Acoustic Analysis of the Uvular Unvoiced Fricative

Ashraf Alkhairy

KACST
P. O. Box 6086
Riyadh 11442, Saudi Arabia

ABSTRACT

An acoustic analysis of the uvular unvoiced fricative in intervocalic sequences, as spoken in Arabic, is presented. Properties of both the noise region and voiced boundaries are examined, and several attributes are discovered: the glottal waveform changes considerably during its transition between the vowels and the fricative boundaries, with the open quotient becoming 100% at the boundaries compared to around 50% at the neighboring vowels; the noise region in the unvoiced fricative exhibits an underlying periodic pattern with a frequency of vibration of around 90Hz; the directions of formant trajectories into the fricative are highly dependent on the intervocalic context and different from predictions of articulatory-acoustic models for most cases; the turbulence noise strongly excites the first formant, in addition to the expected second formant. All these results form the basis for determining perceptual cues, developing aerodynamic models and acoustic models, and building high quality formant synthesizers.

1 Introduction

In this paper we study the acoustic properties of the uvular unvoiced fricative in the context of symmetric intervocalic sequences. Unvoiced fricatives are produced with creation of a turbulent airflow and generation of noise near a narrow constriction; this is unaccompanied with any voicing source excitation at the glottis [1]. The uvular consonants are produced with the back of the tongue raised towards the uvula [2].

Such studies are important because they supplement existing phonetic knowledge with information about guttural sounds; thus providing a more comprehensive understanding of consonants in general and unvoiced fricatives in particular.

The Arabic language is used for conducting our study because it is known to contain many guttural sounds [3]. Several research papers on this subject have been published to date [4, 5, 6, 7, 8], but still substantial research is needed.

Our investigation reveals many characteristics of the uvular unvoiced fricative. The formant trajectories, in general, do not validate the conclusions of articulatory-acoustic models and their directions are dependent on the vocalic context; this is agreement with the observation in [7]. Both the first and second formants get excited during the frication period. This has been observed also in other unvoiced fricatives. A 90Hz quasi periodic waveform is superimposed on the frication noise. At the voice boundaries of the fricative, the glottal waveform is characterized by the absence of the closed phase - which is different from its regular condition. The last two attributes have not been addressed in previous research. Such results necessitate re-examination of articulatory-acoustic models, refinement of glottal models, improvement of synthesizers, physiologic investigation, and research into perceptual cues for the sound.

The results of this paper form points of investigation for a future physiologic analysis of the uvular unvoiced fricative to validate and explain the above observations. This paper focuses only on the acoustic phonetics of the uvular unvoiced fricative in intervocalic sequences. This includes waveform analysis, intensity analysis, formant analysis, and glottal source analysis at the boundaries of the fricative. Issues that relate to physiologic analysis, modeling, inter-speaker differences, and syllable and suprasegmental features such as gemination and stress contrast are outside the scope of this work.

This paper proceeds by describing the methodology used in Section 2. Time-domain analysis related to the properties of the pressure waveform and its intensity is presented in Section 3. Frequency-domain analysis of resonances is discussed in Section 4. The boundary conditions of the fricative with respect to the fundamental frequency and the glottal waveform are examined in Section 5.

2 Methodology

The linguistic material for our study consists of the unvoiced uvular fricative in geminate form embedded between the symmetric sequence /z/V and V/z/ where

 ISBN 1-876346-48-5

z is the voiced alveolar fricative [2]. The geminate form is chosen to allow an accurate study of the steady state characteristics of the unvoiced fricative. The vowel, V, takes the value of all vocals in Arabic, /a/, /u/, and /i/. The recording consists of words spoken by adult native Saudi Arabian male subjects. Each utterance was digitized with a sampling frequency of 10,000 Hz and quantization of 15 bits. In addition, the sequences /z/ /aa/ /z/, /z/ /uu/ /z/, and /z/ /ii/ /z/, were recorded to facilitate the comparison between the fricative and the steady state vowels. The sequences are initiated with a consonant to satisfy Arabic phonological rules, and they end with the same consonant to provide symmetry. The voiced fricative /z/ is chosen as the start and end consonant because it can be clearly demarcated. It also serves the purpose of embodying suprasegmental features that mark the beginning and end of a word, thereby reducing prosodic effects on the VCV segment.

The recorded data were pitch marked to allow time-domain analysis, frequency-domain analysis and the determination of the glottal waveform at the consonant boundaries. The pitches were determined using a peak detection algorithm to mark the zero-crossings in the region of fastest transition from negative to positive amplitudes.

Five locations were marked on each utterance: V_-, the steady state of the preceding vowel, B_-, the start boundary of the fricative, C_{ss}, the steady state of the frication noise region, B_+, the end boundary of the fricative, and V_+, the steady-state of the following vowel. The beginning boundary, B_-, is located within the last well-defined pitch period before frication noise, and the end boundary, B_+, is placed within the first well-defined pitch period after the frication noise.

3 Time-Domain Analysis

Analysis of the pressure waveform plotted against time was conducted. First, all the waveforms were subjected to visual inspection to allow determination of the underlying pattern. Second, the intensity of the waveform at various time-marks were computed. These issues are discussed in the next two subsections.

3.1 Waveform

The pressure waveforms of all the utterances were plotted for the regions in the neighborhood of B_-, C_{ss}, and B_+, representing the boundaries and the steady state of the fricative.

All the noise waveforms, without exception, exhibited an underlying periodic pattern as shown in Figure 1. The plot illustrates the mid section of the noise region of the fricative spoken in the context of /i/. The waveforms are characterized by three attributes: 1) Vibration is absent at the vowel-consonant and consonant-vowel boundaries (not shown in plot). 2) A substantial portion of the energy is attributed to the quasi-periodic part of the signal, with the random noise contributing no more than half the energy. 3) The frequency of vibration is variable and its nominal value is around 90Hz. This pattern cannot be attributed to environmental factors because other unvoiced fricatives, such as the pharyngeal (the subject of another study) spoken by the same set of speakers in the same experiment did not exhibit any vibration. A clarification of the causes of this vibration can be determined from a physiologic study.

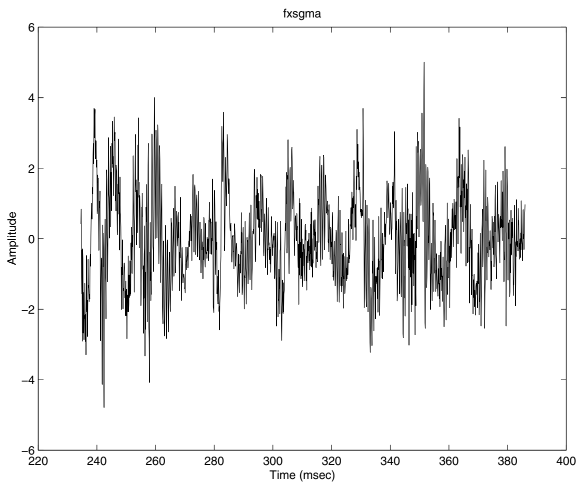

Figure 1: Pressure waveform of fricative

3.2 Intensity

Intensity is computed at the five locations, V_-, B_-, C_{ss}, B_+, V_+ in each utterance. For the noise region, intensity is computed as the power in a 13-millisecond region around its steady state. This is large enough to yield accurate results and is small enough to avoid any transitory effects. The intensity for the four voiced locations in each utterance is calculated using the power in a pitch period.

Statistics for the deviation in intensity from the value at V_+ are shown in Table 1. The results reveal a difference in intensity between the unvoiced fricatives and the vowels of around 25 db and a symmetry in intensity between the preceding and the succeeding vowels. This is consistent with figures commonly found for unvoiced fricatives.

Table 1: Intensity Deviations

	/a/	/u/	/i/
Δ Inten(V_-)	0	-1	-2
Δ Inten(B_-)	-4	-10	-7
Δ Inten(C_{ss})	-21	-28	-21
Δ Inten(B_+)	-1	-6	-3

ISBN 1-876346-48-5

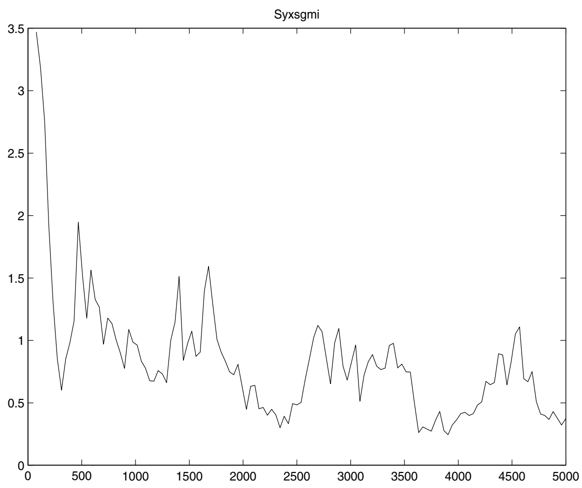

Figure 2: Spectrum at steady state

4 Frequency-Domain Analysis

In order to study the noise spectral characteristics of the uvular unvoiced fricative, the spectra were estimated by calculating the covariance of the noise sequence, multiplying it by a Hamming window of lag 128 and then Fourier transforming the result. The lag length is less than 10% of the total size of the noise sequence thereby resulting in accurate estimate of the spectrum. The frequency-domain bandwidth of the window is around 130Hz which is small enough to resolve any peaks of interest. An example of the spectrum estimate is shown in Figure 2. The peaks clearly show the excitation of both the first and second formants.

The formant frequencies were computed at C_{ss} using the autocorrelation fixed length LPC method with preemphasis of 0.6, Hamming window of 12 ms, and filter of size 12. For comparative purposes, the formants were computed at the corresponding steady-state vowel positions using the synchronized autocovariance LPC method with a preemphasis of 0.95 and filter of size 10. The formant frequencies in Hz for the vowels are as follows: /aa/: 671, 1563, 2662, 3660, 4475. /ii/: 393, 2153, 2752, 3587, 4630. /uu/: 456, 992, 2480, 3235, 4220.

Table 2 presents the mean values of the first five formant frequencies for the fricative spoken in the context of the three vowels. Numbers with an asterisk indicates that it is not significantly different from the value of the corresponding steady state vowel.

Comparison of these formant values to the values of the steady state vowels shows that F_1 is higher for the fricative than the vowel in the context of /i/, F_2 is lower in the context of /a/ and /i/, and F_3 is higher in the context of /a/ and /u/. No other significant changes were observed. Two points may be noted: First, the formant trajectory directions are substantially dependent on the vowel context, making articulatory-acoustic modeling more difficult and dependent on context. Second, the predictions of articulatory-acoustic models are only applicable to the context of /a/ [9].

Table 2: Formant Frequencies

	/a/	/u/	/i/
F_1	792*	469*	638
F_2	1459	956*	1718
F_3	2899	2681	2602*
F_4	3551*	3472	3291
F_5	4356*	4414*	4039*

To determine the excitation characteristics of the fricative, the amplitudes in dB of the spectrum peaks at C_{ss} were computed. The intensities of the noise region, as computed in Subsection 3.2 were subtracted from the amplitudes to normalize for variations. The results of the average differences are shown in Table 3. As can be seen, both F_1 and F_2 get excited, while existing models predict that only F_2 should be excited [9]. These models need to be refined to explain the reason for excitation of F_1, and synthesizers need to be modified to include F_1 amplitude [10].

Table 3: Amplitude of Frication

	/a/	/u/	/i/
AF_1	7	8	4
AF_2	12	9	8
AF_3	-3	-12	-2
AF_4	1	-8	1
AF_5	2	-4	4

5 Boundary Conditions

Analysis of the conditions that prevail at the beginning and end of the fricative can facilitate inclusion of precise transitional effects in the development of synthesizers. In addition, it can yield information on the physiological changes that occur and the cues that effect the perception of the fricative. Since our study is conducted on an unvoiced fricative bounded by (voiced) vowels, it is meaningful to study the changes that occur at the glottis. The next two subsections examine this within the context of fundamental frequency and characteristics of the glottal waveform.

5.1 Fundamental Frequency

At the transition from voiced vowels to unvoiced fricatives, the vocal folds move from a state of being close together and vibrating to a state of being pulled apart. A study of the pressure waveforms indicates that this is an abrupt change, with voicing disappearing within a single pitch period.

ISBN 1-876346-48-5 © 2003 UAB

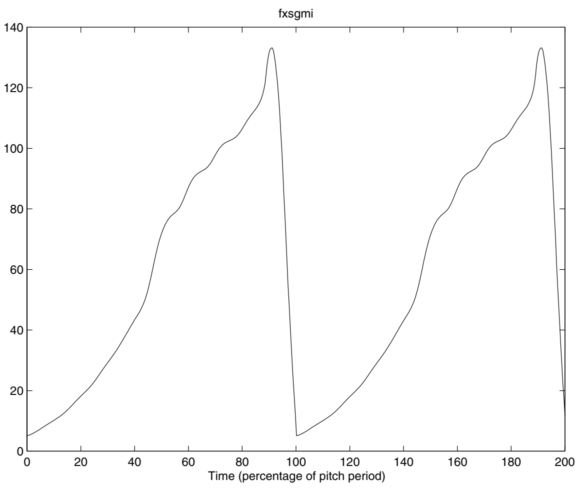

Figure 3: Glottal waveform at boundary

In order to examine the changes that occur at the boundary, the fundamental frequencies were computed at four locations within the voiced regions: V_-, B_-, B_+ and V_+. In addition, the fundamental frequencies of the corresponding long vowels spoken by the same speakers were computed. The long vowel fundamental frequencies were subtracted from the corresponding fundamental frequencies at the four locations. Statistics of the differences were computed and are shown in Table 4, with means appearing for each vowel and the standard deviations shown in the brackets. The high standard deviations indicate that there is no reason to believe that any consistent changes in the fundamental frequency occur at the boundaries from that of vowel positions.

Table 4: Fundamental Frequency

	/a/	/u/	/i/
$\Delta F_0(V_-)$	3 (21)	5 (22)	-1 (13)
$\Delta F_0(B_-)$	-19 (34)	-15 (14)	-22 (19)
$\Delta F_0(B_+)$	-3 (22)	-2 (19)	-8 (11)
$\Delta F_0(V_+)$	-10 (21)	-5 (24)	-20 (10)

5.2 Voicing Source

The glottal volume velocity waveforms at the fricative boundaries, B_- and B_+, were computed using a factorization algorithm [11]. A sample graph of the waveform is shown in Figure 3. The open quotient, representing the ratio in duration for the open phase to the closed phase, was computed for each waveform, and was found in all cases to be 100%. Thus substantial differences exist between the glottal waveform at the boundaries of the fricative and at the neighboring vowels, with the absence of the closed phase at the boundaries. This is consistent with the observation that the vocal folds have to prepare to remain open during the unvoicing of the fricative.

REFERENCES

[1] P. Badin, “Acoustics of voiceless fricatives: Production theory and data,” *STL Quarterly Progress and Status Report*, vol. 3, pp. 33–55, 1989.

[2] P. Ladefoged, *A Course in Phonetics*, Harcourt Brace, 1993.

[3] I. Maddieson, *Patterns of Sound*, Cambridge University Press, 1984.

[4] S. Al-Ani, *Arabic Phonology*, Mouton, 1970.

[5] S. Ghazeli, *Back Consonants and Backing Coarticulation in Arabic*, PhD Dissertation - University of Texas at Austin, 1977.

[6] P. Delattre, “Pharyngeal features in the consosonants of arabic, german, spanish and american english,” *Phonetics*, vol. 23, pp. 129–155, 1971.

[7] A. Alwan, *Acoustic and Perceptual Correlates of Pharyngeal and Uvular Consonants*, Master’s Thesis - MIT, 1986.

[8] A. Butcher and K. Ahmad, “Some acoustic and aerodynamic characteristics of pharyngeal consonants in iraqi arabic,” *Phonetica*, vol. 44, pp. 156–172, 1987.

[9] D. H. Klatt and K. N. Stevens, “Pharyngeal consonants,” *RLE Quarterly Progress Report - MIT*, vol. 93, pp. 207–215, 1969.

[10] D. H. Klatt, “Review of text-to-speech conversion for english,” *JASA*, vol. 82, pp. 737–793, 1987.

[11] A. Alkhairy, “An algorithm for glottal volume velocity estimation,” *IEEE ICASSP*, 1999.