Source characteristics of voiceless dorsal fricatives

Charles Redmon[a] and Allard Jongman
Department of Linguistics, University of Kansas, 1541 Lilac Lane, Lawrence, Kansas 66045, USA

(Received 1 February 2018; revised 8 May 2018; accepted 18 June 2018; published online 13 July 2018)

Aerodynamic and acoustic data on voiceless dorsal fricatives [x/χ] in Arabic, Persian, and Spanish were recorded to measure the extent to which such productions involve trilling of the uvula, thus exhibiting a sound source which, contrary to assumptions for voiceless fricatives, is *mixed* rather than *aperiodic*. Oscillation in airflow indicative of uvular vibration was present more often than not in Arabic (63%) and Persian (75%), while Spanish dorsal fricatives were more commonly produced with unimodal flow indicative of an aperiodic source. When present, uvular vibration frequencies averaged 68 Hz in Arabic and 67 Hz in Persian. Rates of uvular vibration were highly variable, however, and ranged between 40 and 116 Hz, with oscillatory periods averaging 4–5 cycles in duration, with a range of 1–12. The effect of these source characteristics on dorsal fricative acoustics was to significantly skew the spectral shape parameters (M1–M4) commonly used to characterize properties of the anterior filter; however, spectral peak frequency was found to be stable to changes in source characteristics, suggesting the occurrence of trilled tokens is not due to velar-uvular allophony, but rather is more fundamental to dorsal fricative production.
 https://doi.org/10.1121/1.5045345

[ZZ] Pages: 242–253

I. INTRODUCTION

A defining characteristic of voiceless fricative consonants is the dominant presence of noise in the radiated acoustic signal, the source of that noise being turbulence in airflow generated at a constriction in the vocal tract which is too narrow to allow laminar flow. The implications of this narrow-constriction definition for the acoustics of fricatives are further elaborated in Stevens (1998, pp. 176 and 379), where a secondary glottal abduction gesture is identified which has the consequence of amplifying the noise source at the supralaryngeal constriction and effectively decoupling the anterior and posterior cavities. As a result, the frequency characteristics of the radiated spectrum become primarily a function of the spectral properties *downstream* of the constriction.

For much of the research on the fundamental acoustic parameters of fricative consonants, and the production characteristics underlying those parameters, the above definition and its corollaries in Fant (1960), Shadle (1985), Stevens (1998), and others, holds. Voiceless fricatives are in large part defined by the amplitude of the sound source and the resonance properties of the anterior cavity in the vocal tract, with the resulting acoustic parameterization successfully applied to the modeling of fricative contrasts in both acoustic (Forrest *et al.*, 1988; Jongman *et al.*, 2000) and perceptual (Behrens and Blumstein, 1988; Hedrick and Ohde, 1993; McMurray and Jongman, 2011) domains. However, velar and uvular fricatives, henceforth referred to collectively as *dorsal fricatives* and transcribed as /X/, pose a problem for such models in that the passive articulator potentially includes a mobile structure in the uvula.[1]

The fact that the uvula is mobile, unlike other passive articulators in the vocal tract, such as the hard palate (ç), alveolar ridge (s, ʃ), and teeth (θ), means the combination of high-velocity airflow and a narrow constriction can introduce sufficient conditions for the Bernoulli force to induce vibration of the uvula (Solé, 1998; Yeou and Maeda, 2011). This vibration disrupts the expectation in voiceless fricatives of a fully aperiodic source by introducing a periodic component which, when combined with the noise generated at the constriction, results in a *mixed* source signal. And while the occurrence of uvular vibration during voiceless dorsal fricative production has been reported in previous studies (Fant, 1960; Shosted and Chikovani, 2006; Shosted, 2008b; Yeou and Maeda, 2011), as well as in studies on allophonic variation in rhotic production (see Barry, 1997; Sebregts, 2015; Solé, 1998, among others), this phenomenon has not been directly studied to date.

The present study addresses three primary questions related to characteristics of the sound source in dorsal fricative production: (1) how common are mixed-source productions of voiceless dorsal fricatives, where the periodic component is due to uvular vibration, both within and across speakers; (2) in such mixed-source tokens, what are the essential frequency, amplitude, and timing characteristics of the periodic component; and (3) what effect does the presence of uvular vibration have on the radiated acoustic spectrum, and, when present, to what degree are certain acoustic parameters sensitive to the prominence of that periodic component.

A. Fricative acoustics and source-filter assumptions

The acoustics of voiceless fricatives have been modeled within source-filter theory as a function of an aperiodic sound source generated either at the point of constriction or at an obstacle upon which the turbulent jet

[a]Electronic mail: redmon@ku.edu

 0001-4966/2018/144(1)/242/12/$30.00

impinges (Alwan, 1986; Fant, 1960; Shadle, 1985).[2] That their voiced counterparts require mixed-source models (i.e., an aperiodic source at the constriction and a periodic source at the larynx) to adequately represent the temporal and spectral characteristics of the acoustic output was acknowledged early on (Fant, 1960) and has been modeled thoroughly more recently in Jackson (2000). However, acoustic models of mixed-source speech sounds have generally been limited to such cases of voicing during frication/plosion, and noise at the glottis in non-modal phonation. Were the uvula to vibrate during production of /X/, two disturbances in flow would be present: the turbulence from the narrow constriction between the tongue dorsum and the soft palate, and broad oscillations in flow from periodic contact between the uvula and the tongue.

Existing source-filter theoretic models of dorsal fricatives, while not accounting for the potential presence of a periodic component due to uvular vibration, do provide important empirical estimates for the acoustics of what we consider the baseline production: that with an aperiodic source. Shadle (1985) modeled velar fricatives as mechanical models with a sound source generated at both the point of constriction and a surface (the hard palate) downstream, where front cavity lengths of 4 and 6 cm were used to model the resonance characteristics of what Shadle refers to as *palatal-velar* and *dorso-velar* places of articulation for /x/.

At a distance of 4 cm from the lips, a broad peak between 1.5 and 2 kHz was generated in the spectrum, while increasing the front cavity length by 2 cm reduced the frequency of this peak by approximately 0.5 kHz. These model results are consistent with earlier acoustic studies which noted that the spectra of velar and uvular fricatives exhibited an overall concentration of energy in the lower frequencies, with a global spectral peak generally below 2 kHz (Jassem, 1965; Strevens, 1960). Extending this framework to more posterior places of articulation, Alwan (1986) modeled uvular and pharyngeal fricatives with synthetic tube models and Arabic speech data as a baseline. The uvular model in Alwan (1986) critically differs from Shadle's models in manipulating constriction length, l_c, which combined with the cross-sectional constriction area determines the Helmholtz resonance, which generally corresponds to $F1$ for uvulars and $F2$ for pharyngeals.

Computed frequencies of $F1$ and $F2$ for the uvular model were 483–582 Hz and 1232–1255 Hz, respectively, for a constriction length of 1 cm (equivalent to Shadle's model), and 447–542 Hz and 1379–1386 Hz, respectively, for $l_c = 2$ cm. Compared to productions of uvular fricatives by four male speakers of Arabic, spectral peak frequencies were found to reflect primarily $F2$, with the computed values of $F2$ within the range of peak frequencies for uvular fricatives in /i/ (1.5–1.8 kHz), /a/ (1.1–1.5 kHz), and /u/ (0.7–1.0 kHz) contexts.

Further, the model predicts dorsal fricatives to be notably "peaked," exhibiting a single spectral prominence which is related to the (compact) feature of velar stop bursts in Jakobson *et al*. (1951) (see also Stevens and Blumstein, 1978). Additionally, models of posterior fricatives in Shadle (1985) and Alwan (1986) differ from their more anterior counterparts /s, ʃ, f, θ/ in showing clear formant structure below $F4$. More recent studies of Swedish (Shosted, 2008b) and Arabic (Al-Khairy, 2005), among others, have replicated these findings. Shosted (2008b), for instance, describes the velar in Swedish as exhibiting a clearer peak and a steeper high-frequency spectral slope than other voiceless fricatives in the language, and Al-Khairy (2005) finds similar results for the Arabic uvular fricative in reporting a relatively high value for the spectral kurtosis parameter.

These models provide a good approximation of the data while assuming an aperiodic pressure source, yet no predominance of low-frequency energy consistent with uvular vibration is predicted by either model. While Jackson (2000) demonstrates that for mixed-source /z/ models the peak resonance is unaffected by the low-frequency harmonic component (and thus mirrors that for /s/, but with additional low-frequency peaks corresponding to $f0$ and its harmonics), it is not clear whether the same independence will be obtained for a mixed source of the type examined in the present study. Given that the periodic component in trilled dorsal fricatives is generated at the same point as the turbulence generating the aperiodic component (i.e., the uvula), we might expect a more complex interaction between source and filter for these sounds. In Sec. III B we address this question by analyzing the response of filter characteristics such as the spectral peak frequency ($F2$) to changes in source type.

The above acoustic data agrees with aerodynamic data from Moroccan Arabic (Yeou and Maeda, 2011; Zeroual, 2003) and Georgian (Shosted and Chikovani, 2006), where mean airflow in dorsal fricatives is between 400 and 700 ml/s (in contrast with the 100–300 ml/s range for alveolars), providing a greater volume of air for front cavity resonance, similar to expectations for approximant productions (Stevens, 1971). However, given that some of these dorsal fricatives may have been trilled, as Yeou and Maeda (2011) and Shosted and Chikovani (2006) suggest, the underlying source of this higher flow rate is not definite. On the one hand, the general trend from previous studies is for volume velocity to increase with more posterior places of articulation (e.g., 553 and 799 ml/s for epiglottal and glottal fricatives, respectively, in Zeroual, 2003), results which are compatible with both the generally larger constriction diameters for these sounds and less direct obstruction from downstream obstacles such as the teeth and lips (Flanagan, 1972; Shadle, 1985). On the other hand, spanning this range are values of peak oral airflow reported for voiceless apical trills (660–1340 ml/s; Solé, 1998)—the closest articulatory analogue of trilled dorsal fricatives—averaging between the expected range for voiceless fricatives and voiceless stop bursts. Thus, in experiment 1 we report flow characteristics of individual cycles of the oscillatory portion of mixed-source productions of dorsal fricatives in part to clarify this source of uncertainty in the literature.

B. Cross-linguistic distribution of fricatives and trills

Dorsal fricatives are reported in the phonemic inventories of 153 of the 451 languages in the UPSID database

(Maddieson, 1992). However, of these, only 22 languages exhibit a contrast between velar and uvular fricatives, meaning that in the absence of phonetic contrast, specifications of place of articulation (i.e., the precise point on the palate where the primary constriction is made) require greater certainty in the articulatory and acoustic data, or must be motivated on phonological grounds. Yet, with the exception of a handful of studies, primarily on Arabic (Ghazeli, 1977; Zeroual, 1999), data of this sort are relatively sparse.

In contrast with the wide distribution of dorsal fricatives, dorsal (uvular) trills are cross-linguistically rare. Only four languages in the UPSID database (<1%) are described as having a uvular trill, and in all such cases the trill is voiced. While voiced trills are the cross-linguistic norm, comprising 36% of the languages in UPSID, as compared with <0.3%, this relation is not an articulatory necessity. Solé (1998), for instance, motivates the apical trill voicing contrast in Icelandic on aerodynamic grounds. Further, allophonic voicing alternations in apical trills are reported in a number of languages, including Spanish and Swedish (Colantoni, 2006; Solé, 1998).

For languages with uvular trills, such as German and French, frication is more commonly reported to co-occur with trilling (Solé, 1998), leading in German, for instance, to ongoing debate over the appropriate phonological analysis (Ladefoged and Maddieson, 1996; Schiller, 1999). Further sociophonetic variability in the manner of articulation of /ʁ/ is reported in Dutch (Sebregts, 2015) and Belgian French (Demolin, 2001), and ranges between voiceless trills, voiced/voiceless fricatives, and approximants, with the baseline production in a given system not always evident (Demolin, 2001).

The above observation is critical to the ultimate interpretation of the results to follow, as it raises the question for cases of variation in the source mechanism as to what should be taken as the underlying category/gesture. This paper does not take a position on this question, but rather aims to provide the necessary aerodynamic and acoustic data to inform the debate.

C. Sound systems of Arabic, Persian, and Spanish

Three languages with sound systems containing voiceless dorsal fricatives were chosen for this study as representative of key phonotactic and inventory patterns: Saudi Arabic, Iranian Persian, and Castillian Spanish. Regarding *phonotactics*, Arabic and Persian allow dorsal fricatives in both onset and coda positions, while Spanish is restricted to onset position. This set is also differentiated according to *inventory distribution*. Spanish may be categorized as exhibiting dorsal fricatives in a *peripheral* position within the inventory, as the language lacks any sounds more posterior in the vocal tract (e.g., glottals). Arabic and Persian, on the other hand, have both anterior contrasts, such as /s, ʃ/, and posterior contrasts, /h, ħ/, though /ħ/ is only present in Arabic. Finally, this set is differentiated according to *contrast density*, with Arabic the most dense (seven fricative places), Spanish the least dense (four places), and Persian intermediate at five places.

All three languages exhibit similar triangular vowel systems of five (Spanish) or six (Arabic, Persian) vowels. The corner vowels /i, a, u/ are present in each language, with Arabic completing its set with short/lax counterparts of each corner vowel (Thelwall and Sa'Adeddin, 1990; Watson, 2002), while Persian and Spanish have the mid vowels /e, o/ (Lazard, 1992; Martínez-Celdrán *et al.*, 2003), the sixth vowel in Persian being the short/lax counterpart of /a/. Further, in none of the three languages are dorsal fricatives phonotactically restricted with regard to the vocalic environments in which they occur, though Martínez-Celdrán *et al.* (2003) note that before back vowels /x/ may be retracted to a uvular position in some dialects. The opposing coarticulatory pattern, palatalization of velar fricatives in the context of high-front vowels (a phenomenon characteristic of German; Ladefoged and Maddieson, 1996) is not reported in any of these languages.

In addition to potential velar-uvular variation as a function of vowel context in Spanish (Martínez-Celdrán *et al.*, 2003), descriptions of Arabic and Persian are not entirely consistent with a single place of articulation (POA)—*velar* or *uvular*—for the dorsal fricative. Lazard (1992), for instance, describes the dorsal fricative in Persian as velar /x/, but notes that this sound is articulated "appreciably farther back" than its stop counterpart /k/. Similarly, in Arabic there is some debate as to whether the uvular fricative is in fact phonetically velar, with Cairene Arabic being one example of a dialect said to have velars (Watson, 2002). One of the aims of this study is that in focusing on the sound source, greater clarity will be brought to the question of place of articulation.

II. EXPERIMENT 1: AERODYNAMICS

The goal of this experiment is to determine, by way of oral airflow data, the essential characteristics of the sound source in dorsal fricative production. Namely, we are concerned with the generation of turbulent airflow at a constriction between the tongue dorsum and the soft palate, and whether that flow is unimodal, as predicted from aperiodic source assumptions (Shadle and Scully, 1995). As noted in Sec. I, we have not made a distinction between velar and uvular places of articulation primarily because the literature on the target languages in this study is inconsistent in identifying the passive articulator for these sounds. One conjecture of this study, however, which is related to the indeterminate POA for dorsal fricatives, is that the presence of oscillation in the flow signal outside of the frequency range attributable to voicing entails a uvular place,[3] as vibration of the uvula in such instances is assumed to be the cause of that oscillation, making the uvula the relevant passive articulator in the description of the sound.

In measuring the flow characteristics of voiceless dorsal fricatives, a single overarching question guides the analysis: *is the source fully aperiodic*? The aerodynamic experiment reported below is critical to answering that question because the acoustic signal reflects turbulence noise and radiation losses which may obscure any underlying periodicity for mixed-source sounds.

A. Methods

1. Participants

Four native speakers (2 female, 2 male) of each of the three target languages—Arabic, Persian, and Spanish—were recruited from the University of Kansas student population for participation in the study. Arabic speakers were restricted to those who grew up in Saudi Arabia (i.e., were born in Saudi Arabia and lived there until at least 12 yrs of age), all Persian speakers grew up in Iran, and all Spanish speakers grew up in Spain. The 12 participants were paid volunteers and reported no speech or hearing impairments.[4]

2. Materials

Word lists with three items exhibiting each combination of the voiceless fricatives /s, X/ adjacent to the corner vowels /i, a, u/ in word-initial and word-final position, as well as 14 filler words not composed of the target phoneme sequences, were prepared in each language. While the consonant of focus in this study is /X/, the voiceless alveolar fricative, /s/, was included in the recordings to serve as a reference where unimodal airflow is expected, and on which there is thorough articulatory and aeroacoustic data (Shadle and Scully, 1995).

Each word was elicited in a sentence frame where the pre-target word ended in the low vowel /a/, and repeated in 4 separate randomized blocks for a total of 144 target utterances per participant (2 consonants × 3 vowels × 2 positions × 3 items × 4 repetitions) in the Arabic and Persian groups, and 72 utterances per participant in the Spanish group, yielding 1440 total utterances in the experiment. Spanish word lists had half the number of stimuli, including half as many filler words, due to phonotactic constraints restricting the target consonants to word-initial position. For elicitation, sentences were presented on a computer monitor with a 4 s inter-stimulus interval, while each repetition was presented in a separate randomized block with a 1 min break given between blocks.

3. Recording

Simultaneous acoustic, oral airflow, and nasal airflow signals were recorded in a quiet room in the KU Phonetics & Psycholinguistics Laboratory with a pneumotachograph equipped with separate nose and face masks integrated through a Macquirer 516 multi-channel transducer (Scicon R&D Inc., 2015).[5] All signals were digitized at 11 kHz and 16-bit resolution, with the airflow signals low-pass filtered at 200 Hz, a threshold chosen because it was well below the range of vocal tract resonances expected for dorsal fricatives.[6] For female speakers this threshold meant the fundamental frequency of vocal fold vibration was often filtered out, but since the target fricatives are voiceless this artifact was not critical for the analysis presented below.

4. Analysis

All signals were first annotated in Praat 6.0 (Boersma and Weenink, 2016), with frication intervals then extracted and imported to MATLAB (MathWorks, 2016) for analysis. Nasal airflow was recorded to allow for future study of the impact of velo-pharyngeal leakage on uvular vibration, but is not reported in the present study.

The target fricatives /s, X/ were segmented from the surrounding utterance as follows. Frication onset from the preceding vowel in the utterance, or offset into the following vowel in the target word itself, was defined as the joint occurrence of three features in the acoustic signal: (1) loss of a clear first formant in the spectrogram, (2) loss of periodicity in the waveform associated with vocal fold vibration, and (3) onset of noise in both the waveform and spectrogram.[7]

This procedure is conservative in attributing any periodicity contiguous with vocalic pulses, however noisy, to the vowel itself. Such a procedure was necessary, however, to avoid erroneously attributing (quasi-)periodic flow due to voicing to perturbations at the uvula. Figure 1 shows representative segmentations of acoustic and oral airflow signals.

Oral airflow signals were analyzed for the presence and rate of oscillation by first computing the autocorrelation function of the signal, R_x, and then taking the power spectral density (PSD) of that function.[8] By spectrally decomposing R_x, rather than applying the Fourier transform directly to the flow signal, the effect of the random component on the estimation of the base oscillation frequency is reduced, making this method more robust at handling the phenomenon of focus, i.e., uvular vibration under turbulent flow (Newland, 1984).

This two-step procedure was performed over 60 ms windows shifted in 10 ms steps from onset to offset of frication. Further, global $f0$ minima from acoustic data collected in

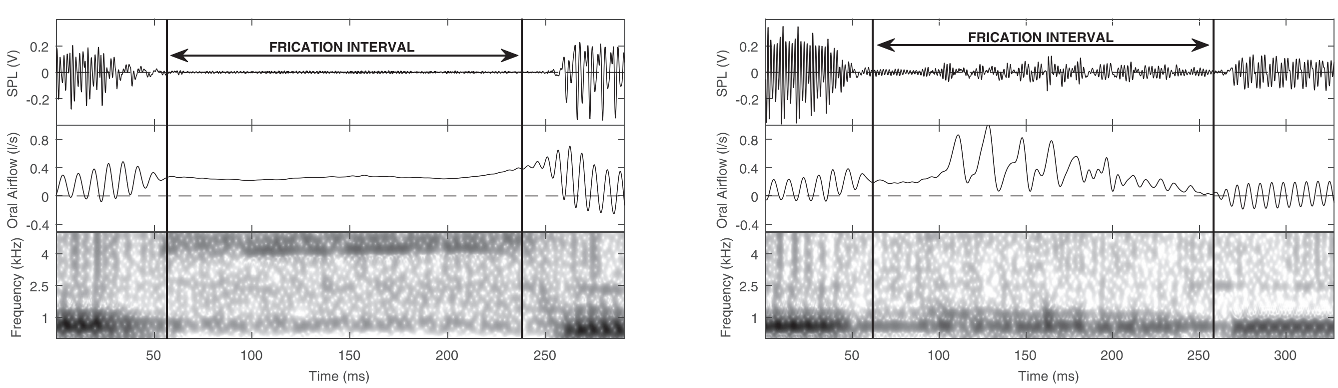

FIG. 1. Sample segmentation of the frication interval in the Persian words /suːt/ "whistle" (left) and /Xuːn/ "blood" (right) by speaker PM01.

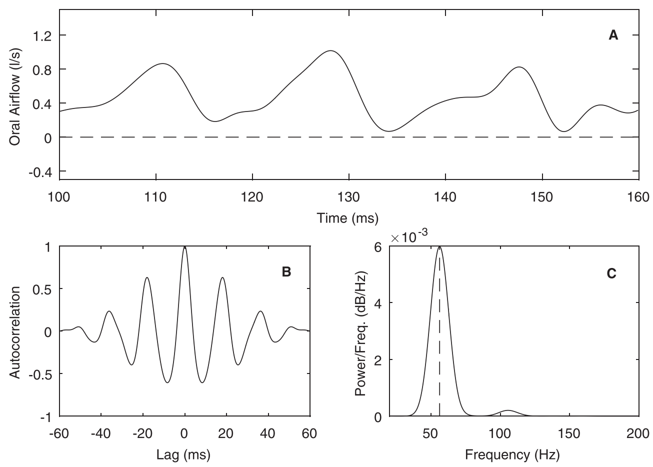

FIG. 2. Demonstration of the quantification of periodicity in oral airflow (A) via the autocorrelation function R_x (B) and the PSD of R_x (C). This sample is taken from the production of /Xuːn/ blood by PM01 in Fig. 1. The computed frequency of oscillation in this token (taken from Panel C) is 55 Hz.

experiment II were computed for each speaker and used as a reference such that any estimated oscillation frequency within a speaker's *f*0 range (as estimated from the global *f*0 minimum in preceding/following vowels across all items) was excluded to ensure such values were not reflective of possible voicing during frication. Figure 2 displays the application of this procedure to a 60 ms window from the /X/ sample in Fig. 1.

From the autocorrelation data, the distribution of mixed-source productions in Arabic, Persian, and Spanish was analyzed by classifying a given fricative interval as *mixed* if at least one analysis frame showed clear periodicity (visible oscillation in airflow and a prominent peak in the PSD of the autocorrelation function) as described above and illustrated in Fig. 2; all other productions were classified as *aperiodic* (Sec. II B 1).[9] Productions from a subset of speakers exhibiting consistent mixed-source dorsal fricatives (100% of all items) were then analyzed for the frequency of oscillation (Sec. II B 2), cycle amplitude, and timing characteristics (Sec. II B 3). All measurements were studied as a function of position (CV, VC) and vowel context (i, a, u) to examine the degree to which coarticulation may modulate characteristics of the sound source; however, separate analyses are conducted for each language, as not only are the items not comparable, but as was noted in Sec. I C, in addition to the phonotactic restriction on consonant position in Spanish, all three languages exhibit moderate differences in vowel systems which make the direct modeling of cross-linguistic differences problematic.

B. Results

1. Distribution of mixed-source productions

The distribution of mixed-source productions of /X/ in Arabic, Persian, and Spanish was analyzed as a function of Language, Speaker, Position, and Vowel Context to determine the extent to which source type is contextually predictable. Table I displays the proportion of tokens exhibiting a mixed source (as defined by the procedure outlined in Sec. II A 4) for each combination of the above factors. Notable trends identifiable in Table I include the more common occurrence of mixed-source productions in Arabic and Persian relative to Spanish (63% and 75%, respectively, as compared with 40%, with three speakers above 50% overall in Persian, two in Arabic, and one in Spanish), and the presence in each language of one speaker with productions which are consistently mixed-source across items; namely, AM02, PF01, and SF01.[10]

TABLE I. Proportions of mixed-source dorsal fricative productions. Word-final (VC) productions were not recorded in Spanish.

		CV			VC		
Language	**Speaker**	**i**	**a**	**u**	**i**	**a**	**u**
Arabic	AF01	0.42	0.58	0.67	0.58	0.08	0.00
	AF02	0.58	0.67	0.75	0.67	0.67	1.00
	AM01	0.42	0.25	0.33	0.82	0.15	0.67
	AM02	1.00	1.00	1.00	1.00	1.00	1.00
Persian	PF01	1.00	1.00	1.00	1.00	1.00	1.00
	PF02	0.92	0.92	0.92	1.00	0.83	0.83
	PM01	0.25	0.50	0.50	0.58	0.42	0.25
	PM02	0.92	0.67	0.83	1.00	0.00	0.67
Spanish	SF01	1.00	1.00	1.00	—	—	—
	SF02	0.17	0.08	0.42	—	—	—
	SM01	0.00	0.50	0.00	—	—	—
	SM02	0.00	0.42	0.17	—	—	—

Modeling source type as a binary outcome, effects of context (Position and Vowel) were analyzed in separate logistic regression models for each language, where Speaker was included as a fixed effect rather than a random intercept in a multilevel model because the assumption of a normal distribution from which the random intercept variance is estimated would be untenable given the small number of speakers. Further, for each language, speakers with 100% mixed-source productions (i.e., AM02, PF01, and SF01) were excluded to improve model stability, and because such cases would not elucidate any Position or Vowel effects. Given the sample size and number of items, Bayesian estimation (Hamiltonian Monte Carlo; Stan Development Team, 2016) was used in this and all subsequent models.[11]

In Arabic, mixed-source productions were equally likely among the three vowel contexts in word-initial position ($CI_{a/i} = [0.4, 3.0]$, $CI_{u/i} = [0.6, 4.3]$, $CI_{a/u} = [0.3, 1.9]$); however, in word-final position mixed-source productions are significantly less common following /a/ than following /i/ ($e^{\beta} = 0.119$, $CI = [0.1, 0.5]$) or /u/ ($e^{\beta} = 0.293$, $CI = [0.1, 0.8]$). No significant difference between /u/ and /i/ vowels in VC position was found ($CI_{u/i} = [0.2, 1.5]$). Equivalently, logistic model results suggest the effect of Position is limited to the /a/ vowel context, though greater uncertainty in the Position estimate meant the predicted lower odds of mixed-source productions in VC position relative to CV did not reach significance ($CI = [0.1, 1.0]$).

In Persian, vowel context effects remain insignificant in word-initial position ($CI_{a/i} = [0.3, 3.3]$, $CI_{u/i} = [0.5, 4.7]$, $CI_{a/u} = [0.2, 2.1]$). In word-final position, both /a/ ($e^{\beta} = 0.061$, $CI = [0.0, 0.2]$), and /u/ ($e^{\beta} = 0.158$, $CI = [0.0, 0.5]$) contexts show a significantly lower odds of exhibiting oscillation in dorsal fricative airflow relative to /i/. No difference between /a/ and /u/ was obtained in VC position ($CI_{a/u} = [0.1, 1.1]$). This interaction between Vowel Context and Position is present also as a significantly lower mixed-source odds in VC than in CV position for the /a/ context ($e^{\beta} = 0.204$, CI = [0.1, 0.7]); this difference, while in the same direction for the vowel /u/, is not significant (CI = [0.1, 1.2]).

In Spanish, mixed-source productions are significantly more likely for consonants preceding /a/ than /i/ ($e^{\beta} = 2.304$, CI = [2.1, 69.8]); however, neither the /u/ > /i/ ($CI_{u/i} = [0.8, 31.3]$) nor the /a/ > /u/ ($CI_{a/u} = [0.7, 6.9]$) relations were significant. This result—particularly evident in the wide intervals around the above estimates—speaks to the substantial variability in the occurrence of uvular trilling in dorsal fricative production in Spanish.

2. Oscillation frequency

Where oscillation in airflow was present, its frequency was measured in two ways. First, for all speakers, the frequency of the peak of the PSD of the autocorrelation function was recorded for each frame where oscillation was present (following the procedure in Sec. II A 4). Multiple frames from a single consonant interval were then combined into a *power-weighted mean frequency* by weighting each frequency by its relative power from the PSD, and taking the weighted average across frames. This weighting has the effect of making the measurement of average frequency for a given item more reflective of stable regions of oscillation (high autocorrelation) than of unstable regions. Given the inherent noise in spectral decomposition of signals from random vibrations, this procedure was held to more reliably recover the base oscillation frequency of that production than would an unweighted mean or median.

Second, for the subset of speakers exhibiting 100% mixed-source productions (AM02, PF01, and SF01), the periods of individual cycles were measured by hand and used to compute more precise frequency values over the consonant interval.[12] From these values, estimates of the *mean cycle frequency*, as well as the range of frequencies in a given interval, were obtained for each item.

Figure 3 displays results of Bayesian linear regression fits to each speaker's data. Separate speaker models were run, as opposed to a single model, because the wide variability in mixed-source distribution by vowel and position meant that in addition to assuming speaker differences in mean frequencies, differences in the error term must be assumed (this latter heterogeneity cannot be modeled in a single-level linear regression).

In general, greater differences between speakers are observed than are within-speaker differences due to position or vowel context. Median oscillation frequencies per speaker, as estimated from the PSD, range from 40 Hz (AM01) to 116 Hz (AF01), with a cross-speaker mean of 75 Hz (68 Hz in Arabic, 67 Hz in Persian, and 90 Hz in Spanish). Of the context effects which were significant in the linear models, the high-front vowel context tended to elicit lower frequencies relative to /a/ and /u/, primarily in word-initial position (AF02: $\beta_{u\text{-}i} = 17.71$, $CI_{u\text{-}i} = [4.4, 31.4]$; AM02: $\beta_{a\text{-}i} = 15.07$, $CI_{a\text{-}i} = [6.3, 23.5]$, $\beta_{u\text{-}i} = 13.14$, $CI_{u\text{-}i} = [4.2, 21.4]$; SF01: $\beta_{a\text{-}i} = 14.09$, $CI_{a\text{-}i} = [0.1, 27.5]$), though this effect was also present for one speaker in word-final position (PM01: $\beta_{a\text{-}i} = 18.0$, $CI_{a\text{-}i} = [2.6, 34.3]$). Word-finally, PM01 also showed a significant difference between /a/ and /u/ ($\beta_{a\text{-}u} = 32.59$, $CI_{a\text{-}u} = [13.5, 53.5]$); however, this effect is not replicated in data from any other speaker. All other comparisons were not significant.

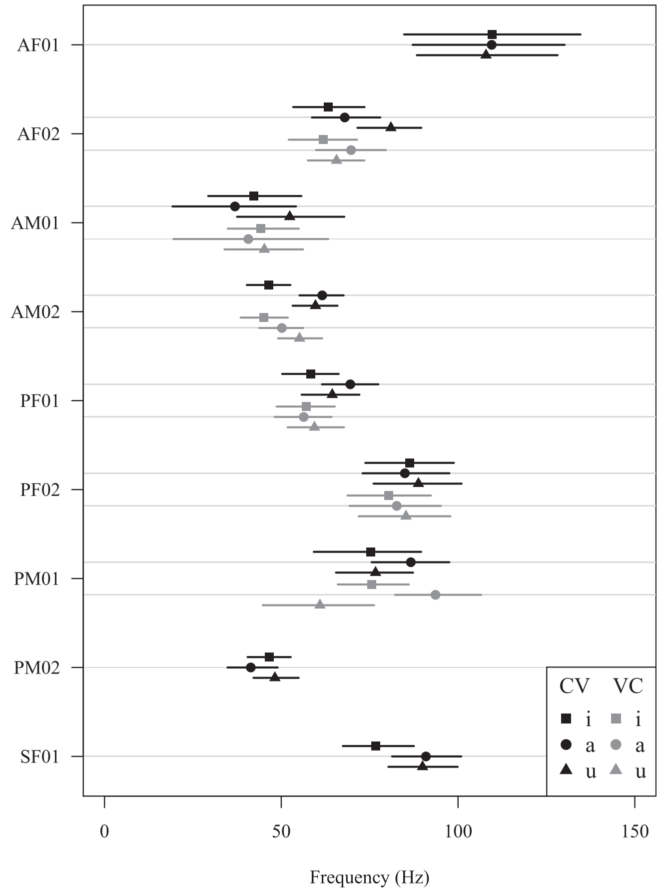

FIG. 3. Power-weighted mean frequency of oscillation in dorsal fricative airflow by speaker (vertical axis), position (CV, VC), and vowel context (i, a, u). Points represent medians of the posterior distribution in a Bayesian regression, with lines spanning the 95% credible interval (HPDI) over that distribution. Speakers SF02, SM01, and SM02 are not shown due to their sparse productions of mixed-source tokens, while AF01 and PM02 estimates are limited to CV position due to sparsity in mixed-source productions in VC position.

For the subset of productions from AM02, PF01, and SF01, where oscillation frequency was determined from the measurement of individual cycles in airflow, mean trill rates per consonant interval correlated significantly with the corresponding estimates from the PSD [$r = 0.89$, $t(177) = 25.5$, $p < 0.001$], with a root-mean-square error (RMSE) of 8.95 Hz. Much of this difference may be attributed to a 23% reduction in between-item variance when frequencies are calculated directly from cycle periods as opposed to estimated from the autocorrelation PSD.

Analyses of context effects on mean oscillation frequencies largely replicated the patterns reported above. Namely, in productions from AM02 and SF01, the rate of uvular vibration was significantly lower for dorsal fricatives preceding /i/ than for those preceding /a/ (AM02: $\beta = 11.07$, CI = [2.3, 19.2]; SF01: $\beta = 14.50$, CI = [4.9, 24.2]) or /u/ (AM02: $\beta = 12.27$, CI = [3.1, 20.6]; SF01: $\beta = 10.42$, CI = [0.6, 20.3]). As before, oscillation frequencies were not found to vary significantly as a function of context in PF01's productions, with the single exception being a significant reduction in frequency in VC position (relative to CV) for /a/-context productions ($\beta = -9.08$, CI = [−18.1, −0.4]). Thus the largely automated, autocorrelation-based method provides a close approximation to hand measurement.

3. Cycle amplitude and timing

Given the large variance in the above estimates of oscillation frequency, and the current lack of available data on uvular vibration in dorsal fricative production more generally, we examined a number of characteristics of the individual cycles comprising the oscillatory flow in the data from AM02, PF01, and SF01. Each cycle was measured by calculating the local maxima and minima over the consonant airflow signal in MATLAB (MathWorks, 2016), and then hand-checking the resulting periods, from which measures of peak-to-peak amplitude, oscillation onset/offset (as a proportion of consonant duration), and duration of oscillation (in number of cycles and in ms), were made. Mean values of these parameters by Speaker, Position, and Vowel Context are shown in Table II.

Mean oscillation frequency was reported in Sec. II B 2; however, we have yet to discuss the volume of airflow expelled during a given trill cycle. Peak-to-peak amplitude was found to average approximately 310 ml/s (364 ml/s for AM02; 261 ml/s for PF01; 301 ml/s for SF01), though values ranged from 67 to 788 ml/s (AM02[136,788]; PF02[67,526]; SF01[80,679]). Mean cycle amplitude largely did not vary by Vowel Context or Position, the one notable exception being that in the high vowel context, uvular vibration was consistently greater in amplitude post-vocalically (VC) than prevocalically (CV), though this effect was only significant in the /i/ context for AM02 ($\beta = 79.69$, CI = [4.5, 155.8]), and the /u/ context for PF01 ($\beta = 94.33$, CI = [13.1, 179.9]). Within a given position, mean cycle amplitude did not vary significantly by vowel context for any of the three speakers.

TABLE II. Mean cycle peak-to-peak amplitude (ml/s), oscillation onset/offset time (normalized as a percentage of duration), and duration of the periodic portion of the consonant (in cycles and ms).

		CV			VC		
		i	a	u	i	a	u
Amplitude	AM02	342	384	341	457	335	334
(ml/s)	PF01	252	216	219	283	273	324
	SF01	385	262	255	—	—	—
Onset	AM02	0.14	0.12	0.13	0.16	0.19	0.15
(norm. *t*)	PF01	0.10	0.11	0.11	0.19	0.25	0.16
	SF01	0.06	0.09	0.07	—	—	—
Offset	AM02	0.80	0.91	0.90	0.90	0.92	0.82
(norm. *t*)	PF01	0.70	0.69	0.75	0.85	0.82	0.81
	SF01	0.80	0.88	0.85	—	—	—
Duration	AM02	3.3	6.3	5.3	4.3	5.1	4.8
(cycles)	PF01	3.3	3.4	4.0	3.3	3.1	3.6
	SF01	5.4	5.8	6.8	—	—	—
Duration	AM02	67	103	87	95	98	81
(ms)	PF01	51	53	61	58	53	56
	SF01	73	61	80	—	—	—

The above values span the range of what has previously been found for voiced and voiceless apical trills (Solé, 2002), though the lack of vowel context effects diverges from Solé's findings on apical trills, where greater flow volume was elicited in the /i/ context than in the /a/ context. Aerodynamic data on voiced uvular trills is needed to provide compatible coarticulatory expectations for trilling at a dorsal POA.

Regarding the timing of uvular vibration, on average, oscillation in flow begins within the first 15% of the consonant interval (mean: 13.6, median: 11.1), and ends in the final 20% (mean: 82.6, median: 86.4). This pattern of symmetric oscillation around the consonant midpoint is consistent with trajectories of changes in constriction area and posterior cavity pressure for both fricatives and trills (Shadle and Scully, 1995; Solé, 2002, 2010), and holds across position and vowel contexts. In a beta regression on relative oscillation timing, the one notable trend that was consistent in productions from both AM02 and PF01, though only significant in PF01, was a later cessation of oscillatory flow in VC than in CV position ($CI_{VC/CV|i} = [1.1, 4.0]$, $CI_a = [1.1, 4.1]$, $CI_u = [1.0, 3.9]$), which may reflect different gestural timing constraints in the two positions.

The duration of such oscillatory intervals, based on the longest contiguous sequence in cases where oscillation begins and ends at multiple points in a given consonant, was found to average between 4 and 5 cycles (range: 1–12), or approximately 72 ms (range: 13–152). With the exception of word-initial productions from AM02, where vowel context effects were observed for duration in milliseconds (/a/ > /i/, $\beta = 34.69$, CI = [7.8, 61.8]) and cycle count (/a/ > /i/, $e^{\beta} = 1.903$, $CI_{a/i} = [1.3, 2.8]$; /u/ > /i/, $e^{\beta} = 1.623$, $CI_{u/i} = [1.1, 2.4]$), productions did not differ in the duration of oscillation by position or vowel context.[14]

C. Discussion

Experiment 1 details a number of key characteristics of dorsal fricative production and uvular vibration under turbulent flow. First, Arabic and Persian were found to exhibit uvular vibration in /X/ productions more often and more consistently across speakers than Spanish. Yet when present, characteristics of this oscillation were quite similar cross-linguistically. Notably, the frequency and time course of oscillation remained fairly constant across vowel contexts /i, a, u/ and position (CV, VC). However, due to the random nature of the flow source, this periodic component proved to be highly unstable, as reflected in the high overall variance in oscillation frequencies, both within and across speakers, relative to previous studies of voiced uvular trills where standard deviations are generally between 3 and 5 Hz (Ladefoged *et al.*, 1977; Shosted, 2008a; Solé, 2002).

It should also be noted that the vibration rates were much higher in the present study than have been reported previously for voiced uvular trills, which generally average 25–35 Hz, and though higher trill frequencies have been reported for voiceless versus voiced apical trills, such differences are generally on the order of 1–2 Hz (Solé, 1998). We should point out that this does not lead to the expectation of a similar difference in vibration rate between voiced and voiceless uvular trills because the mass and tissue properties of the uvula, as well as properties of the constriction location and shape, are quite different from those of the tongue tip. Verhoeven (1994), for instance, reports frequencies of trill variants of the uvular fricative in Dutch at approximately 60 Hz, while voiced uvular trills remained around 26 Hz in frequency. Similar rates may be derived from vertical striation patterns in spectrograms of Russian (Fant, 1960), Belgian French (Demolin, 2001), and Dutch (Sebregts, 2015, p. 64), suggesting two possible sources higher vibration rate: the turbulence inherent to the fricative productions, which may disrupt the uvula's natural vibration frequency and lead to a more complex oscillation pattern, and the lesser muscular tension in the uvula relative to the tongue tip, which may cause it to be less resistant to changes in aerodynamic conditions. Nevertheless, this area deserves further study and controlled measurement of uvular vibration at different volume velocities (U) and Reynolds numbers (Re). For further consideration, sample airflow signals representing different oscillation frequencies from each speaker are provided in the supplementary material.[13]

In Sec. III, results of the acoustic experiment are presented to examine the overall effect of periodicity in the sound source on the radiated acoustics of dorsal fricatives, providing estimates of the degree to which previously reported measures of /X/ spectra in languages like Arabic are dependent on assumptions regarding the nature of the sound source.

III. EXPERIMENT 2: ACOUSTICS

In this experiment we make two predictions regarding the dependence of the radiated output on the characteristics of the source. First, there is the direct mathematical consequence of adding a prominent low-frequency component in the spectrum to the shape of that spectrum; namely, the spectral mean will lower and spectral skewness will increase, and as the two other moments (M2 and M4) are correlated with M1 and M3, we should see concomitant effects on those parameters as well. The second prediction is conditional on what the cause of the production difference between mixed-source and aperiodic-source dorsal fricatives is. If this difference derives from allophonic variation in velar and uvular places of articulation, then a difference in spectral peak frequency by source type should emerge, with mixed-source productions exhibiting lower peak frequencies due to the longer front cavity anterior to the constriction. However, if this difference is rather dependent on aerodynamic characteristics such as the cross-sectional area of the constriction and the rate of airflow, then a source difference could emerge without a difference in POA, and thus yield comparable spectral peak frequencies.

A. Methods

1. Participants and materials

The same participants and stimuli were used in the acoustic experiment, with no change in the item order or method of presentation. All participants were paid volunteers and reported no speech or hearing impairments.

2. Recording

Stimuli were recorded in frame sentences with a head-worn cardioid condenser microphone (Shure SM-35) on a solid-state recorder (Marantz PMD671) in an anechoic chamber at the University of Kansas. The position of the microphone was approximately 4 cm from the side of the participant's mouth. Microphone levels were calibrated to approximately 80% of the input voltage during a practice reading of material not part of the experimental stimuli. Audio signals were sampled at 22.05 kHz with 16 bit resolution.

3. Analysis

All stimuli were annotated, segmented, and analyzed in Praat 6.0 (Boersma and Weenink, 2016). The segmentation procedure followed that outlined in Sec. II A 4 and demonstrated in Fig. 1. Following segmentation, a diagnostic of source type was developed based on the expectation that the introduction of periodicity into the acoustic signal, i.e., the emergence of a low frequency base signal onto which noise is overlaid (as in Fig. 4), would result in a spectrum with an overall negative tilt, similar to that of vowels and resonant consonants. The measurement proposed here as a diagnostic of source type we refer to as the *source-filter ratio* (SFR), which is simply the difference in maximum amplitudes of two spectral regions: 0–200 Hz, which broadly comprises those frequencies influenced by any periodic source, and 0.5–10 kHz, which covers the remainder of the spectrum

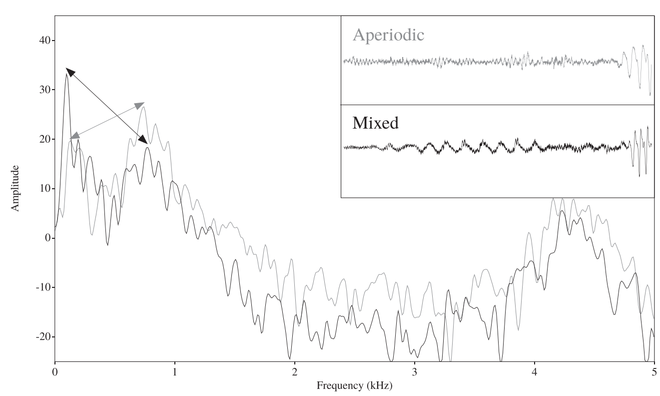

FIG. 4. Spectra from the middle 60 ms of aperiodic (gray) and mixed-source (black) dorsal fricatives in two repetitions of /Xutˤba/ by AF01. Arrows indicate amplitude peaks corresponding to source and filter components used to compute the SFR for each spectrum. Corresponding waveforms for the full consonant and initial periods of the following vowel are shown in the inset.

influenced by resonance characteristics of the vocal tract filter.

Figure 4 illustrates the manner in which this index, the relative difference between the source and filter components of the spectrum, delineates aperiodic and mixed-source dorsal fricatives. This ratio, with its threshold of 0, was chosen for its computational simplicity, and validated in a logistic regression on a subset of the data (20% of all tokens, balanced by speaker and position) that was independently and blindly rated by the authors as either mixed or aperiodic, according to auditory impression and visual inspection of periodicity in the waveform. Inter-rater agreement on the classification was at 88%, with the pooled model yielding a classification boundary (the point of 50% predicted probability of a *mixed* source) at an SFR of 0.5 dB.[15] Thus there was close agreement between our initial threshold based on source-filter-theoretic assumptions (0 dB) and that which may be derived from the data.

In addition to the SFR, which was used to classify fricatives by source type and quantify the relative amplitude of the source signal in mixed tokens, further measures of dorsal fricative acoustics in Arabic, Persian, and Spanish are reported to provide a relation between the present data and that from prior research on these and similar languages, and to quantify the degree to which these parameters are influenced by characteristics of the sound source. Five spectral measures—spectral peak frequency and the four spectral moments (M1–M4) characterizing the overall shape of the spectrum—were computed from a Hamming window over the middle 60 ms of frication. Measurement of these five parameters was made following the procedures in Jongman *et al*. (2000).

B. Results

1. Spectral peak frequency

The analysis of spectral peak frequency as a function of source type addresses the second prediction stated in Sec. III; namely, if the peak frequency varies significantly with the nature of the source we have some evidence that allophony in POA (velar vs uvular) remains a potential explanation for the presence of uvular vibration in dorsal fricative production.

In examining this relationship, variation in spectral peak frequency by source type was modeled both with source type as a derived dichotomous variable ($\text{SFR} \leq 0 = \textit{aperiodic}$, $\text{SFR} > 0 = \textit{mixed}$), and directly as the continuous SFR. Each variable was interacted with Position and Vowel Context, while controlling for speaker mean differences, in a separate linear regression on Spectral Peak Frequency per Language. As in Sec. II B, Bayesian estimation was used to fit all models.

Results did not support the allophonic hypothesis. As illustrated in the top row of Fig. 5, spectral peak frequency generally did not differ by source type or vary strongly as a function of the relative amplitude of the source component. Where significant differences in source type were found, such as in the /u/ context word-initially ($\beta = -406.1$, CI = [−564, −250]) and word-finally ($\beta = 427.4$, CI = [175,677]) in Arabic, and in the /i/ context in Spanish ($\beta = 297.7$, CI = [167,428]), the directionality of the effect was not consistent, nor compatible with predictions, i.e., mixed-source items were not consistently lower in peak frequency than aperiodic items.

Similar results were obtained in the model with the continuous SFR as a predictor. Generally, no relationship with peak frequency was found, and when present (Arabic, CV:

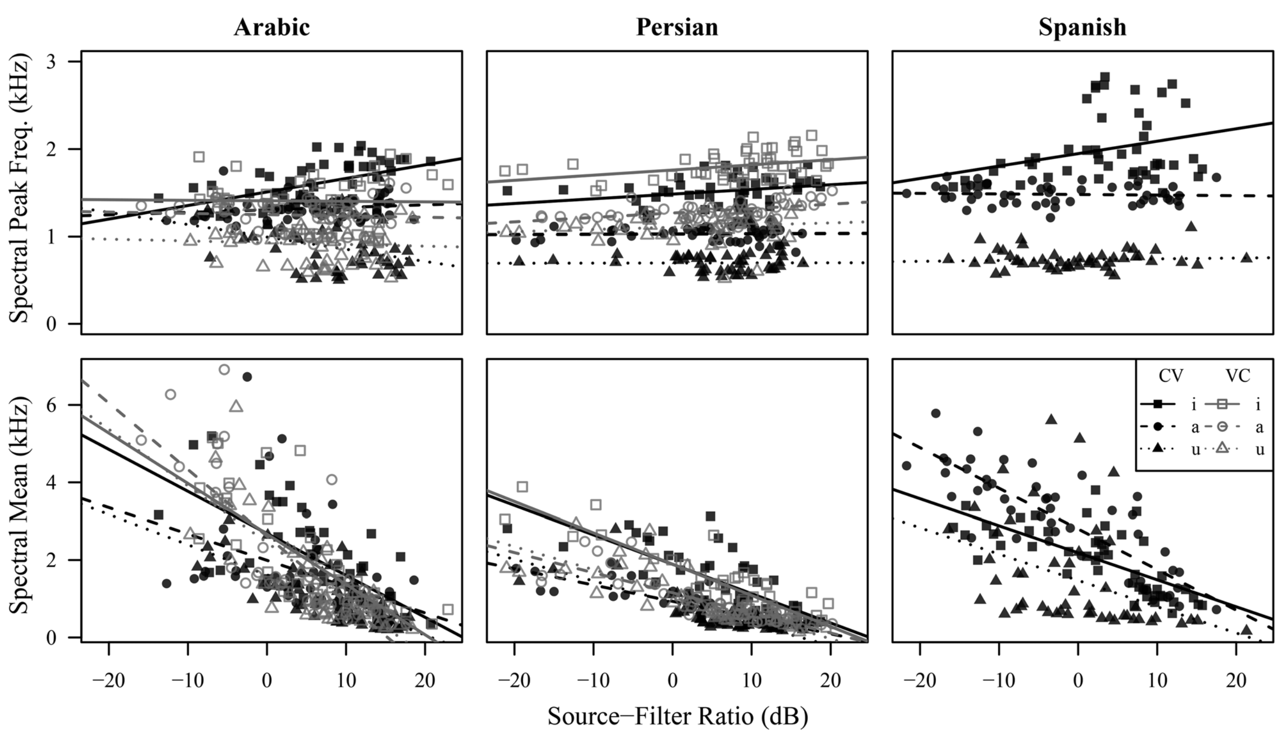

FIG. 5. Relationship between SFR and spectral peak frequency/spectral mean in Arabic, Persian, and Spanish.

$CI_i = [1.5, 16.5]$, $CI_u = [-34.2, -18.5]$; Arabic, VC: $CI_a = [0.9, 26.5]$, $CI_u = [15.2, 43.0]$; Persian, VC: $CI_a = [0.1, 14.7]$; Spanish, CV: $CI_i = [3.7, 18.7]$) the relationship was inconsistent across contexts. In fact, when there was a significant effect of SFR on spectral peak frequency the tendency was for that relationship to be positive, the opposite of what the allophonic hypothesis would predict.

2. Spectral shape

Unlike the spectral peak frequency analysis, characteristics of the spectral shape (M1–M4), as functions of the energy distribution in the entire spectrum, are expected to co-vary with source component characteristics (see Jongman *et al.*, 2000, for comparable effects of voicing on spectral moments). The question then for this section is not whether an effect of source characteristics will be observed, but to what degree will the various spectral moments be affected by changes in the sound source, and how is this covariation modulated by position and vowel context. Table III summarizes the expected values and maximum predicted change in each spectral moment over the range of SFRs comprising the *mixed* source type. The relationship between spectral mean and SFR for each context in each language is shown in the bottom row of Fig. 5.

In Arabic, estimates of the average change in spectral mean with a 1 dB increase in the relative amplitude of the source ranged between −20 and −171 Hz, with a generally greater influence of SFR word-finally ($\beta_i = -89.60$, $CI = [-117, -62.5]$; $\beta_a = -124.0$, $CI = [-171, -77.3]$) than word-initially ($\beta_i = -72.02$, $CI = [-99.0, -44.5]$; $\beta_a = -76.53$, $CI = [-104, -48.4]$), though the opposite relation was observed for /u/ ($\beta_{CV} = -91.69$, $CI = [-120, -62.3]$; $\beta_{VC} = -68.32$, $CI = [-117, -19.7]$). In Persian, consistently steeper slopes between SFR and M1 were observed in the VC position, with relative differences between vowel contexts differing from β_a (-41.68, $CI = [-53.9, -29.4]$) $\leq \beta_u$ (-57.50, $CI = [-72.2, -42.5]$) $\leq \beta_i$ (-72.38, $CI = [-85.4, -58.9]$) in CV, to β_u (-72.87, $CI = [-95.9, -49.2]$) $\leq \beta_i$ (-80.08, $CI = [-92.2, -68.3]$) $\leq \beta_a$ (-80.91, $CI = [-102.4, -59.7]$) in VC. Finally, in Spanish, the largest negative relationship was observed in the /a/ context ($\beta = -79.51$, $CI = [-102, -57.3]$), followed by /i/ ($\beta = -49.47$, $CI = [-77.2, -23.3]$), then /u/ ($\beta = -47.30$, $CI = [-74.4, -19.7]$).

Yet, despite relative differences in the magnitude of the effect according to context, all combinations of Position and Vowel Context in all three languages show significant negative effects of the relative source amplitude (SFR) on the overall mean of the spectrum at greater than a 20 Hz decrease per 1 dB increase. Concomitant effects for the three additional spectral moments (M2–M4) are shown in Table III.

TABLE III. Mean values of spectral moments M1–M4. The predicted change in each parameter over the range of SFR values above zero (i.e., for *mixed-source* items) is shown in parentheses.

		CV			VC		
		i	a	u	i	a	u
M1	Arabic	1.96	1.60	1.02	1.88	1.93	1.52
kHz		(−2.2)	(−1.3)	(−1.4)	(−3.0)	(−3.0)	(−2.7)
	Persian	1.41	0.74	0.57	1.27	0.91	0.96
		(−1.4)	(−0.7)	(−0.9)	(−1.6)	(−1.0)	(−1.0)
	Spanish	2.00	3.05	1.50	—	—	—
		(−1.1)	(−1.8)	(−1.4)	—	—	—
M2	Arabic	2.47	2.18	1.84	2.45	2.29	2.20
kHz		(−0.8)	(−0.2)	(−1.0)	(−1.2)	(−1.2)	(−1.8)
	Persian	1.83	1.15	1.07	1.78	1.31	1.35
		(−0.5)	(−0.4)	(−0.8)	(−0.2)	(−0.3)	(−0.6)
	Spanish	2.18	2.55	1.84	—	—	—
		(+0.1)	(−0.2)	(−1.0)	—	—	—
M3	Arabic	2.32	2.93	4.78	2.29	2.57	3.44
		(+2.2)	(+1.3)	(+5.0)	(+2.8)	(+2.9)	(+4.7)
	Persian	2.88	5.18	7.52	3.13	4.33	4.62
		(+1.9)	(+1.8)	(+4.8)	(+1.6)	(+2.6)	(+2.4)
	Spanish	2.01	1.31	4.27	—	—	—
		(+0.9)	(+1.3)	(+3.6)	—	—	—
M4	Arabic	7.6	11.8	37.2	6.3	9.5	19.7
		(+15)	(+11)	(+70)	(+15)	(+17)	(+47)
	Persian	11.9	38.1	79.9	13.0	26.4	33.1
		(+17)	(+22)	(+79)	(+13)	(+28)	(+33)
	Spanish	4.7	2.1	32.4	—	—	—
		(+5)	(+5)	(+40)	—	—	—

C. Discussion

In experiment 2 we examined the extent to which characteristics of the radiated acoustic signal depend on changes in characteristics of the sound source. A few critical results came out of the analysis above. First, analysis of spectral peak frequency as a function of the type and relative amplitude of the sound source demonstrated that the main resonance of the vocal tract remains constant with changes in source characteristics. This result lends support to the hypothesis that uvular vibration in dorsal fricative production is not a consequence of allophonic variation between velars and uvulars, but rather emerges likely as a complex function of constriction diameter and oral airflow rate.

Second, the ensemble of spectral moments, particularly M1, M3, and M4 (Table III), were highly sensitive to source characteristics, and in some instances (such as the spectral mean) the degree of change associated with the source independent of the filter was on the order of contrastive shifts in place of articulation. For example, in Arabic, a 1 dB increase in the amplitude of the source component relative to that of the filter led to a median reduction of 83 Hz in M1, which, considering the 22.9 dB range of SFR values in mixed-source items, means a 1.9 kHz drop in M1 may result purely from a difference in source characteristics, a value which is well within the range of cross-category differences attributed to place of articulation [e.g., the difference between /χ/ and /ħ/ spectral means reported in Al-Khairy (2005) is 1.1 kHz]. Thus, not only are the source effects on the acoustics predicted to be highly salient, but they also have the potential to be misinterpreted as constituting a feature change that is due to an entirely different mechanism, thus motivating further attention to source characteristics in the study of posterior fricative systems.

IV. GENERAL DISCUSSION

The present study has demonstrated, by way of aerodynamic and acoustic data, that uvular vibration is a pervasive phenomenon in dorsal fricative production, and that the resulting mixed source signal has robust effects on the acoustic characteristics of these sounds. More often than not, aerodynamic and acoustic data indicative of a vibrating uvular source was present in Arabic and Persian. When present, the rate of oscillation in airflow from uvular vibration was on average twice that which has been reported in studies of voiced apical and uvular trills, but also exhibited much greater variability, motivating further study of uvular vibration under turbulent airflow conditions. Most critically for the role of this study within the phonetic literature in studying the acoustic consequences of uvular vibration for the radiated acoustic spectrum, not only were the previous observations of spectral tilt (M3) and peakedness (M4) in dorsal fricative acoustics strongly correlated with the presence and prominence of a periodic component in the spectrum, but all other spectral shape parameters investigated were shown to be highly sensitive to differences in source characteristics. Notably insensitive to changes in the sound source was the spectral peak frequency.

Among the open questions raised by the results above are aspects of the production and perception of dorsal fricatives. On the production end, to adequately model the aerodynamic conditions generating uvular vibration in the mixed-source tokens identified in experiment 1, pharyngeal pressure measurements are needed to study the time course of pressure changes behind the constriction. Further, imaging data are necessary to determine precisely where contact between the uvula and tongue dorsum is being made, and how this contact depends on overall tongue body position, particularly as a function of coarticulation with the surrounding vowel context.

Regarding perception, the finding that the two languages which were consistently produced with trilling were Arabic and Persian, the two languages with consonant inventories containing other posterior fricatives like /h/, suggests that vibration of the uvula during dorsal frication has the potential to serve as a contrast-enhancing feature in perception. Category identification data are therefore needed to determine whether the lack of a periodic component in these sounds, particularly in degraded audio conditions, would cause significant confusions with similar fricative categories like /ħ/ or /h/.

Finally, the present data, particularly that of Arabic and Persian, where uvular vibration is evident in greater than 60% of productions, raises the important phonological question as to whether such sounds are better considered as trills than fricatives. In languages like German, where uvular trills may be attributed historically to both rhotic and fricative origins (Schiller, 1999), such decisions are not without controversy (Ladefoged and Maddieson, 1996). We leave such questions to be answered in the specific phonological contexts of the languages in question, but note that regardless of the position adopted, the evidence above suggests that a thorough account of the relevant acoustics of dorsal fricatives requires analytical considerations from both manner classes.

ACKNOWLEDGMENTS

We would like to thank the Associate Editor and two anonymous reviewers for their helpful feedback, as well as Joan Sereno, Jie Zhang, Anders Löfqvist, Doug Whalen, and the members of the KU Experimental Linguistics Seminar for their input on earlier versions of this work.

[1]The uvula is included as a relevant surface in velar fricative production based on observations in Flanagan (1972), Shadle (1985), and others that /x/ is articulated with a long constriction sometimes extending over the entire soft palate.

[2]The identification and analysis of obstacle sound sources in fricative production is complex and beyond the scope of the present introduction. We refer the reader to Shadle (1990) for further discussion.

[3]See Sec. II A 4 for details on the manner in which flow oscillation is attributed to different articulatory sources.

[4]The sample size in this study, both in terms of speakers and number of languages representing a given typological feature, is understood to be sufficient to provide a window on the phenomenon and motivate further large-sample studies on individual groups, not to directly generalize to either population.

[5]The nasal mask was held in place over the participant's nose via a strap extending around the back of the head, while the oral mask was held in place by the participant via a rod attached to the back of the transducer. While the nasal mask always maintained a tight seal, the oral mask often needed to be adjusted to fit the participant (e.g.,, for shorter faces the mask occasionally needed to be angled downward to maintain a seal). In all cases the seal of the mask was checked by the researcher prior to each block of recording.

[6]This low-pass filtering approach follows that of Scully (1990), Solé (2002), and others, though with a higher filtering threshold (Scully and Solé use 50 Hz cutoffs) because preliminary recordings suggested the oscillations from uvular vibration could be as high as 125 Hz. The specific filter used was a one-sided Hann filter that was 6 dB down at 200 Hz and had a 40 Hz range (180–220) between pass and stop values.

[7]Inter-rater agreement (between C.R. and A.J.) on segmentation of a representative subset of the data (5% of items) showed a median absolute deviation (MAD) in CV boundary marking of 5 ms, and an 8 ms MAD for VC.

[8]MATLAB code for these computations is provided on C.R.'s website at redmonc.github.io/matlab.

[9]While this procedure introduces some degree of subjectivity in the assessment of periodicity, it was chosen over an objective, threshold-based measure because we are uncertain at this stage as to the reliability of the precise autocorrelation computed from irregular oscillations.

[10]The unique pattern of productions exhibited by SF01 may be due to Galician influence, as she is from southern Galicia, and uvular fricatives have previously been observed in related Portuguese (Jesus and Shadle, 2005).

[11]Unless otherwise stated, all point estimates are reported as the median of the posterior distribution, and credible intervals as the 95% highest posterior density interval (HPDI). All coefficients for logistic regressions are reported as odds ratios, where subscripts define the levels compared in the ratio (e.g., $CI_{a/b}$ is the credible interval for the probability of category *a* relative to category *b*). Unlike linear regression coefficients, the null value for an odds ratio is 1 (i.e., equal probabilities in *a* and *b*), and thus confidence intervals excluding 1 would be considered "significant" evidence against the null.

[12]Based on the power-weighted frequency analysis, an optimal filtering threshold was defined, 120 Hz, that was above all oscillation rates for AM02, PF01, and SF01, and which, being lower than the initial cutoff of 200 Hz made the onset and offset of individual cycles clearer.

[13]See the supplemental material at https://doi.org/10.1121/1.5045345 for representative acoustic and oral airflow signals from each speaker.

[14]While the occurrence of each cycle is not independent, Cycle Count was modeled as a Poisson distribution due to its skewness and the equality between its mean and variance.

[15]Sample signals at either end of the SFR spectrum, as well as more ambiguous tokens around 0 dB, are provided in the supplemental material (footnote 13).

Al-Khairy, M. A. (**2005**). "Acoustic characteristics of Arabic fricatives," Ph.D. thesis, University of Florida.
Alwan, A. (**1986**). "Acoustic and perceptual correlates of pharyngeal and uvular consonants," MS thesis, MIT.
Barry, W. J. (**1997**). "Another R-tickle," J. Int. Phonetic Assoc. **27**(1–2), 35–45.
Behrens, S., and Blumstein, S. E. (**1988**). "On the role of the amplitude of the fricative noise in the perception of place of articulation in voiceless fricative consonants," J. Acoust. Soc. Am. **84**(3), 861–867.
Boersma, P., and Weenink, D. (**2016**). "Praat: Doing phonetics by computer" [Computer software], http://www.praat.org/.
Colantoni, L. (**2006**). "Increasing periodicity to reduce similarity: An acoustic account of deassibilation in rhotics," in *Selected Proceedings of the 2nd Conference on Laboratory Approaches to Spanish Phonetics and Phonology* (Cascadilla, Somerville, MA), pp. 22–34.
Demolin, D. (**2001**). "Some phonetic and phonological observations concerning /r/ in Belgian French," in *r-atics: Sociolinguistic, Phonetic and Phonological Characteristics of /r/*, edited by H. Van de Velde and R. van Hout (Etudes and Travaux, Brussels), pp. 63–73.
Fant, G. (**1960**). *Acoustic Theory of Speech Production: With Calculations Based on X-Ray Studies of Russian Articulations* (Mouton & Co., The Hague).
Flanagan, J. L. (**1972**). *Speech Analysis*, 2nd ed. (Springer, New York).
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.
Ghazeli, S. (**1977**). "Back consonants and backing coarticulation in Arabic," Ph.D. thesis, University of Texas at Austin.
Hedrick, M. S., and Ohde, R. N. (**1993**). "Effect of relative amplitude of frication on perception of place of articulation," J. Acoust. Soc. Am. **94**(4), 2005–2026.
Jackson, P. J. (**2000**). "Characterisation of plosive, fricative and aspiration components in speech production," Ph.D. thesis, University of Southampton.
Jakobson, R., Fant, C. G., and Halle, M. (**1951**). *Preliminaries to Speech Analysis: The Distinctive Features and Their Correlates* (The MIT Press, Cambridge, MA).
Jassem, W. (**1965**). "The formants of fricative consonants," Lang. Speech **8**(1), 1–16.
Jesus, L. M., and Shadle, C. H. (**2005**). "Acoustic analysis of European Portuguese uvular [χ, ʁ] and voiceless tapped alveolar [ɾ̥] fricatives," J. Int. Phonetics Assoc. **35**(1), 27–44.
Jongman, A., Wayland, R., and Wong, S. (**2000**). "Acoustic characteristics of English fricatives," J. Acoust. Soc. Am. **108**(3), 1252–1263.
Ladefoged, P., Cochran, A., and Disner, S. (**1977**). "Laterals and trills," J. Int. Phonetics Assoc. **7**(2), 46–54.
Ladefoged, P., and Maddieson, I. (**1996**). *The Sounds of the World's Languages* (Oxford Publishers, New York).
Lazard, G. (**1992**). *A Grammar of Contemporary Persian* (Mazda Publishers, Costa Mesa, CA).
Maddieson, I. (**1992**). UCLA Phonological Segment Inventory Database, UCLA, Los Angeles, CA.
Martínez-Celdrán, E., Fernández-Planas, A. M., and Carrera-Sabaté, J. (**2003**). "Castilian Spanish," J. Int. Phonetics Assoc. **33**(2), 255–259.
MathWorks (**2016**). MATLAB version 9.0 (R2016a). MathWorks, Inc., Natick, MA.
McMurray, B., and Jongman, A. (**2011**). "What information is necessary for speech categorization? Harnessing variability in the speech signal by integrating cues computed relative to expectations," Psychol. Rev. **118**(2), 219–246.
Newland, D. E. (**1984**). *An Introduction to Random Vibrations and Spectral Analysis*, 2nd ed. (Longman, New York).
Schiller, N. O. (**1999**). "The phonetic variation of German /r/," in *Variation und Stabilität in der Wortstruktur* (Georg Olms, Verlag), pp. 261–287.
Scicon R&D Inc. (**2015**). X16 series hardware system, Scicon R&D, Inc., Los Angeles, CA, http://www.sciconrd.com/x16.aspx (Last viewed July 2, 2018).
Scully, C. (**1990**). "Articulatory synthesis," in *Speech Production and Speech Modelling*, edited by W. Hardcastle and A. Marchal (Springer, Dordrecht), pp. 151–186.
Sebregts, K. (**2015**). *The Sociophonetics and Phonology of Dutch r* (LOT Publishers, Utrecht).
Shadle, C. H. (**1985**). "The acoustics of fricative consonants," Ph.D. thesis, MIT.
Shadle, C. H. (**1990**). "Articulatory-acoustic relationships in fricative consonants," in *Speech Production and Speech Modelling*, edited by W. Hardcastle and A. Marchal (Springer, Dordrecht), pp. 187–209.
Shadle, C. H., and Scully, C. (**1995**). "An articulatory-acoustic-aerodynamic analysis of [s] in VCV sequences," J. Phonetics **23**(1), 53–66.
Shosted, R. (**2008a**). "An aerodynamic explanation for the uvularization of trills," in *Proceedings of the 8th International Speech Production Seminar*, pp. 421–424.
Shosted, R. (**2008b**). "Acoustic characteristics of Swedish dorsal fricatives," J. Acoust. Soc. Am. **123**(5), 3888.
Shosted, R. K., and Chikovani, V. (**2006**). "Standard Georgian," J. Int. Phonetics Assoc. **36**(2), 255–264.
Solé, M.-J. (**1998**). "Phonological universals: Trilling, voicing, and frication," Ann. Meet. Berkeley Ling. Soc. **24**(1), 403–416.
Solé, M.-J. (**2002**). "Aerodynamic characteristics of trills and phonological patterning," J. Phonetics **30**(4), 655–688.
Solé, M.-J. (**2010**). "Effects of syllable position on sound change: An aerodynamic study of final fricative weakening," J. Phonetics **38**(2), 289–305.
Stan Development Team (**2016**). RStan: The R interface to Stan. R package version 2.14.1.
Stevens, K. N. (**1971**). "Airflow and turbulence noise for fricative and stop consonants: Static considerations," J. Acoust. Soc. Am. **50**(4B), 1180–1192.
Stevens, K. N. (**1998**). *Acoustic Phonetics* (MIT Press, Cambridge, MA).
Stevens, K. N., and Blumstein, S. E. (**1978**). "Invariant cues for place of articulation in stop consonants," J. Acoust. Soc. Am. **64**(5), 1358–1368.
Strevens, P. (**1960**). "Spectra of fricative noise in human speech," Lang. Speech **3**(1), 32–49.
Thelwall, R., and Sa'Adeddin, M. A. (**1990**). "Arabic," J. Int. Phonetics Assoc. **20**(2), 37–39.
Verhoeven, J. (**1994**). "Fonetische Eigenschappen van de Limburgse huig-r" ("Phonetic characteristics of the Limburg uvular r"), Taal en Tongval **46**(7), 9–21.
Watson, J. C. (**2002**). *The Phonology and Morphology of Arabic* (Oxford University Press, New York).
Yeou, M., and Maeda, S. (**2011**). "Airflow and acoustic modelling of pharyngeal and uvular consonants in Moroccan Arabic," in *Instrumental Studies in Arabic Phonetics*, edited by Z. M. Hassan and B. Heselwood (John Benjamins Publishing Co., The Netherlands) pp. 141–162.
Zeroual, C. (**1999**). "A fiberscopic and acoustic study of guttural and emphatic consonants of Moroccan Arabic," in *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco, CA, pp. 997–1000.
Zeroual, C. (**2003**). "Aerodynamic study of Moroccan Arabic guttural consonants," in *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, Spain, pp. 1859–1862.