequency distribution between 0.5 and 18 kHz (with 0.5 kHz being used to exclude effects of $f0$ and/or harmonics). Dynamic amplitude (A_d) was then estimated to evaluate potential differences in effort level that may be associated with the singleton *vs* geminate contrast. Dynamic amplitude (A_d) represents the difference between the maximum amplitude occurring between 0.5–18 kHz and the minimum amplitude occurring between 0 and 2 kHz (Jesus and Shadle, 2002).

e. Voicing patterns in medial fricatives. To quantify voicing patterns in fricatives, phonologically voiced and voiceless fricatives were separated in order to investigate the degree of devoicing for voiced fricatives or voicing shadows for voiceless fricatives. The sound file was first low-pass filtered at 500 Hz and $f0$ estimation was used (as described above). Then PRAAT's VUV function was used with an average duration of a complete glottal cycle adapted to each speaker (see above) and with a minimum of 20 ms for continuous voiced or unvoiced intervals. Automatic estimation of the voiced/unvoiced frames was then manually checked for potential errors, by investigating both a narrowband spectrogram and $f0$ tracks (see Fig. 1, narrowband spectrogram and tier "V/U"). Errors constituted less than 5% of the data. Then the percentage of voicing in each fricative was computed.

f. Formant frequencies of surrounding vowels. Formant frequencies (F_1, F_2, and F_3) of surrounding vowels were used to evaluate potential qualitative differences linked with the singleton vs geminate environments. These were obtained from a 25 ms Gaussian window with a 5 ms time step and interpolation. A maximum of five formants were requested in the formant analysis using the default Burg algorithm for formant estimation with a maximum frequency of 5 kHz for male and 5.5 kHz for female speakers. Then PRAAT's formant track function was used in order to limit errors in automatic formant estimation. Formant frequencies were obtained at the midpoint and offset of the preceding vowel and at the onset and midpoint of the following vowel. Formant frequencies were then verified manually to prevent potential errors obtained from automatic extraction (errors constituted less than 5% of the data).

g. Voice quality correlates of surrounding vowels. The following four voice quality measures were used: $*H_1$-$*H_2$, $*H_1$–A_1, $*H_1$–A_2, and $*H_1$-$*A_3$, with the asterisk reflecting normalized measures to correct for the boosting effect of formants on these harmonics (following Iseli *et al.*, 2007). These four acoustic measures were shown to reflect differences in voice quality (Arvaniti and Tserdanelis, 2000; Iseli *et al.*, 2007; Idemaru and Guion, 2008). The sound file was low-pass filtered with an anti-aliasing filter with cut-off frequency of 5 kHz for male and 5.5 kHz for female, downsampled to 10 kHz for male and 11 kHz for female speakers, and pre-emphasized by a factor of 0.98. Intervals 40 ms long were defined, left-aligned at the offset of the preceding vowel and right-aligned at the onset of the following vowel, and windowed using Kaiser-2 window function. The DFT was computed for each 40 ms windowed signal and the logarithmic power spectral density, with a bin size of 11 Hz, was computed. Then amplitudes of the first and second harmonics and of the first to the third formants were automatically obtained by detecting the highest peaks for a particular harmonic. Starting with H_1 and H_2, the maximum amplitude was obtained in the region from $f0*0.9$ to $f0*1.1$ for the former and in the region from $2*f0*0.95$ to $2*f0*1.05$ for the latter. For the amplitude of the harmonics closest to the three first formants, the maximum amplitude was obtained in the region from $F_1 - 0.5*\text{Bandwidth}_1$ to $F_1 + 0.5*\text{Bandwidth}_1$ for A_1, from $F_2 - 0.5*\text{Bandwidth}_2$ to $F_2 + 0.5*\text{Bandwidth}_2$ for A_2, and from $F_3 - 0.5*\text{Bandwidth}_3$ to $F_3 + 0.5*\text{Bandwidth}_3$ for A_3. The automatic detection of the highest peak for H_1, H_2, A_1, A_2, and A_3 was manually checked to prevent errors (errors constituted less than 5% of the data).

C. Statistical analyses

1. Z-score transformation

In order to reliably compare acoustic measurements varying in scales in the extent to which they play a role in the singleton vs geminate contrast, all acoustic measurements were Z-scored. To obtain Z-scores for a particular acoustic measure, for example, the duration of a particular fricative, we started by subtracting the duration of that fricative from the average duration and the result was then divided by the standard deviation of duration (average and standard deviation were obtained from all fricatives and vowels in the four syllabic shapes, for each speaker). 99.7% of Z-scores range between +3 and −3 and can be interpreted by evaluating the distance of a Z-score from zero; for example, a Z-score of −0.59 (see Z-score duration of the medial fricative in CVCVC, Fig. 2) indicates a low absolute duration compared to all other segments. Then we compare this Z-score with that of a comparable geminate environment, e.g., medial fricative in CVC:VC; +1.20. The difference between the singleton and geminate categories is then evaluated as the percentage difference between the left-tailed probability percentile of each Z-score, which yields a percentage of +217.3% rise in Z-score duration in the geminate environment compared to the singleton. When the probability percentile difference is close to 0%, no significant differences are obtained. Graphical results are based on Z-scores and reference to raw data will be made to evaluate absolute differences.

2. Analysis of variance and effect size measures

To examine potential differences between the singleton and geminate environments, several three-way UNIVARIATE analyses were applied on all the measurements. Effects of the word type as a random factor were evaluated and results were non-significant from low to high level interactions, suggesting a homogeneous set of realizations. Therefore we averaged all the productions from a particular speaker, phoneme and syllable structure into one value in order to reduce the error in the final statistical model. The data were separated by preceding vowel, fricative and following vowel, and statistical analyses were applied where appropriate. All statistical analyses were applied using the UNIVARIATE Generalized Linear Model in SPSS 19. Three independent variables were included as fixed factors, as we are using all possible combinations of consonant length, vowel length and fricative phonemes in LA: Consonant length (two levels, singleton vs geminate), vowel length (two levels, short vs long vowel preceding the fricative), and phonemes (ten fricative consonants). The dependent variables in each three-way UNIVARIATE analyses were each acoustic measurement [statistical results reported are significant at $p \leq 0.023$ after the false discovery rate (FDR) alpha correction]. T-tests were used to evaluate the differences observed in consonant length in the two short and long vowel environments, as this interaction is supposed to show more differences (results reported are significant at $p \leq 0.02$ after FDR alpha correction). And finally, two effect size measures were used to evaluate the real contribution of each factor to the singleton vs geminate contrast: the unbiased measure of strength of association, omega-squared ω^2 for the UNIVARIATE analyses and Cohen *d* for the t-test (Cohen, 1988, Chaps. 2 and 8). Cohen benchmarks are used to evaluate the size of the effect: ω^2 is "large" when > 0.14, "moderate" when >0.06 and <0.14, and "small" when <0.01 and Cohen *d* is large when >0.8, moderate when >0.5 and <0.8, and small when <0.2 (henceforth "L," "M," and "S" are used for large, moderate, and small).

3. Discriminant analysis and ROC curves

To evaluate the robustness of the observed differences between the singleton and geminate categories, all the acoustic measures were submitted to several linear discriminant function analyses. As grouping variables, we used consonant length in the two short and long vowel environments separately. Independent variables were each acoustic measurement. The training and testing sets were determined automatically by SPSS 19, by using the *leave-one-out* cross-validation method, wherein each case is classified by the functions derived from all the cases other than that case. Once the classification was obtained, we recorded the probability score obtained for each of the singleton and geminate cases. Then all classification models were evaluated based on the receiver operating characteristics (ROC) curves (see, e.g., Swets *et al.*, 2000). We used the probability scores of the geminate category obtained from the discriminant analyses to obtain each ROC curve. The shape of the curve (a bowed curve raising from the 45 degree line to the upper left corner), and the percentage score of the area under the curve (AUC), will show which classification model is better at distinguishing between the singleton and the geminate categories. An AUC close to 100% indicates a highly accurate model.

D. Expectations

We expect to obtain high classification and AUC rates from highly significant differences with high effect sizes. If the singleton vs geminate contrast in LA is based on temporal differences with subsequent secondary effects of non-temporal acoustic cues, we expect to obtain higher classification and AUC rates when duration is used. If on the other hand, the singleton vs geminate contrast in LA is based on the tense vs lax (fortis vs lenis) distinction with secondary consequences of temporal differences, we expect to obtain higher classification and AUC rates with non-temporal acoustic cues. Based on findings from the literature, geminate environments were expected to involve the following:

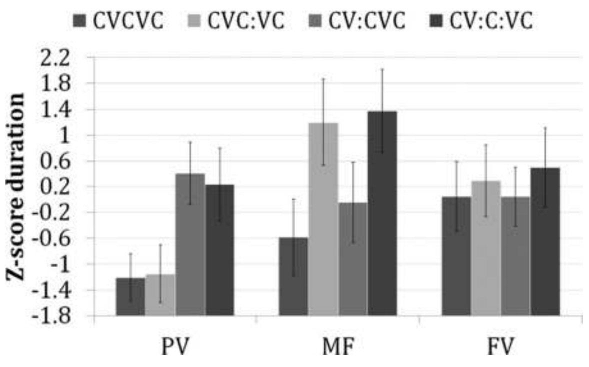


FIG. 2. Z-score duration of the vowel preceding the singleton-geminate consonants (PV), the medial fricative (MF), and the vowel following the singleton-geminate consonants (FV), presented in the four syllabic shapes.

Longer duration of the medial fricative (Sec. III A), potential reduction in the duration of the vowel preceding the medial fricative (Sec. III A), and centralization based on formant measures (Sec. III F); higher intensity and *f*0 in the post-geminate syllable and potentially of the medial fricative and the preceding vowel (Secs. III B and III C); difference in spectral moments, the peak and the dynamic amplitude linked to potential variation in how the medial fricative is realized (Sec. III D); fewer voiced portions in the medial fricative (Sec. III E); and creaky phonation in surrounding vowels with lower dB difference in voice quality correlates (Sec. III G).

III. RESULTS

Our results section summarizes the effects of the different measurements on the singleton vs geminate contrast. All the data discussed refers to Z-score results. As the consonant length factor was of primary interest in this study, the vowel length and the phoneme factors were used as control factors to evaluate if the effect of consonant length would still be observed across vowel length and phonemes. We only discuss when necessary results linked to consonant length and to the two-way interaction of consonant length × vowel length and the three-way interaction of consonant length × vowel length × phoneme with a summary of all results obtained (see Table I).

A. Absolute duration

Average durations of the medial fricative and the surrounding vowels are presented in Fig. 2 in addition to the summary of statistical results (see Table I). These results suggest that consonant length was the main contributor to the UNIVARIATE model in the medial fricative and in the following vowel, as these accounted for 38% and 8%, respectively, of the variance associated with absolute duration (see effect size in Table I). Statistical results revealed no significant effects of consonant length on the duration of the preceding vowel, however, graphical results suggest that the long vowel preceding the singleton consonant was significantly longer in CVːCVC compared to CVːCːVC by around +10%, corresponding to approximately +9 ms difference, with a small to moderate effect size (158 vs 149 ms, $p < 0.004$, $d = 0.33$ "SM"). Moving on to the medial fricative, graphical and statistical results showed consonant length and consonant length × vowel length to be significantly different with respect to the duration of the medial fricative. Singleton medial fricatives were significantly shorter than geminate medial fricatives with a very large effect size (on average −83 ms). The duration of the medial fricative in CVCVC was shorter than in CVCːVC by −217%, equivalent to approximately −91 ms (107 vs 198 ms, $p < 0.0001$, $d = -2.83$ "L") and the duration of the medial fricative in CVːCVC was shorter than in CVːCːVC by −90%, equivalent to approximately −73 ms (135 vs 208 ms, $p < 0.0001$, $d = -2.24$ "L"). Results obtained for the following vowel revealed the same pattern of lengthening in the geminate environments compared to singletons with a moderate to large effect sizes (on average −17 ms in singletons; see Fig. 2). Duration of the following vowel was significantly shorter in CVCVC compared to CVCːVC by around −19%, equivalent to approximately −12 ms (140 vs 152 ms, $p < 0.0001$, $d = -0.44$ "M") and significantly shorter in CVːCVC compared to CVːCːVC by around −33.5%, equivalent to approximately −24 ms (140 vs 163 ms, $p < 0.0001$, $d = -0.85$ "L").

These results show that the duration of the medial geminate fricative is the main contributor to the variance associated with the UNIVARIATE model, which is compatible with previous research (see Sec. I) and (the same patterns are observed for relative duration, whereby the geminate consonant contributes to almost 50% of the duration of the whole VCV syllable, see Khattab and Al-Tamimi, 2014). Our results on the preceding vowel are in accordance with previous results that show temporal compensation in the geminate environment only in the long vowel context, although the difference is below the just noticeable difference (JND) in duration discrimination (see, e.g., Stevens, 1998, pp. 228–229). And finally, the vowel following the singleton/geminate fricatives showed lengthening of its duration in the two geminate environments which is close to the JND in duration discrimination (see, e.g., Stevens, 1998, pp. 228–229).

B. Intensity

Graphical results of the intensity (dB) obtained at the different positions of the medial fricative and the surrounding vowels are presented in Fig. 3. Statistical results showed that consonant length accounted for 9% of the variance associated with intensity at the offset of the medial fricative and at the onset of the following vowel, with a moderate effect size (see Table I). All the other positions revealed a mixed picture due to vowel length being the main contributor to the model. Statistical results also revealed no significant two-way and three-way interactions between consonant length and vowel length, indicating that most of the observed differences in Fig. 3 are due to quality differences between the short and long vowel (see Sec. II A). With respect to consonant length effects split by vowel length, results suggest no differences throughout the vowel preceding the singleton/geminate consonants. Graphical results obtained in the

TABLE I. ANOVA summary table for consonant length (CL) × vowel length (VL) × phoneme (P), with degrees of freedom of each factor (df_1) and of the error (df_2), with F (values in bold are significant at $p \leq 0.023$ after FDR alpha correction) and ω^2 values (small = <0.01; moderate = >0.06 and <0.14; large = >0.14). (PV = preceding vowel, MF = medial fricative, FV = following vowel, On = onset, Md = mid, Of = offset, Vd = voiced, Vs = voiceless.)

		df_2	CL $df_1 = 1$	VL $df_1 = 1$	P $df_1 = 9$	CL × VL $df_1 = 1$	CL × P $df_1 = 9$	VL × P $df_1 = 6$	CL × VL × P $df_1 = 6$
Duration	PV	*639*	3.9, 0.00	**2958.3**, 0.38	**71.8**, 0.13	**15.8**, 0.00	**2.8**, 0.00	**2.7**, 0.00	**2.5**, 0.00
	MF	*639*	**2694.8**, 0.35	**53.5**, 0.01	**123.7**, 0.22	**43.2**, 0.01	**2.9**, 0.00	1.0, 0.00	**2.5**, 0.00
	FV	*639*	**70.6**, 0.08	3.2, 0.00	**7.0**, 0.07	**6.5**, 0.01	1.0, 0.00	1.2, 0.00	1.8, 0.01
Intensity	PV Md	*639*	1.9, 0.00	2.2, 0.00	**5.1**, 0.05	0.3, 0.00	1.2, 0.00	2.0, 0.01	0.9, 0.00
	PV Of	*639*	0.4, 0.00	**34.8**, 0.04	**7.5**, 0.07	2.1, 0.00	1.2, 0.00	1.3, 0.00	0.6, 0.00
	MF On	*639*	0.1, 0.00	**37.5**, 0.04	**16.0**, 0.15	2.6, 0.00	1.0, 0.00	1.7, 0.00	0.6, 0.00
	MF Md	*639*	**21.9**, 0.02	**11.8**, 0.01	**49.3**, 0.36	0.2, 0.00	**2.7**, 0.01	0.8, 0.00	0.1, 0.00
	MF Of	*639*	**117.8**, 0.09	**156.7**, 0.12	**13.3**, 0.09	1.9, 0.00	**4.2**, 0.02	**2.5**, 0.01	1.3, 0.00
	FV On	*639*	**68.6**, 0.07	**143.1**, 0.13	**4.3**, 0.03	3.2, 0.00	**3.6**, 0.02	**3.4**, 0.01	0.9, 0.00
	FV Md	*639*	**14.5**, 0.02	**120.4**, 0.12	**3.1**, 0.02	1.1, 0.00	1.9, 0.01	1.4, 0.00	1.6, 0.00
*F*0	PV Md	*639*	**8.1**, 0.01	0.4, 0.00	**3.0**, 0.02	3.6, 0.00	**2.8**, 0.02	**2.6**, 0.01	1.4, 0.00
	PV Of	*639*	**21.8**, 0.02	**117.2**, 0.11	**9.0**, 0.08	3.2, 0.00	**2.9**, 0.02	1.5, 0.00	1.6, 0.00
	MF On	*617*	**26.5**, 0.03	**104.4**, 0.11	**10.6**, 0.10	1.5, 0.00	**2.4**, 0.01	1.1, 0.00	1.4, 0.00
	MF Md	*186*	**7.6**, 0.02	3.7, 0.01	**15.2**, 0.31	0.0, 0.00	**2.7**,[a] 0.04	1.0, 0.00	**3.3**,[b] 0.02
	MF Of	*352*	**8.6**, 0.01	**12.2**, 0.01	**33.9**, 0.38	1.4, 0.00	1.6, 0.01	1.0, 0.00	**2.7**,[c] 0.01
	FV On	*639*	**38.5**, 0.05	**13.0**, 0.01	**7.0**, 0.07	3.8, 0.00	1.1, 0.00	**2.9**, 0.01	0.7, 0.00
	FV Md	*639*	**8.2**, 0.01	1.8, 0.00	1.6, 0.01	1.0, 0.00	1.1, 0.00	1.4, 0.00	0.2, 0.00
Peak	MF	*639*	4.5, 0.00	0.3, 0.00	**202.9**, 0.65	0.2, 0.00	0.6, 0.00	1.4, 0.00	0.6, 0.00
M1	MF	*639*	**5.9**, 0.00	0.2, 0.00	**348.4**, 0.72	1.7, 0.00	0.8, 0.00	0.9, 0.00	0.4, 0.00
M2	MF	*639*	1.6, 0.00	0.0, 0.00	**64.6**, 0.43	1.2, 0.00	**3.2**, 0.02	0.9, 0.00	0.3, 0.00
L3	MF	*639*	1.1, 0.00	0.0, 0.00	**27.2**, 0.25	0.0, 0.00	0.9, 0.00	0.4, 0.00	0.3, 0.00
L4	MF	*639*	1.2, 0.00	0.0, 0.00	**12.1**, 0.13	0.1, 0.00	1.3, 0.00	0.5, 0.00	0.1, 0.00
A_d	MF	*639*	**109.9**, 0.04	2.1, 0.00	**165.9**, 0.57	1.0, 0.00	**4.9**, 0.01	1.3, 0.00	0.4, 0.00
Voicing	MF Vd	*184*	**44.2**, 0.14	**18.4**, 0.06	2.8,[d] 0.01	4.1, 0.01	1.0,[d] 0.00	1.0,[e] 0.00	2.3,[e] 0.00
	MF Vs	*455*	**89.0**, 0.06	0.4, 0.00	**123.0**,[f] 0.50	0.3, 0.00	1.3,[f] 0.00	2.0,[b] 0.00	2.8,[b] 0.01
F_1	PV Md	*639*	**23.2**, 0.02	**237.2**, 0.15	**18.7**, 0.12	**23.5**, 0.02	**4.2**, 0.02	**15.4**, 0.07	**4.3**, 0.02
	PV Of	*639*	1.1, 0.00	**243.5**, 0.09	**127.1**, 0.44	1.5, 0.00	**2.7**, 0.01	**11.9**, 0.03	2.3, 0.00
	FV On	*639*	**5.5**, 0.00	**216.2**, 0.07	**168.2**, 0.46	**28.2**, 0.01	**6.1**, 0.01	**8.3**, 0.01	**9.2**, 0.02
	FV Md	*639*	**7.1**, 0.00	**1326.5**, 0.29	**32.1**, 0.09	0.1, 0.00	**25.0**, 0.07	**17.3**, 0.03	**36.3**, 0.06
F_2	PV Md	*639*	**39.9**, 0.01	**398.4**, 0.12	**117.0**, 0.36	**58.2**, 0.02	**11.5**, 0.03	**48.1**, 0.10	**15.7**, 0.03
	PV Of	*639*	**12.4**, 0.01	**318.9**, 0.13	**87.7**, 0.34	**49.3**, 0.02	**5.2**, 0.02	**30.1**, 0.08	**5.8**, 0.01
	FV On	*639*	3.0, 0.00	**812.2**, 0.23	**75.9**, 0.24	**18.3**, 0.01	**16.3**, 0.05	**16.7**, 0.03	**18.8**, 0.04
	FV Md	*639*	**5.8**, 0.00	**1759.2**, 0.29	**66.3**, 0.14	1.0, 0.00	**50.1**, 0.10	**24.9**, 0.03	**61.8**, 0.08
F_3	PV Md	*639*	0.0, 0.00	**45.6**, 0.06	1.5, 0.01	0.5, 0.00	**2.8**, 0.02	**4.0**, 0.02	2.2, 0.01
	PV Of	*639*	1.3, 0.00	1.5, 0.00	**12.2**, 0.12	1.8, 0.00	**3.5**, 0.03	**3.8**, 0.02	0.7, 0.00
	FV On	*639*	0.8, 0.00	1.4, 0.00	**23.9**, 0.23	1.6, 0.00	1.3, 0.00	1.7, 0.00	0.4, 0.00
	FV Md	*639*	0.1, 0.00	0.1, 0.00	**9.9**, 0.10	0.5, 0.00	1.3, 0.00	**3.2**, 0.02	0.9, 0.00
$*H_1$-$*H_2$	PV Of	*624*	0.2, 0.00	0.3, 0.00	1.6, 0.01	2.8, 0.00	0.9, 0.00	0.3, 0.00	1.1, 0.00
	FV On	*638*	1.4, 0.00	0.0, 0.00	0.5, 0.00	0.0, 0.00	0.3, 0.00	1.4, 0.00	0.3, 0.00
$*H_1$-$*A_1$	PV Of	*624*	1.5, 0.00	3.3, 0.00	1.3, 0.00	**10.2**, 0.01	0.8, 0.00	1.4, 0.00	0.7, 0.00
	FV On	*638*	**8.1**, 0.01	3.8, 0.00	**3.0**, 0.03	3.3, 0.00	1.0, 0.00	2.3, 0.01	1.2, 0.00
$*H_1$-$*A_2$	PV Of	*624*	2.0, 0.00	**8.3**, 0.01	1.1, 0.00	0.2, 0.00	1.2, 0.00	1.9, 0.01	1.5, 0.00
	FV On	*638*	4.3, 0.00	**7.0**, 0.01	1.5, 0.01	1.7, 0.00	0.8, 0.00	0.3, 0.00	0.6, 0.00
$*H_1$-$*A_3$	PV Of	*624*	3.3, 0.00	0.4, 0.00	**2.2**, 0.02	0.1, 0.00	0.9, 0.00	0.5, 0.00	1.1, 0.00
	FV On	*638*	0.4, 0.00	0.2, 0.00	**2.8**, 0.02	0.5, 0.00	0.6, 0.00	1.4, 0.00	1.0, 0.00

[a] $df_1 = 8$.
[b] $df_1 = 4$.
[c] $df_1 = 5$.
[d] $df_1 = 2$.
[e] $df_1 = 1$.
[f] $df_1 = 6$

medial fricative show a U-shape pattern with the highest values at the onset and the lowest at the midpoint, with the former potentially linked to the trochaic stress pattern (see Fig. 3). At the onset of the medial fricative, no differences were observed. At the midpoint, singleton environments showed higher intensity values (on average +2 dB); intensity values were significantly higher only in CVCVC compared with CVCːVC by around +31%,

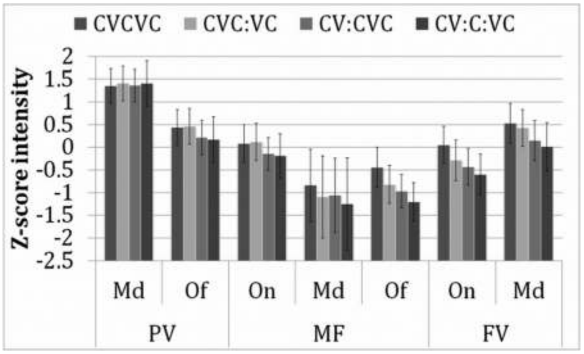


FIG. 3. Z-score intensity of preceding vowel (PV), medial fricative (MF), and following vowel (FV) (On = onset, Md = midpoint, Of = offset).

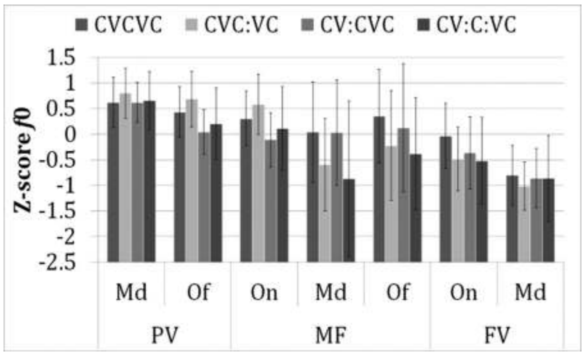


FIG. 4. Z-score *f*0 of preceding vowel (PV), medial fricative (MF), and following vowel (FV) (On = onset, Md = midpoint, Of = offset).

equivalent to approximately +1 dB (63 vs 62 dB, $p < 0.002$, $d = 0.29$ "SM"). At the offset of the consonant, intensity was significantly higher in singleton environments (on average +2 dB); a significantly higher intensity was found in CVCVC compared to CVC:VC by around +38%, equivalent to approximately +2 dB, with a large effect size (65 vs 63 dB, $p < 0.0001$, $d = 0.88$ "L") and a significantly higher intensity at the offset of the medial fricative was observed in CV:CVC than in CV:C:VC by around +32%, equivalent to approximately +1 dB (62 vs 61 dB, $p < 0.0001$, $d = 0.60$ "ML"). Moving on to the following vowels, graphical and statistical results show the same patterns whereby singleton environments show significantly higher intensity compared to geminates (on average +2 dB, see Fig. 3 and Table I). Looking at the results by vowel quality, intensity values were significantly higher at the onset of the following vowel in CVCVC compared to CVC:VC by around +26%, equivalent to approximately +2 dB (67 vs 65 dB, $p < 0.0001$, $d = 0.79$ "L") and a significantly higher intensity in CV:CVC than in CV:C:VC by around + 18%, equivalent to approximately +1 dB (65 vs 64 dB, $p < 0.001$, $d = 0.40$ "SM"). At the midpoint of the following vowel, intensity values were significantly higher in CVCVC compared to CVC:VC by around +6%, equivalent to approximately +1 dB difference (70 vs 69 dB, $p < 0.01$, $d = 0.25$ "SM") and significantly higher in CV:CVC compared to CV:C:VC by around +10%, equivalent to approximately +1 dB (68 vs 67 dB, $p < 0.01$, $d = 0.28$ "SM").

Intensity results therefore showed the *opposite* of what is expected for the singleton vs geminate contrast (Local and Simpson, 1988), with the medial fricative and the following vowel having higher intensity in the singleton than in the geminate context. Although the differences were in the region of +1 to +2 dB difference, which is equivalent to the JND in amplitude discrimination (see, e.g., Stevens, 1998, pp. 225–226), they were still relatively small and the main patterns for the intensity measure merely show the influence of the trochaic stress pattern.

C. Fundamental frequency

*f*0 results were obtained at different positions of the medial fricative and surroundings vowels, and the results are presented graphically in Fig. 4. In a separate UNIVARIATE analysis, we included voicing as a fourth factor, and none of the high order interactions were significant, which in turn suggests that even though voicing of the consonant may affect the overall *f*0 results, it is not one of the main contributors to the singleton vs geminate contrast, and so voicing was omitted from the final model. The same insignificant differences were observed at the high order interactions between consonant length × vowel length × phoneme, suggesting that the observed patterns are present regardless of the voicing/place of articulation distinction (see Table I). Consonant length was not the main contributor to the UNIVARIATE analysis as only 2% to 5% of the variance was accounted for by this factor; instead, vowel length and/or phoneme were the main contributors to the model (see Table I).

Graphical results showed an overall rise in *f*0 in the geminate environments from the midpoint of the preceding vowel up to the onset of the medial fricative, followed by an overall lowering of frequencies from the midpoint of the medial fricative up to the onset of the following vowel. When looking at the effect of consonant length, results obtained in the preceding vowel showed lower *f*0 in singleton environments (at midpoint: on average −3 Hz for female and −2 Hz for male in singletons; at offset: on average −6 Hz for female subjects and −2 Hz for male subjects in singletons). At the midpoint of the preceding short vowel, *f*0 was significantly lower in CVCVC compared to CVC:VC by around −8%, equivalent to approximately −2 Hz for female subjects (231 vs 233 Hz) and to approximately −3 Hz for male subjects (128 vs 131 Hz) ($p < 0.0001$, $d = -0.36$ "SM"). At the offset of the preceding vowel, the same pattern was observed, i.e., singleton environments showed significantly lower *f*0 in CVCVC compared to CVC:VC by around −13%, equivalent to approximately −5 Hz for female subjects (227 vs 231 Hz) and approximately −4 Hz for male subjects (126 vs 130 Hz) ($p < 0.0001$, $d = -0.49$ "M") and significantly lower *f*0 in CV:CVC compared to CV:C:VC by around −12%, equivalent to approximately −8 Hz for female subjects (214 vs 222 Hz) and no differences for male subjects (122 vs 122 Hz) ($p < 0.012$, $d = -0.28$ "SM").

Moving on to the medial fricative, our results seem to show an on-off phonetic voicing regardless of phonological

TABLE II. Number of phonologically voiced or voiceless medial fricative consonants that are realized as phonetically voiced (with percentages in brackets), analyzed at onset, midpoint, and offset. The total number of fricatives analyzed is 1726, consisting of 488 phonologically voiced and 1238 phonologically voiceless fricatives.

	Onset	Midpoint	Offset
Phonologically voiced	470 (96%)	269 (55%)	391 (80%)
Phonologically voiceless	1085 (88%)	149 (12%)	322 (26%)
Total phonetically voiced	**1555 (90%)**	**418 (24%)**	**713 (41%)**

status of the fricative. At the onset, almost all the consonants turn out to be phonetically voiced (both phonologically voiced and voiceless); while less than one quarter are voiced mid-fricative and one third are voiced at the offset (see Table II). At the onset of the medial fricative, lower *f*0 was observed in singleton environments (on average −8 Hz for female subjects and −3 Hz for male subjects in singletons). Significantly lower *f*0 was obtained in CVCVC compared to CVCːVC, by around −16%, equivalent to approximately −5 Hz for female subjects (225 vs 230 Hz) and approximately −4 Hz for male subjects (124 vs 128 Hz), ($p < 0.0001$, $d = -0.50$ "M"), and significantly lower *f*0 was obtained in CVːCVC compared to CVːCːVC, by around −19%, equivalent to approximately −12 Hz for female subjects (210 vs 222 Hz) and no differences for male subjects (121 vs 121 Hz) ($p < 0.004$, $d = -0.33$ "SM"). At the midpoint of the medial fricative, a reverse pattern is observed; singleton environments showed significantly higher *f*0 than geminates (on average +17 Hz for female and +8 Hz for male). *F*0 was significantly higher in CVCVC compared to CVCːVC, by around +47%, equivalent to approximately +16 Hz for female subjects (216 vs 200 Hz) and to approximately +8 Hz for male subjects (121 vs 113 Hz) ($p < 0.0001$, $d = 0.67$ "ML"), and significantly higher *f*0 was obtained in CVːCVC compared to CVːCːVC, by around +63%, equivalent to approximately +20 Hz for female subjects (212 vs 192 Hz) and to approximately +10 Hz for male subjects (121 vs 111 Hz), ($p < 0.003$, $d = 0.77$ "L"). *f*0 at the offset of the medial fricative showed the same patterns; singleton environments showed significantly higher *f*0 than geminates (on average +5 Hz for female subjects and +9 Hz for male subjects). Looking at effects of vowel length, *f*0 was significantly higher in CVCVC compared to CVCːVC, by around +35%, equivalent to approximately +8 Hz for female subjects (212 vs 204 Hz) and to approximately +10 Hz for male subjects (129 vs 119 Hz) ($p < 0.0001$, $d = 0.58$ "ML"), and significantly higher *f*0 frequencies in CVːCVC compared to CVːCːVC, by around 36%, equivalent to approximately +1 Hz for female subjects (204 vs 205 Hz) and to approximately +8 Hz for male subjects (127 vs 119 Hz) ($p < 0.009$, $d = 0.42$ "SM").

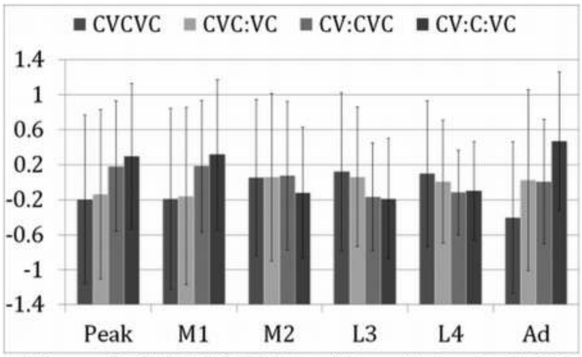
Graphical results on the following vowel showed the same pattern reported above, with singleton environments showing higher *f*0 only in the short preceding vowel environments (on average +7 Hz for female and +5 Hz for male). Statistical results showed that *f*0 was significantly higher at the onset of the vowel following the fricative in CVCVC compared to CVCːVC, by around +36%, equivalent to approximately +11 Hz for female subjects (205 vs 194 Hz) and to approximately + 5 Hz for male subjects (123 vs 118 Hz) ($p < 0.0001$, $d = 0.71$ "ML"). At the midpoint, significantly higher *f*0 was observed in singleton environments (on average +3 Hz for female subjects and +1 Hz for male subjects), with higher *f*0 values observed in CVCVC compared to CVCːVC, by around 27%, equivalent to approximately +6 Hz for female subjects (188 vs 182 Hz) and to approximately +2 Hz for male subjects (114 vs 112 Hz) ($p < 0.0001$, $d = 0.41$ "SM").

*f*0 results showed the same pattern linked to the singleton vs geminate contrast and to the trochaic stress pattern (see also Sec. III B). From the midpoint of the preceding vowel to the onset of the medial fricative, significantly higher *f*0 was obtained in the geminate context, followed by low values from the midpoint of the medial fricative to the midpoint of the following vowel, suggesting that gemination accentuates the high-low *f*0 pattern that is typical of the trochaic context. These patterns were the same in both voiced and voiceless fricatives, although the singleton vs geminate differences were larger in the latter. In almost all cases, raw *f*0 frequency differences seem to be linked to sex-based differences with larger differences in female subjects, although these were substantially reduced when *f*0 values were Z-scored (with an average difference of 3% to 7%). Even though absolute *f*0 frequency differences are in some cases low, they seem to be close to the JND in pitch discrimination, which is close to 1 Hz difference for complex tones with frequencies between 80 and 500 Hz (Kollmeier *et al.*, 2008, p. 65). The observed differences between female and male subjects tend to fall within the JND of pitch discrimination (see, e.g., Stevens, 1998, pp. 227–228).

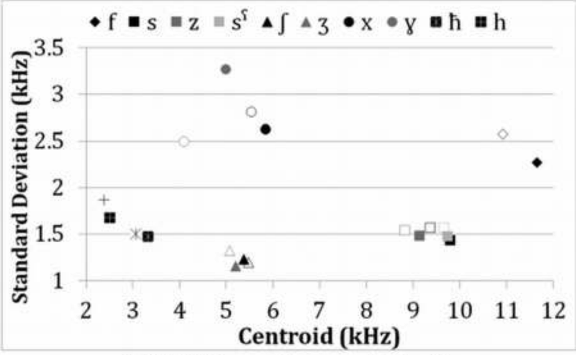
D. Spectral moments in the medial fricative

Graphical results and a summary of statistical results for the peak, the centroid (M1), the standard deviation (M2), the skewness (L3), the kurtosis (L4), and the dynamic amplitude (A_d) are presented in Fig. 5(a) and Table I, respectively. Dynamic aspects of the fricative show that consonant length accounted for some of the variance associated with the centroid (M1) and the dynamic amplitude (A_d); peak frequency, standard deviation (M2), skewness (L3), and kurtosis (L4) showed no significant differences (Table I). Geminate fricatives showed significantly higher centroid (M1) (on average +221 Hz) and a higher dynamic amplitude (A_d) (on average +8 dB), and, although not statistically significant, higher peak (on average +273 Hz) and lower standard deviation (on average −70 Hz) than singleton fricatives [see Fig. 5(a) and Table I].

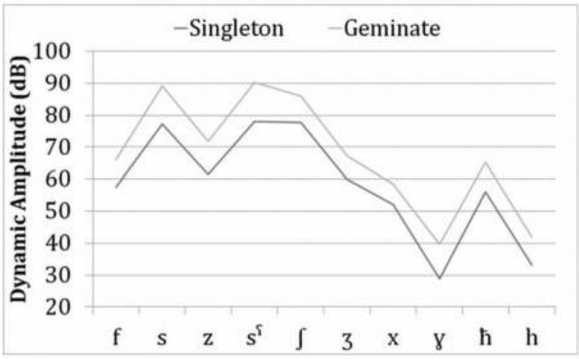
Although statistical significance was not reached in the two-way interaction of consonant length × vowel length, there were minor differences linked to vowel type; these suggest that the same patterns are observed, albeit smaller/larger differences. Starting with the peak, singleton environments showed lower peak frequencies in CVCVC compared to CVCːVC by around −5%, equivalent to approximately



a) Z-score Peak, M1, M2, L3, L4 and Dynamic Amplitude (A_d)



b) M1*M2 acoustic space of fricatives



c) Dynamic Amplitude (A_d) differences

FIG. 5. Z-score of the peak, the four spectral moments and dynamic amplitude (A_d) of fricative consonants (a), the (M1*M2) acoustic space of fricative consonants, with geminates in black/gray filled symbols (b), and results of dynamic amplitude (A_d) by fricative phonemes in singleton and geminate environments (c).

−208 Hz (6302 vs 6510 Hz, ns, $d = -0.05$ "S") and lower peak frequencies in CVːCVC compared to CVːCːVC by around −8%, equivalent to approximately −400 Hz (7610 vs 8010 Hz, ns, $d = -0.15$ "S"). Centroid (M1) frequencies showed the same patterns as the peak, whereby singleton environments showed lower centroid (M1) frequencies in CVCVC compared to CVCːVC by around −3%, equivalent to approximately −116 Hz (6605 vs 6721 Hz, ns, $d = -0.02$ "S") and a lower centroid (M1) frequencies in CVːCVC compared to CVːCːVC by around −9%, equivalent to approximately −400 Hz (7787 vs 8187 Hz, ns, $d = -0.16$ "S"). The standard deviation (M2) results suggest that singleton environments show significantly higher standard deviation (M2) values only in CVːCVC compared to CVːCːVC by around +15%, equivalent to approximately +185 Hz (1832 vs 1647 Hz, $p < 0.02$, $d = 0.25$ "SM"). And finally, results obtained for the dynamic amplitude (A_d) showed significantly lower dynamic amplitude (A_d) values in CVCVC compared to CVCːVC by around −49%, equivalent to approximately −8 dB (60 vs 68 dB, $p < 0.0001$, $d = -0.44$ "M") and a significantly lower dynamic amplitude (A_d) values in CVːCVC compared to CVːCːVC by around −18%, equivalent to approximately −8 dB (68 vs 76 dB, $p < 0.0001$, $d = -0.6$ "M").

Looking at spectral moments, results suggest that geminate environments are associated with an overall higher centroid (M1) and lower standard deviation (M2). To evaluate potential effects of phoneme identity on these patterns, the centroid*standard deviation (M1*M2) acoustic space of fricative consonants show most fricative phonemes as having different positions between singletons (white filled symbols) and geminates (black/gray filled symbols) either on the two-axes or on only one [see Fig. 5(b)]. Although the differences seem minimal, it is consistent across phonemes with large differences for /f s z ʒ x ɤ ħ h/ (ns to $p < 0.001$, $d > 0.2$ to < 1.1, "S to L").

Dynamic amplitude (A_d) results seem also to suggest that geminate environments are associated with an overall higher dynamic amplitude (A_d) than singletons and this can be seen across all phonemes [see Fig. 5(c)]. The observed differences ranged between +6 dB for /x/ and +12 dB for /sˤ/ in the geminate environment ($p < 0.01$ to $p < 0.0001$, $d > 0.5$ to < 1 "M to L").

Geminate environments show significantly higher centroid (M1) and dynamic amplitude (A_d) values, higher peak frequency and lower standard deviation (M2), with the same pattern present in most of the individual fricative phonemes [see Figs. 5(a), 5(b), and 5(c)]. Our results suggest that, compared with singletons, geminate fricatives show a significant increase in the centroid (M1) and dynamic amplitude (A_d), which can be correlated with an increase in flow velocity in the constriction; this in turn can be the result of either (a) the constriction area decreasing, (b) the volume velocity increasing, or a combination of both (e.g., Shadle, 2012, p. 521). A high frequency boost can also be correlated with the increase in effort level of sustained fricatives (e.g., Shadle, 2012, p. 521); which in the case of geminate fricatives is potentially due to their longer durations. The higher centroid (M1) indicates a shorter front resonating cavity leading to more front articulations (Forrest *et al.*, 1988; Jongman *et al.*, 2000; Li *et al.*, 2009; Maniwa *et al.*, 2009, among others). These results show that geminate fricatives behave as a different set compared to singleton fricatives in the sense that they potentially show hyper-articulated productions (Lindblom, 1990; Maniwa *et al.*, 2009) reflecting stronger articulations (Kohler, 1984).

E. Medial fricative voicing patterns

The influence of gemination on voicing patterns in voiced and voiceless medial fricatives was explored. Starting with phonologically voiced fricatives, statistical results show that consonant length accounted for 14% of the variance associated with voicing patterns (see Table I). Overall, there were fewer voiced frames in the two geminate environments

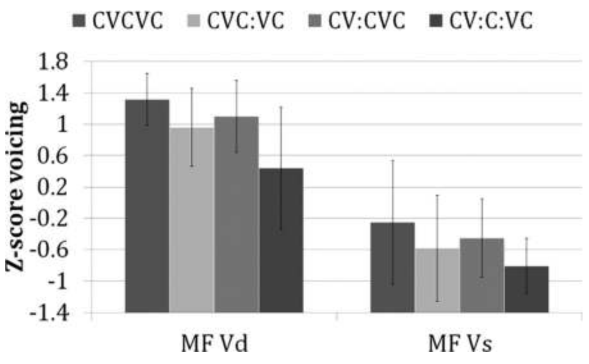


FIG. 6. Z-score of the proportion of voicing (in %) relative to the length of the fricative in voiced (Vd) and voiceless (Vs) medial fricative (MF).

(on average +17% of the total proportion of voicing in singletons, see Fig. 6). The proportion of voiced frames was significantly larger in CVCVC compared to CVCːVC, by around +8%, equivalent to approximately +12% of the total proportion of voicing in the medial fricative (97% vs 85%, $p < 0.0001$, $d = 0.86$ "L") and significantly larger in CVːCVC compared to CVːCːVC, by around +23%, equivalent to approximately +24% of the total proportion of voicing in the medial fricative (89% vs 65%, $p < 0.0001$, $d = 1.07$ "L").

In phonologically voiceless fricatives, consonant length accounted for only 6% of the variance associated with voicing patterns (see Table I). There were again fewer voiced frames in the two geminate environments compared to the singletons with a moderate effect size (on average +12% of the total proportion of voicing in singleton, see Fig. 6). The proportion of voiced frames was significantly larger in CVCVC compared to CVCːVC, by around +30%, equivalent to approximately +12% of the total proportion of voicing in the medial fricative (41% vs 29%, $p < 0.0001$, $d = 0.45$ "M") and significantly larger in CVːCVC compared to CVːCːVC, by around +36%, equivalent to approximately +12% of the total proportion of voicing in the medial fricative (33% vs 21%, $p < 0.0001$, $d = 0.83$ "L").

In sum, fewer voiced frames were observed in geminate environments compared with singletons ones; this was exhibited in both phonologically voiced and voiceless fricatives, with more devoicing in voiced categories and fewer voicing shadows in the voiceless category. These results are comparable with the patterns normally found in tense (fortis) categories (Jaeger, 1983).

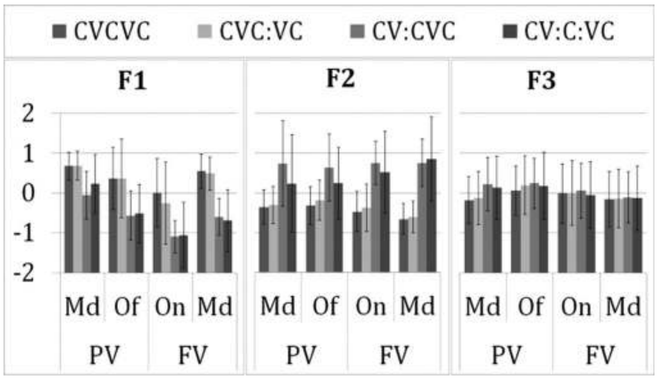


FIG. 7. Z-score of formant frequencies of F_1, F_2, and F_3 at the different positions of preceding vowel (PV) and following vowel (FV) (On = onset, Md = midpoint, Of = offset).

F. Formant frequencies of surrounding vowels

Formant frequencies of F_1, F_2, and F_3 at the different positions in the preceding and following vowels surrounding the singleton/geminate fricatives are presented Fig. 7. Statistical results presented in Table I show that consonant length accounted for less than 2% of the total variance associated with F_1 and F_2, with the highest contributor being vowel length, followed by Fricative phonemes. The observed differences were mainly linked to qualitative differences in the vowel preceding the singleton/geminate fricatives with, overall, higher F_1 frequencies in the two short environments CVCVC and CVCːVC compared to the long environments CVːCVC and CVːCːVC, with the former group being realized as [a] and the latter realized between [eː] and [ɛː] (see Fig. 7). The three-way interaction of consonant length × vowel length × phoneme was significant with small effect size (see Table I), however, the direction of difference was always linked to differences in vowel type differences rather than in consonant length differences.

Differences in the singleton vs geminate environments were not highly significant with regard to consonant length and the consonant length × vowel length interaction. Starting with the vowel preceding the singleton/geminate fricatives, there were no significant differences in F_1 or F_2 measures at the mid-point of the short vowel preceding singleton and geminate fricatives. At the offset, however, both F_2 and F_3 showed significant differences. F_2 at the offset of the preceding vowel shows significantly lower frequencies in the singleton CVCVC compared to CVCːVC, by around −14%, equivalent to approximately −41 Hz in female subjects (1680 vs 1721) and to approximately −38 Hz in male subjects (1314 vs 1352 Hz) ($p < 0.002$, $d = -0.28$ "SM"). F_3 at the offset of the preceding vowel shows significantly lower frequencies in the singleton CVCVC compared to CVCːVC, by around −11%, equivalent to approximately −29 Hz in female subjects (3077 vs 3106) and to approximately −22 Hz in male subjects (2645 vs 2667 Hz) ($p < 0.018$, $d = -0.21$ "SM").

Results obtained for the long preceding vowel context showed that F_1 frequencies at the midpoint were significantly lower in the singleton CVːCVC compared to CVːCːVC, by around −24%, equivalent to approximately −67 Hz in female subjects (647 vs 714) and to approximately −23 Hz in male subjects (563 vs 586 Hz) ($p < 0.0001$, $d = -0.43$ "SM"). F_2 frequencies obtained at the midpoint of the preceding vowel were significantly higher in the singleton CVːCVC compared to CVːCːVC, by around +23%, equivalent to approximately +154 Hz in female subjects (2028 vs 1874) and to approximately +118 Hz in male subjects (1557 vs 1439 Hz) ($p < 0.0001$, $d = 0.44$ "SM"). And at the offset, F_2 frequencies were significantly higher in CVːCVC compared to CVːCːVC, by around +20%, equivalent to approximately +111 Hz in female subjects (1971 vs 1860) and to approximately +101 Hz in male subjects (1545 vs 1444 Hz) ($p < 0.0001$, $d = 0.46$ "M").

Moving on to the vowel *following* the singleton/geminate fricatives (see Fig. 7), statistical and graphical results show significant differences at the onset of F_1 and F_2. Significantly higher F_1 frequencies were obtained at the onset of the following vowel in the singleton CVCVC compared to CVCːVC, by around +20%, equivalent to approximately +26 Hz in female subjects (678 vs 652 Hz) and to approximately +37 Hz in male subjects (560 vs 523 Hz) ($p < 0.003$, $d = 0.27$ "SM"). And significantly higher F_2 frequencies were obtained at the onset of the following vowel in the singleton CVːCVC compared to CVːCːVC, by around +10%, equivalent to approximately +84 Hz in female subjects (2022 vs 1938 Hz) and to approximately +45 Hz in male subjects (1566 vs 1521 Hz) ($p < 0.011$, $d = 0.29$ "SM").

Geminate environments exhibit different effects on surrounding vowels depending on phonological vowel length. Phonologically short preceding vowel environments show no effects on F_1, but a more retracted production (lower F_2) is observed at the offset. Following vowels in this context had closer (lower F_1) and more retracted productions (lower F_2) at their onset. As for the phonologically long preceding vowels, which were shortened in the geminate environment (see Sec. III A), more open (higher F_1) and more retracted (lower F_2) productions were found at the midpoint, suggesting that preceding long vowels were centralized in geminate environments. F_3 did not seem to contribute to the singleton vs geminate contrast, although marginal high F_3 frequencies were found at the offset of the preceding short vowel. Raw frequency differences between the singleton and geminate environments were all in the range of the JND in frequency discrimination which is close to 3 Hz (for frequencies below 500 Hz) and 0.6% for frequencies above 1000 Hz (Kollmeier *et al.*, 2008, p. 65). In sum, vowels preceding geminate environments are only centralized when the preceding vowel is long; this happens to a lesser extent when the vowel is short, which do not show any shortening before geminate fricatives (see Sec. III A). These results point to a relationship between shortening and centralization in vowels preceding geminate environments (see Sec. IV).

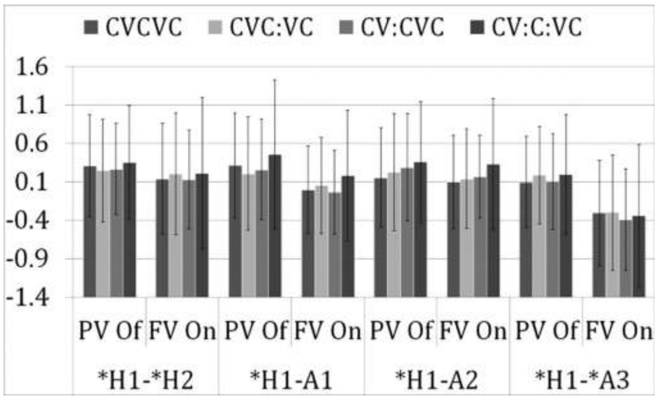


FIG. 8. Z-score voice quality correlates: $*H_1$-$*H_2$, $*H_1$-A_1, $*H_1$-A_2, and $*H_1$-$*A_3$, at offset (Of) of preceding vowel (PV) and onset (On) of following vowel (FV).

G. Phonation measures

Voice quality correlates were used to assess potential effects of the singleton vs geminate environments on surrounding vowels. Statistical and graphical results are presented in Table I and in Fig. 8, respectively. As can be seen from the statistical results, significant differences were obtained for consonant length only on the $*H_1$–A_1 metric in the following vowel, when the preceding vowel is long (on average −1 dB in singletons). When looking at the differences in consonant length by vowel type, results showed that $*H_1$–A_1 is significantly lower in CVːCVC compared to CVːCːVC, by around −18%, equivalent to approximately −2 dB in (−14 vs −12 dB, $p < 0.006$, $d = -0.31$ "SM"). Although statistical significance was not reached (after FDR alpha correcting) there was a tendency for the following metrics to show lower amplitude differences in the singleton environments in the long preceding vowel environment: $*H_1$–A_1 and $*H_1$–A_2. $*H_1$–A_1 values at the offset of the preceding vowel were lower in CVːCVC compared to CVːCːVC, by around −12%, equivalent to approximately −1 dB in (−11 vs −10 dB, $p = 0.026$ (>0.023), $d = -0.25$ "SM") and $*H_1$–A_2 values at the onset of the following vowel were lower in CVːCVC compared to CVːCːVC, by around −11%, equivalent to approximately −2 dB in (−10 vs −8 dB, $p = 0.03$ (>0.023), $d = -0.23$ "SM").

Raw amplitude differences were in the range of the JND in amplitude discrimination (Stevens, 1998, pp. 225–226). These results exhibit a pattern whereby geminate environments show relatively higher amplitude difference in $*H_1$–A_1 at the offset of the preceding long vowel and at the onset of the following vowel (when the preceding vowel is long); a higher amplitude difference in $*H_1$–A_2 is also exhibited at the onset of the following vowel in the long preceding vowel environment, suggesting that a breathy phonation is associated with geminates.

H. Discriminant analysis and ROC curves

To evaluate the degree to which each acoustic measurement contributed to the singleton vs geminate contrast, acoustic measures were submitted to several linear discriminant analyses, followed by ROC curves analyses of geminate probability classification scores. Results for the best ten classification models are presented in Fig. 9 and for all acoustic measures are summarized in Table III. According to ROC curves, the Duration of the medial fricative was the highest contributor to the contrast with a classification rate of 89% (SD 2%), and an AUC of 96% (SD 1%). This indicates that the duration of the medial fricative predicted the geminate category highly accurately. All subsequent models are linked to non-temporal acoustic cues of medial fricatives such as voicing (of voiced and voiceless consonants); *f*0 (at midpoint and offset); intensity (at offset); dynamic amplitude (A_d) of

TABLE III. Classification rates (SD), AUC (standard error), and asymptotic *p* value for all measurements in the geminate category. Order of results is by highest to lowest AUC rate. (PV = preceding vowel, MF = medial fricative, FV = following vowel, Vd = voiced, Vs = voiceless, Fric = fricative, On = onset, Md = midpoint, Of = offset, ns = not significant.)

	Rate	AUC	p
MF Duration	89%(2%)	96%(1%)	<0.0001
MF $f0$ Md	71%(3%)	76%(4%)	<0.0001
MF voicing Vd Fric	70%(1%)	76%(3%)	<0.0001
MF voicing Vs Fric	62%(8%)	71%(2%)	<0.0001
MF intensity Of	63%(5%)	70%(2%)	<0.0001
FV intensity On	62%(5%)	68%(2%)	<0.0001
MF $f0$ Of	67%(4%)	68%(3%)	<0.0001
FV duration	60%(7%)	67%(2%)	<0.0001
FV $f0$ On	60%(8%)	67%(2%)	<0.0001
MF dynamic amplitude (A_d)	57%(3%)	63%(2%)	<0.0001
MF $f0$ On	58%(0%)	61%(2%)	<0.0001
PV F_2 Of	59%(6%)	61%(2%)	<0.0001
PV $f0$ Of	56%(3%)	60%(2%)	<0.0001
PV F_2 Md	57%(5%)	59%(2%)	<0.0001
FV intensity Md	56%(3%)	58%(2%)	<0.0001
MF intensity Md	56%(1%)	58%(2%)	<0.0001
FV $f0$ Md	54%(3%)	58%(2%)	<0.005
PV $f0$ Md	54%(1%)	57%(2%)	<0.005
PV duration	53%(3%)	56%(2%)	<0.005
FV F_1 Md	56%(6%)	56%(2%)	<0.01
FV F_1 On	53%(5%)	56%(2%)	<0.01
PV F_1 Md	54%(4%)	56%(2%)	<0.01
PV *$H1$-$A1$ Of	56%(5%)	56%(2%)	<0.05
FV F_2 Md	56%(7%)	56%(2%)	<0.05
PV F_3 Of	53%(3%)	55%(2%)	<0.05
FV *$H1$-$A1$ On	54%(5%)	55%(2%)	<0.05
FV F_2 On	51%(0%)	54%(2%)	ns
PV intensity Md	54%(1%)	54%(2%)	ns
PV *$H1$-$A3$ Of	53%(4%)	54%(2%)	ns
PV *$H1$-*$H2$ Of	52%(1%)	54%(2%)	ns
FV *$H1$-$A2$ On	52%(0%)	54%(2%)	ns
MF peak	51%(3%)	54%(2%)	ns
MF intensity On	51%(2%)	53%(2%)	ns
PV F_3 Md	52%(1%)	52%(2%)	ns
PV *$H1$-$A2$ Of	52%(0%)	52%(2%)	ns
MF M1	50%(3%)	52%(2%)	ns
FV F_3 On	52%(3%)	52%(2%)	ns
FV *$H1$-*$H2$ On	54%(0%)	52%(2%)	ns
PV intensity Of	52%(2%)	52%(2%)	ns
FV *$H1$-*$A3$ On	47%(7%)	52%(2%)	ns
MF M2	49%(4%)	51%(2%)	ns
MF L4	50%(2%)	51%(2%)	ns
PV F_1 Of	51%(1%)	51%(2%)	ns
MF L3	49%(0%)	51%(2%)	ns
FV F_3 Md	45%(8%)	51%(2%)	ns

the fricative; and to the duration, intensity, and $f0$ at the onset of the following vowel. However, their contribution to the singleton vs geminate contrast is lower than that of the duration of the medial fricative with classification rates ranging between 51% and 71% and AUC percentages ranging between 51% and 76% (see Table III). These results confirm consonant duration as the main contributor to the singleton vs geminate contrast with secondary non-temporal acoustic cues.

IV. SUMMARY AND DISCUSSION

The aim of this study was to explore the extent to which qualitative differences in consonant realization play a role in the singleton vs geminate distinction in LA. As suggested in the literature on Arabic and other languages, temporal differences are considered to be the primary exponents of the contrast, with consonant duration acting as the main acoustic cue to the distinction between singleton and geminate consonants (Ham, 2001; Hassan, 2003; Khattab, 2007; Ridouane, 2007, among others). However, non-temporal acoustic cues have also been found to play a role in the contrast in other languages and are present as secondary consequences of consonant length (Abramson, 1999; Arvaniti and Tserdanelis, 2000; Esposito and di Benedetto, 1999; Idemaru and Guion, 2008; Local and Simpson, 1999; Payne, 2006; Ridouane, 2007, among others). The prevalence of these cues seems to vary across languages depending on language-specific rules for segment timing and prosodic conditioning. Our aim was therefore to explore how prominent non-temporal cues are in a language where durational contrasts for both consonants and vowels play a major role in the grammar.

Our results showed that both temporal and non-temporal acoustic cues contributed to the phonetic implementation of the phonological contrast. In terms of temporal cues, the duration of the medial geminate fricative was twice as long as that of the singleton; preceding vowel duration showed shortening only when preceding a geminate fricative in the preceding long vowel context (but not in the short vowel context); and the following vowel showed a direct correlation between gemination and lengthening, which seems to suggest longer syllable duration in the geminate categories (Fig. 2). In terms of non-temporal acoustic cues, there were systematic differences linked to the contrast, with geminate environments showing higher $f0$ at the offset of the preceding vowel and at the onset of the medial fricative followed by a significant decrease from the midpoint of medial fricative up to the end of the following vowel (Fig. 4); moreover, a higher centroid (M1) and dynamic amplitude (A_d) and a lower standard deviation (M2) were obtained [Fig. 5(a)]. Our examination of the acoustic space of fricative consonants [Fig. 5(b)], along with dynamic amplitude (A_d) changes [Fig. 5(c)], revealed novel results which suggest systematic qualitative differences linked to the geminate fricatives that are not solely conditioned by phonological place of articulation. Both voiced and voiceless geminates had a voiced part that was proportionately smaller relative to the length of the entire fricative than in singletons (Fig. 6). Surrounding vowels showed centralization (lower F_2) at the midpoint and offset of the preceding long vowel only (Fig. 7). And finally, gemination affected the voice quality of surrounding vowels, showing breathy phonation mainly through the *H_1–A_1 and *H_1–A_2 metrics (Fig. 8). This result seems to be compatible with geminate fricatives being realized with more friction/aspiration (Local and Simpson, 1988), and being associated with breathy phonation as a consequence of [+tense] or fortis articulation (Jessen, 2001).

Our results on non-temporal acoustic cues reveal the differences in articulatory strength between singleton and

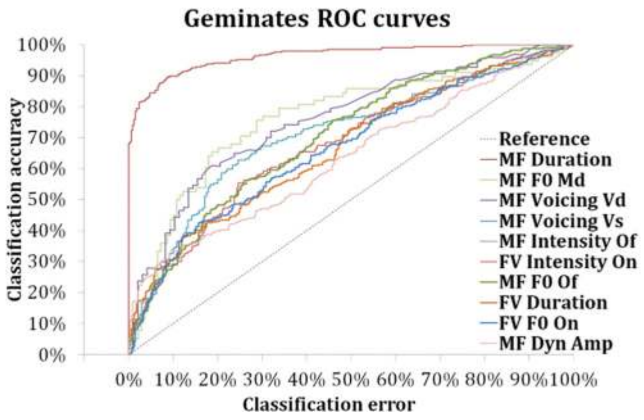


FIG. 9. (Color online) ROC curves of ten best classification models of the geminate category (PV = preceding vowel, MF = medial fricative, FV = following vowel, Vd = voiced, Vs = voiceless, On = onset, Md = midpoint, Of = offset, Dyn Amp = dynamic amplitude A_d).

geminate consonants (Kohler, 1984) whereby geminate fricatives display compatible acoustic cues with those available for [+tense] consonants (Jessen, 2001). This leaves open the question regarding whether these non-temporal acoustic cues are the result of a mechanical effect of lengthening, suggesting that they operate at the phonetic level, or whether the phonological targets for singleton/geminate consonants in LA are fundamentally different, and should be thought of as lenis/fortis pairs. On the one hand, the effect of gemination can be witnessed in the preceding vowel, with higher *f*0 and higher *H_1–A_1 and *H_1–A_2 when the preceding vowel is long. This has been reported for Italian, French and Swedish (see Jessen, 2001, p. 279) and potentially suggests that different gestural targets are in place in the geminate context, and these are implemented in the surrounding vowels. On the other hand, vowel quality either remains unaffected by gemination as in the short vowel context or shows the opposite pattern to that expected, as in the centralization of the long vowel. Moreover, the following vowel mostly shows the influence of the post-stress lengthening in geminate environments compared with singleton ones, leaving the possibility that any changes in the preceding vowel are due to gestural timing and the result of time available for both vowel and consonant to reach their target. This interpretation becomes more likely when one also considers the results from discriminant analysis classification rates and ROC curves, which enabled us to conclude that the duration of the medial fricative is the main contributor to the singleton vs geminate contrast with nearly 90% classification rates and an AUC of 96%; non-temporal acoustic cues contributed to the contrast with much lower classification and AUC rates (see Fig. 9 and Table III). These results suggest that the singleton vs geminate contrast in LA is mainly temporal, with secondary acoustic cues leading to a lax vs tense (lenis vs fortis) distinction.

The primacy of the temporal domain for geminates in LA is explored elsewhere (Khattab and Al-Tamimi, 2014), where we have shown that a moraic account best explains the durational implementation of consonants and preceding vowels in singleton and geminate environments in LA. Moraic segments show greater durational stability than non-moraic ones (Ham, 2001), and this applies to both geminate consonants and their surrounding vowels, which show no temporal compensation unless the resulting syllable is trimoraic (in the context of a long vowel followed by a geminate). A survey on the role of a tense/lax distinction in stops across languages (Jessen, 2001) shows a correlation between languages that tend to show a [+tense] phonological contrast in stops and the presence of aspirated stops in these languages, the lack of voiced geminates and compensatory vowel shortening before phonologically long consonants; these timing patterns exhibit prosodic conditioning which favors post-stress geminates and may be considered the driving factor for gemination in these languages. LA, on the other hand, shows a different prosodic profile: stops are generally unaspirated, short vowels do not shorten before geminate consonants, and gemination can take place both pre- and post-stress. The durational ranges for the singleton and geminate consonants are clearly demarcated, making them less likely to require other cues to the contrast from vowel or consonant quality. This suggests that the contrast is more fundamentally based on stable temporal grounds and that any weakening or strengthening of the consonant may be minimal and due to the phonetic implementation of length.

Future work is needed on non-lexical geminates in order to tease apart any fundamental differences between phonological and phonetic lengthening. Within lexical geminates, a closer look at grammatical categories is required in order to explore any differences between morpheme internal geminates, as found in nouns, and cross-morpheme boundaries which can be seen as the result of two consonants (Hayes, 1989; Local and Simpson, 1988). Work is also currently under way to examine other consonant categories (stops, nasals, and laterals). As Payne (2006) notes, the relatively smaller effect of non-temporal as opposed to temporal cues that was witnessed in fricatives may be due to the fact that there is less room for changes in this crowded class of sounds; for instance, any retraction or palatalization of an alveolar fricative may lead it to become too similar to a post-alveolar fricative. Differences in tongue shape and place of articulation between singleton and geminate consonants have been reported elsewhere for stops and laterals (e.g., Local and Simpson, 1988; Payne, 2006) and would require a different interpretation regarding the phonological targets for each. These are currently being examined for LA.

ACKNOWLEDGMENTS

The authors would like to thank James F. Lynch, Christine Shadle and two anonymous reviewers for their invaluable comments, Paul Boersma for advice on computing time-averaged power spectra, and Paul Foulkes for feedback on an earlier draft of this paper. This research was supported by an ESRC First Grant (RES-061-23-0053) and a British Academy Overseas Conference Grant 2011.

APPENDIX

List of words used in this study with the IPA transcription and the meaning, according to phoneme and syllable structure (Table IV).

TABLE IV. Word list in IPA followed by gloss (*in italic*) as a function of syllable structure and phoneme.

	CVCVC	CVCːVC	CVːCVC	CVːCːVC
/f/	/dafaʕ/ *he paid*	/dafːaʕ/ *he made someone pay*	/daːfaʕ/ *he defended*	/ħaːfːa/ *having rubbed (fem.)*
	/kafal/ *he guaranteed*	/kafːal/ *he delegated*	/kaːfal/ *he vouched*	/xaːfːe/ *having reduced (fem.)*
	/nafaʕ/ *he benefited*	/nafːaʕ/ *he benefited others*	/naːfeʕ/ *he beneficial*	—
/s/	/ɣasal/ *he washed*	/ɣasːal/ *he washed repeatedly*	/raːsal/ *he corresponded*	—
	/ħasab/ *he counted*	/ħasːah/ *he touched her*	/ħaːsab/ *he charged*	/ħaːsːe/ *feeling (fem.)*
	/nasaf/ *he blew up*	/nasːa/ *he made someone forget*	/naːse/ *having forgotten*	—
/z/	/ʕazam/ *he invited*	/ʕazːab/ *he tortured*	/ʕaːzeb/ *bachelor*	/ʕaːzːe/ *being endearing (fem.)*
	/kazab/ *he lied*	/kazːab/ *he lied repeatedly*	/kaːzeb/ *liar*	/kaːzːe/ *gritting her teeth*
	/nazaʕ/ *he spoiled*	/nazːal/ *he lowered*	/naːzel/ *coming down*	/haːzːe/ *shaking (fem.)*
/sˤ/	/ʔasˤam/ *he divided*	/ʔasˤːam/ *he partitioned*	/ʔaːsˤam/ *he shared*	/ʔaːsˤːa/ *having cut (fem.)*
	/ħasˤal/ *it happened*	/ħasˤːal/ *he acquired*	/ħaːsˤal/ *he recuperated*	/ɣaːsˤːa/ *having choked (fem.)*
	/wasˤaf/ *he described*	/wasˤːa/ *he ordered*	/waːsˤef/ *having described*	/xaːsˤːa/ *private (fem.)*
/ʃ/	/baʃar/ *human beings*	/baʃːar/ *he brought good news*	/baːʃar/ *he started*	/ɣaːʃːe/ *having cheated (fem.)*
	/kaʃaf/ *he revealed*	/kaʃːaf/ *he uncovered*	/kaːʃef/ *having uncovered*	/kaːʃːe/ *having frowned (fem.)*
	/naʃar/ *he spread*	/naʃːar/ *he put the laundry out*	/naːʃel/ *having stolen*	/naːʃːe/ *becoming damp (fem.)*
/ʒ/	/ʕaʒan/ *he kneaded*	/ʕaʒːan/ *he kneaded repeatedly*	/ʕaːʒan/ *having kneaded*	/dˤaːʒːe/ *being bored (fem.)*
	/faʒar/ *he exploded*	/faʒːar/ *he caused an explosion*	/faːʒer/ *loud mouthed*	—
	/haʒar/ *he left*	/haʒːar/ *he displaced*	/haːjar/ *he emigrated*	—
/x/	/ʔaxad/ *he took*	/ʔaxːar/ *he delayed*	/ʔaːxer/ *last one*	/naːxːa/ *bending over (fem.)*
	/faxat/ *he pierced*	/faxːat/ *he pierced repeatedly*	/faːxet/ *having pierced*	—
	/saxan/ *he fell ill*	/raxːa/ *he loosened*	/raːxe/ *having relaxed*	/ʃaːxːa/ *having peed (fem.)*
/ɣ/	/laɣam/ *he rigged*	/laɣːam/ *he rigged up*	—	—
	/ʃaɣal/ *he caused concern*	/ʃagːal/ *he employed*	—	—
/ħ/	/faħas/ *he checked*	/waħːad/ *he unified*	—	—
	/laħam/ *he welded*	/laħːan/ *he composed*	—	—
	/saħab/ *he drew*	/saħːab/ *he made someone withdraw*	—	—
/h/	/bahar/ *he sailed*	/bahːar/ *he spiced up*	—	—
	/rahan/ *he bet*	/sahːar/ *he kept someone up late*	—	—
	/sahar/ *he stayed up late*	/sahːal/ *he facilitated*	—	—

Abramson, A. S. (**1999**). "Fundamental frequency as a cue to word-initial consonant length: Pattani Malay," in *Proceedings of the 14th ICPhS*, San Francisco, pp. 591–594.

Al-Ani, S. H. (**1970**). *Arabic Phonology: An Acoustical and Physiological Investigation* (The Hague, Mouton).

Alghamdi, M. (**1998**). "A spectrographic analysis of Arabic vowels: A cross-dialect study," J. King Saud University: Arts **10**(1), 3–24, available online at http://repository.ksu.edu.sa/jspui/handle/123456789/372.

Al-Tamimi, J. (**2007**). "Static and dynamic cues in vowel production: A cross dialectal study in Jordanian and Moroccan Arabic," in *Proceedings of the 16th ICPhS*, Saarbrücken, Germany, pp. 541–544.

Al-Tamimi, J., and Khattab, G. (**2011**). "Multiple cues for the singleton-geminate contrast in Lebanese Arabic: Acoustic investigation of stops and fricatives," in *Proceedings of the 17th ICPhS*, Hong Kong, China, pp. 212–215.

Arvaniti, A., and Tserdanelis, G. (**2000**). "On the phonetics of geminates: Evidence from Cypriot Greek," in *Proceedings of ICSLP*, Beijing, China, pp. 559–562.

Boersma, P., and Weenink, D. (**2009**). "Praat. Doing phonetics by computer. v. 5.1.0 [computer program]," http://www.praat.org/ (Last viewed July 22, 2009).

Broselow, E., Chen, S., and Huffman, M. (**1997**). "Syllable weight: Convergence of phonology and phonetics," Phonology **14**, 47–82.

Catford, J. C. (**1977**). *Fundamental Problems in Phonetics* (Indiana University Press, Bloomington).

Cohen, J. (**1988**). *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. (Lawrence Erlbaum Associates, Hillsdale, NJ).

Cohn, A. (**2003**). "Phonological structure and phonetic duration: The role of the mora," Working Papers of the Cornell Phonetics Laboratory 15, pp. 69–100.

Davis, S. (**2011**). "Geminates," in *The Blackwell Companion to Phonology*, edited by M. van Oostendorp, C. J. Ewen, E. Hume, and K. Rice (Wiley-Blackwell, Malden, MA), Vol. 2, 837–859.

DiCanio, C. T. (**2012**). "The phonetics of fortis and lenis consonants in Itunyoso Trique," Int. J. Am. Ling. **78**(2), 239–272.

Esposito, A., and di Benedetto, M. G. (**1999**). "Acoustical and perceptual study of gemination in Italian stops," J. Acoust. Soc. Am. **106**(4), 2051–2062.

Farinas, J., Rouas, J. L., Pellegrino, F., and André-Obrecht, R. (**2005**). "Extraction automatique de paramètres prosodiques pour l'identification automatique des langues" ("Automatic extraction of prosodic features for automatic language identification"), Traitement Sign. **22**(2), 81–97, available online at http://hdl.handle.net/2042/2443.

Forrest, K., Weismer, G., Milenkovic, P., and Dougall, R. N. (**1988**). "Statistical analysis of word-initial voiceless obstruents: Preliminary data," J. Acoust. Soc. Am. **84**(1), 115–123.

Ham, W. (**2001**). *Phonetic and Phonological Aspects of Geminate Timing* (Routledge, New York).

Hansen, B. B. (**2004**). "Production of Persian geminate stops: Effects of varying speaking rate," in *Proceedings of the 2003 Texas Linguistics Society Conference*, edited by A. Agwuele, W. Warren, and S.-H. Park (Cascadilla Proceedings Project, Somerville, MA), pp. 86–95.

Hassan, Z. M. (**2003**). "Temporal compensation between vowel and consonant in Swedish and Arabic in sequences of CV:C and CVC: and the word overall duration," in *Proceedings of PHONUM*, Sweden, Vol. 9, pp. 45–48.

Hayes, B. (**1989**). "Compensatory lengthening in moraic phonology," Ling. Inq. **20**, 253–306, available online at http://www.jstor.org/stable/4178626.

Hirst, D. (**2011**). "The analysis by synthesis of speech melody: From data to models," J. Speech Sci. **1**, 55–83, available online at http://www.journalofspeechsciences.org/index.php/journalofspeechsciences/article/viewArticle/21.

Idemaru, K., and Guion, S. G. (**2008**). "Acoustic covariants of length contrast in Japanese stops," J. Int. Phon. Assoc. **38**, 167–186.

Iseli, M., Shue, Y.-L., and Alwan, A. (**2007**). "Age, sex, and vowel dependencies of acoustic measures related to the voice source," J. Acoust. Soc. Am. **121**(4), 2283–2295.

Jaeger, J. J. (**1983**). "The fortis/lenis question: Evidence from Zapotec and Jawoñ," J. Phon. **11**, 177–189.

Jakobson, R., Fant, G. M., and Halle, M. (**1976**). *Preliminaries to Speech Analysis: The Distinctive Features and their Correlates* (MIT Press, Cambridge, MA).

Jessen, M. (**2001**). "Phonetic implementation of the distinctive auditory features [voice] and [tense] in stop consonants," in *Distinctive Feature Theory*, edited by T. A. Hall (Mouton de Gruyter, Berlin), pp. 237–294.

Jesus, L. M. T. de, and Shadle, C. H. (**2002**). "A parametric study of the spectral characteristics of European Portuguese fricatives," J. Phon. **30**, 437–464.

Jongman, A., Wayland, R., and Wong, S. (**2000**). "Acoustic charactersitics of English fricatives," J. Acoust. Soc. Am. **108**, 1252–1263.

Khattab, G. (**2007**). "A phonetic study of gemination in Lebanese Arabic," in *Proceedings of the 16th ICPhS*, Saarbrücken, Germany, pp. 153–158.

Khattab, G., and Al-Tamimi, J. (**2013**). "The role of geminates in shaping early word patterns by Lebanese-Arabic speaking children," in *The Emergence of Phonology: Whole-word Approaches and Cross-linguistic Evidence*, edited by M. Vihman and T. Keren-Portnoy (Cambridge University Press, Cambridge), pp. 374–414.

Khattab, G., and Al-Tamimi, J. (**2014**). "Geminate timing in Lebanese Arabic: The relationship between phonetic timing and phonological structure," Lab. Phon. **5**(2), 231–269.

Kohler, K. (**1984**). "The phonetic explanation in phonology: The feature fortis/lenis," Phonetica **41**, 150–174.

Kollmeier, B., Brand, T., and Meyer, B. (**2008**). "Perception of speech and sound," in *Springer Handbook of Speech Processing* (Springer-Verlag, Berlin), pp. 61–82.

Lahiri, A., and Hankamer, J. (**1988**). "The timing of geminate consonants," J. Phon. **16**, 327–338.

Li, F., Edwards, J., and Beckman, M. E. (**2009**). "Contrast and covert contrast: The phonetic development of voiceless sibilant fricatives in English and Japanese toddlers," J. Phon. **37**, 111–124.

Lindblom, B. (**1990**). "Explaining phonetic variation: A sketch of the H and H theory," in *Speech Production and Speech Modelling*, edited by W. J. Hardcastle and A. Marchal (Kluwer Academic Publishers, Dordrecht, the Netherlands), 403–439.

Local, J., and Simpson, A. (**1988**). "The domain of gemination in Malayalam," in *Prosodic Analysis and Asian Linguistics: To Honour R. Sprigg*, edited by D. Bradley, E. J. A. Henderson, and M. Mazaudon, Pacific Linguistics Series C-104, Australian National University, Research School of Pacific Studies, Department of Linguistics, pp. 33–42.

Local, J., and Simpson, A. (**1999**). "Phonetic implementation of geminates in Malayalam nouns," in *Proceedings of the 14th ICPhS*, San Francisco, CA, pp. 595–598.

Maniwa, K., Jongman, A., and Wade, T. (**2009**). "Acoustic characteristics of clearly spoken English fricatives," J. Acoust. Soc. Am. **125**(6), 3962–3973.

McKay, G. (**1980**). "Medial stop gemination in Rembarrnga: A spectrographic study," J. Phon. **8**, 343–352.

Nakai, S., Kunnari, S., Turk, A., Suomi, K., and Ylitalo, R. (**2009**). "Utterance-final lengthening and quantity in Northern Finnish," J. Phon. **37**(1), 29–45.

Nasr, R. T. (**1960**). "Phonemic length in Lebanese Arabic," Phonetica **5**, 209–211.

Nellis, D. G., and Hollenbach, B. E. (**1980**). "Fortis versus lenis in Cajonos Zapotec phonology," Int. J. Amer. Ling. **46**, 92–105.

Payne, E. M. (**2006**). "Non-durational indices in Italian geminate consonants," J. Int. Phon. Assoc. **36**(1), 83–95.

Ridouane, R. (**2007**). "Gemination in Tashlhiyt Berber: An acoustic and articulatory study," J. Int. Phon. Assoc. **37**(2), 119–142.

Shadle, C. H. (**2012**). "On the acoustics and aerodynamics of fricatives," in *The Oxford Handbook of Laboratory Phonology*, edited by Abigail C. Cohn, Cécile Fougeron, and Marie Huffman (Oxford University Press, Oxford), pp. 511–526.

Stevens, K. N. (**1998**). *Acoustic Phonetics* (MIT Press, Cambridge, MA).

Swets, J. A., Dawes, R. M., and Monahan, J. (**2000**). "Better decisions through science," Sci. Am. **283**(4), 82–87.

Tserdanelis, G., and Arvaniti, A. (**2001**). "The acoustic characteristics of geminate consonants in Cypriot Greek," in *Proceedings of the Fourth International Conference on Greek Linguistics*, Thessaloniki, Greece (University Studio Press S.A., Thessaloniki, Greece), pp. 29–36.

Watson, J. (**2007**). "Syllabification patterns in Arabic dialects: Long segments and mora sharing," Phonology **24**(2), 335–356.