International Journal of Linguistics and Translation Studies
Volume 5, Issue 3, 2024
Homepage: http://ijlts.org/index.php/ijlts/index
DOI: https://doi.org/10.36892/ijlts.v5i3.471

The Acoustics of Sibilant Fricatives in Urban Hijazi Arabic

Majed Al Solami
King Abdulaziz University, Kingdom of Saudi Arabia
msolami@kau.edu.sa

How to cite:
Al Solami, M.(2024). The Acoustics of Sibilant Fricatives in Urban Hijazi Arabic. *International Journal of Linguistics and Translation Studies* 5(3).19-27. https://doi.org/10.36892/ijlts.v5i3.471

ARTICLE HISTORY

Received: 21/03/2024

Accepted: 20/05/2024

KEYWORDS

sibilants, center of gravity, noise duration, intensity, gender

Abstract

This paper examines the spectral characteristics of sibilants [s], [z], [ʃ], and [ʒ] in the Urban Hijazi Dialect. The analysis investigates intensity, noise duration and centre of gravity (COG). These acoustic cues vary based on gender, places of articulation and voicing, as reported in other studies. The study aims to fill a gap in the literature of Arabic fricatives and to see whether these characteristics are influenced by places of articulation, voicing or gender in the aforementioned dialect. The dataset includes word-initial fricatives followed by the short vowels [ɪ], [a], and [ʊ]. The results suggest that intensity is not very informative as an acoustic cue in differentiating between sibilants. The other two spectral measurements, noise duration and COG, are more reliable. The female speakers produced fricatives with longer durations and higher COG values compared to the male speakers. In addition, for male speakers, voiced fricatives are shorter in duration and have less acoustic energy compared to the voiceless fricatives. As for the female speakers, voiceless sibilants have longer durations, similar to the male speakers, but COG values are not influenced by voicing. In both genders, the place of articulation influences COG values, where alveolars have higher COG compared to post-alveolars.

1. INTRODUCTION

Fricatives are found in all world languages and they are produced with a high acoustic energy that results from forcing air through a narrow aperture. Arabic dialects have a number of fricatives including [θ, ð, f, s, z, ʃ, ʒ, h][1]. Arabic fricatives have been the subject of multiple studies, but studies that explored different spectral mesurements in both genders are very limited. This paper explores the acoustics of the natural class of sibilants [s], [z], [ʃ], and [ʒ] by analyzing their spectral characteristics including intensity, centre of gravity, henceforth COG, and noise duration. The analysis focuses on the Urban Hijazi Arabic Dialect by recruiting two male and two female speakers from Jeddah city in Saudi Arabia. It aims to examine potential influence of voicing, place of articulation, and gender on these sibilants.

Previous studies examined the acoustic characteristics of fricatives, here we review some of these studies. For COG values, previous works suggested that [s] and [z] have greater COG compared to [ʃ] in English language (Jongman et al., 2000; Koffi and Bloch, 2017). Haley et al. (2010) reported

[1] Arabic dialects also have emphatic fricatives [ðˤ, sˤ] that have a secondary articulation.

that places of articulation influenced fricatives since fricative [ʃ] is articulated further back in the mouth than the fricative [s] and, as a result, has a longer front cavity and lower spectral means. In other words, COG values decrease as the place of articulation moves further back into the vocal tract (Jongman et al., 2000). COG of [ʃ] in American English is suggested to be around 4200 Hz (Jongman et al., 2000; Haley et al., 2010).

Wallin and Koffi (2017) reported that for [s] the COG is 5137 Hz and for [z] it is 4015 Hz, which means that [s] has a higher COG. As for [ʃ] and [ʒ], on the other hand, [ʃ] has higher COG (4738 Hz) than [ʒ] (3094 Hz). This indicates that voiceless fricatives are expected to have higher frequency levels compared to voiced fricatives. Other studies reported that the COG of [ʃ] in American English is around 4200 Hz (Jongman et al., 2000; Haley et al., 2010).

In the branch of speech perception in psychology, sounds are usually compared within just noticeable difference limits, henceforth JND, see for example Stevens (2000) and Moore (2007). JND proposed the amount of minimum difference needed between speech sounds in order to be noticeable. Within the function of JND, the frequency difference between two distinct sounds needs to be ≥ 630 Hz (Bloch, 2017).

Intensity is reported to be the most reliable cue for differentiating between voiced and voiceless fricatives (Ladefoged and Maddieson, 1996; Thomas, 2011). It is also said to be influenced by the gender of the speaker, where male speakers tend to have higher intensity compared to female speakers for the fricative [ʃ] (Koffi and Bloch, 2017). The intensity of [ʃ] is around 60 dB (Jongman et al., 2000; Haley et al., 2010). Koffie and Lopez-Backstrom (2018) suggested that there is no difference based on intensity between [z] and [s] in their study due to devoicing effects. Based on JND boundaries, at least a difference of 3 dB is needed to differentiate between the segments (Koffi and Lopez-Backstrom, 2018).

As for noise duration, many studies indicated that voiceless fricatives have longer noise durations than voiced fricatives (Klatt, 1976; Jongman et al., 2000; Koffi and Bloch, 2017; Wallin and Koffi, 2017). However, some other studies downplayed the role of noise duration as it is not very reliable in discriminating between fricatives. For example, Wallin and Koffi (2017) showed that voiced [z] is longer than voiceless [s] while voiceless [ʃ] is longer than voiced [ʒ].

Gender differences were cited in some studies. Gender influences the production of fricatives, and other sounds for that matter, due to physiological differences between males and females since males usually have longer vocal tracts compared to females (Kochetov and Lobanova, 2007). Male speakers usually produce fricatives with longer noise duration compared to female speakers. In addition, male speakers produce fricatives with higher intensity compared to female speakers (Jongman et al., 2000; Koffi and Bloch, 2017). COG is also influenced by the gender of the speaker. Jongman et al. (2000) suggested that females produce fricatives with higher frequencies compared to males. Similar results were also reported by Haley et al. (2010). Not all studies found duration to be a reliable factor in separating the two genders. Gordon et al. (2002), for example, reported no such differences in duration.

The remainder of the paper is divided into a number of sections. The following section introduces the participants in this study and data analysis process. This is followed by the results in section three. Section four discusses the main outcomes of the results and, finally, concludes the paper.

2. PARTICIPANTS AND DATA

The study recruited two male and two female speakers from Jeddah City, Saudi Arabia. They spoke the local dialect known as Urban Hijazi Arabic Dialect and reported no speech or hearing problems. Each participant produced 12 words containing the four sibilant consonants [s], [z], [ʃ] and [ʒ] in the word-initial position followed by one of the three short vowels [ɪ], [a], and [ʊ], see Table 1. Each participant repeated each word in Table 1 four times at a normal speech rate. The total number of tokens is 192 (12 words x 4 speakers x 4 repetitions).

Table 1. Stimuli

Fricative C	Cɪ	Ca	Cʊ
[s]	[sɪn] 'tooth'	[sab] 'cursed'	[sʊb] 'curse!'
[z]	[zɪr] 'button'	[zar] 'pressed'	[zʊr.ga] 'blueness'
[ʃ]	[ʃɪ.bɪl] 'lion cub'	[ʃar] 'evil'	[ʃʊr.fa] 'porch'
[ʒ]	[ʒɪn] 'jinn'	[ʒar] 'pulled'	[ʒʊr] 'pull!'

Each data elicitation session was recorded using a Marantz PMD660 digital recorder connected to a cardioid microphone in a quiet room to ensure the best possible audio quality. The audio was stored in a WAV format and then imported into a computer to be analyzed using Praat software (Boersma and Weenink, 2024). The sampling rate for all the recordings was 44100 Hz at 16 bits per sample.

The sibilant segments were marked using Praat text grids. The following measurements were taken. Duration, which is the length of the noise frication of each fricative, was obtained by manually selecting it from the spectrograms. Frication was selected from its onset, the point when high-frequency striation first appeared on the spectrogram, till the end of the frication noise. The intensity was obtained from the *get intensity* algorithm function from Praat. The centre of gravity of the fricatives was measured by selecting 45 ms from the centre of the frication noise. When the frication noise is less than 45 ms in length, the entire frication noise was selected. Then, FFT spectra were produced to measure COG. The results were entered into an Excel spreadsheet to calculate the mean of duration, intensity and the COG. Figure 1 shows a sample of how [s] fricative was segmented and annotated in the token [sʊb].

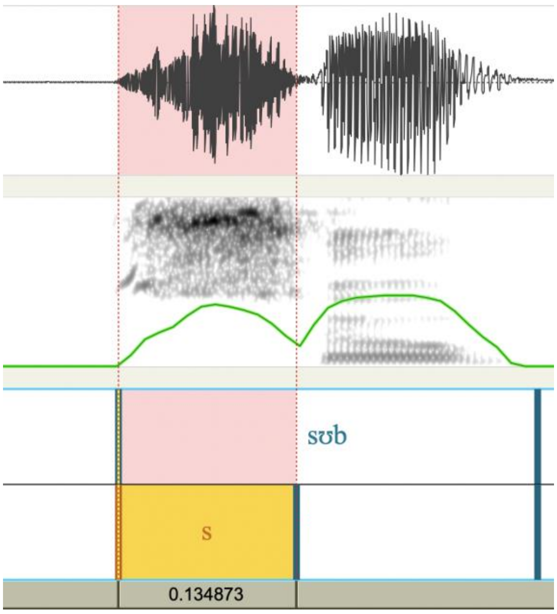

Figure 1. Annotation and segmentation of one of the tokens of [sʊb] ‘curse!’

3. RESULTS

3.1. Male Speakers’ Data

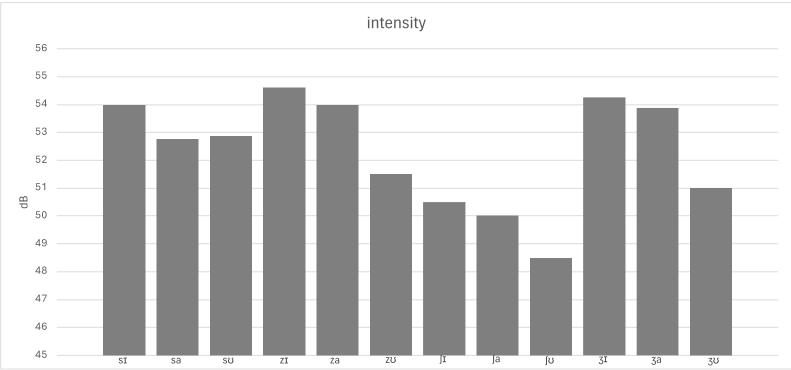

Figure 2: Averages of the intensity of the fricatives produced by the male speakers

Figure 2 shows the averages of the intensity of the sibilants preceding different vowels in the male speakers’ data. A general observation is that the voiced sibilants have higher intensity compared to the voiceless ones.

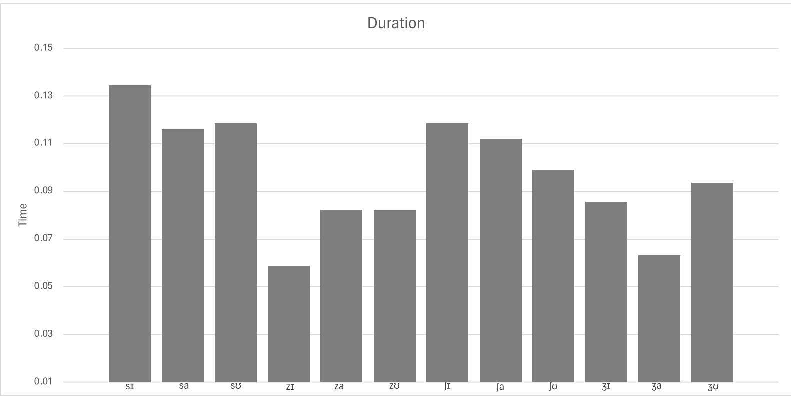

Figure 3: Averages of noise duration of the fricatives produced by the male speakers

Figure 3 presents the average noise durations of the fricatives produced by the male speakers. The results show that the voiceless sibilants have longer noise durations compared to the voiced sibilants.

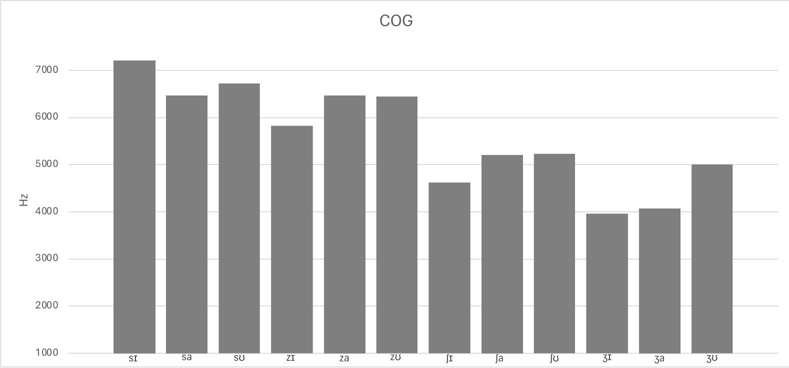

Figure 4: Averages of COG of the fricatives produced by the male speakers

Figure 4 shows the average COG of the male speakers. As a general trend, the voiceless sibilants have higher frequency energies compared to the voiced fricatives. Another observation is that fricatives lose energy as the place of articulation is further back in the vocal tract. This means that the alveolar fricatives have higher frequencies than the post-alveolar fricatives.

3.2.Female spakers' results

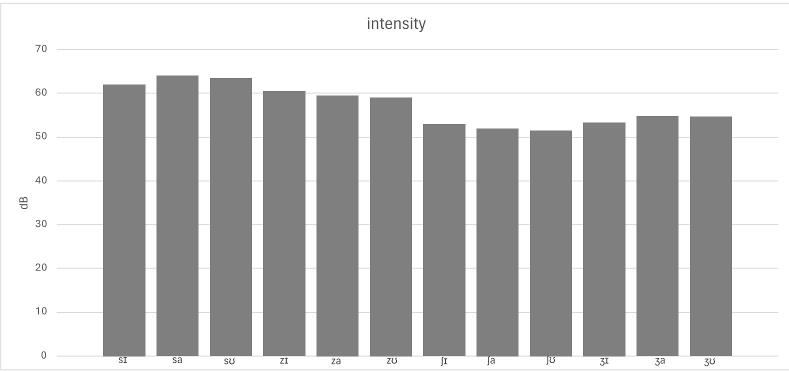

Figure 5: Averages of the intensity of the fricatives produced by the female speakers

Figure 5 illustrates the intensity averages of the sibilants preceding different vowels produced by the female speakers. The intensity levels are somewhat consistent and do not indicate any noticeable patterns. Voice and place of articulation do not influence the intensity values.

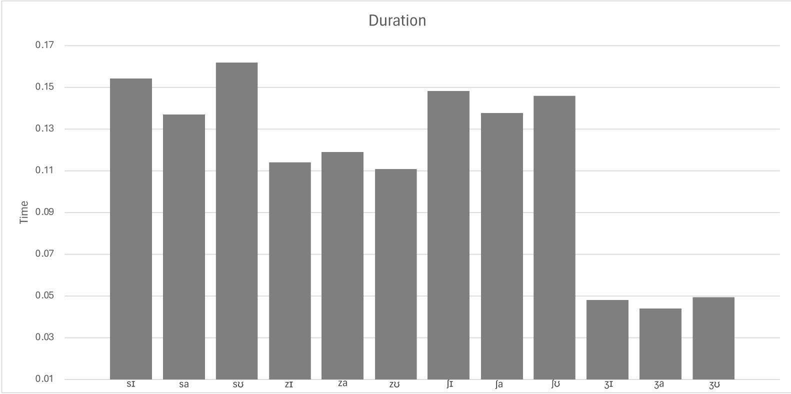

Figure 6: Averages of the duration of the fricatives produced by the female speakers.

Figure 6 presents the averages of noise duration of the fricatives in the data of the female speakers. Duration is a robust cue among these sibilants since the voiced sibilants are considerably shorter than voiceless sibilants.

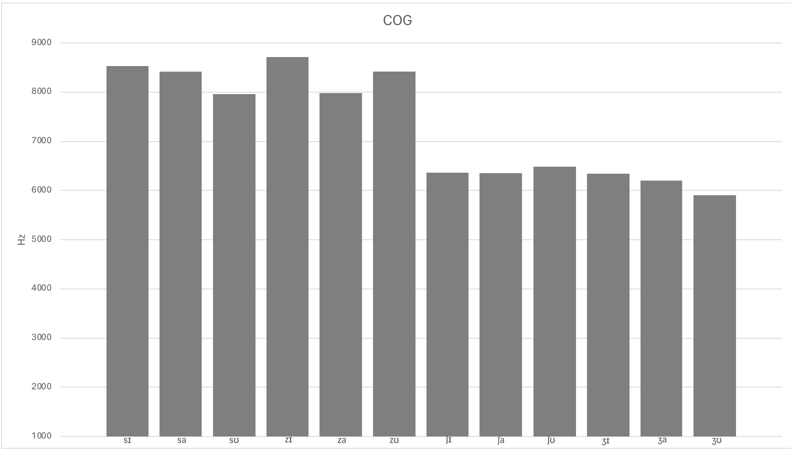

Figure 7: Averages of COG of the fricatives produced by the female speakers

Figure 7 shows the average COG values of the fricative in the female speakers' data. The frequency levels of the alveolar fricatives are quite similar. The same observation can be made about the post-alveolar fricatives. A noticeable pattern is that the alveolars have higher frequency concentrations compared to the post-alveolars. Voicing does not seem to influence COG values since both voiced and voiceless sibilants have similar values.

Table 2 shows the combined mean differences in intensity, duration, and COG of the examined fricatives in the male and female speakers. This table is very helpful in the discussion that follows in the next section.

Table 2. Combined mean values of intensity, duration, and COG in male and female speakers

Intensity (dB)		
	Male speakers	Female speakers
[s]	53	54
[z]	53	59
[ʃ]	50	52
[ʒ]	52	53
Duration (ms)		
	Male speakers	Female speakers

[s]	121	151
[z]	71	114
[ʃ]	109	143
[ʒ]	81	47

COG (Hz)

	Male speakers	Female speakers
[s]	6975	8300
[z]	6245	8370
[ʃ]	5021	6393
[ʒ]	4344	6146

To summarise, based on Table 2, the female speakers produced longer noise durations in fricatives compared to the male speakers, except for the voiced post-alveolar [ʒ]. In addition, COG is significantly higher in the fricatives produced by female speakers. Intensity is very similar in both genders, except for [z] which is louder in the data of the female speakers.

Voicing is a factor in discriminating between the fricatives in the data of both genders based on duration values where voiced fricatives are notably shorter than the voiceless fricatives. COG values, on the other hand, differ based on voicing only in the male speakers' data. The female speakers' data do not show much variation based on voicing in COG values.

4. DISCUSSION AND CONCLUSION

The outcome of this study is similar to the findings of other studies, however, there are some differences. This is expected since the acoustics of fricatives can vary from speaker to speaker (Ladefoged and Maddieson, 1996). The findings of this study are discussed in this section.

Overall, noise duration and COG values are more reliable in categorizing fricatives, while intensity seems less important. Gender-based values show some variations where the female speakers utilize duration to differentiate between voiced and voiceless fricatives and COG as a cue of different places of articulation. The male speakers, on the other hand, follow a similar line in addition to COG to separate voiced from voiceless fricatives.

In terms of intensity, many studies suggested that it is a robust cue for separating voiced and voiceless fricatives (Ladefoged and Maddieson, 1996; Thomas, 2011). However, the intensity values in this study do not adhere to the same outcome. The only outcome where the intensity was influential in differentiating between voiced and voiceless consonants was when the female speakers produced the the fricative [z] louder than the fricative [s], which is different from the findings of Koffi and Lopez-Backstrom (2018) who reported no significant difference in intensity between [s] and [z].

For the female speakers, the difference in dB between [s] and [z] and other sibilants, as can be seen in Table 2, is more than 3 dB, which is the required acoustic distance to differentiate between sounds according to JND. Other pairs do not have the needed acoustic distance of 3 dB.

In both genders, the intensity of [ʃ] is the lowest compared to other fricatives. The value of intensity of [ʃ] is > 60, which is similar to the findings of Jongman et al. (2000).

To highlight gender-based differences, overall, female speakers produce louder fricatives, but the difference is not significant since it is not 3 dB, except for [z]. This result contradicts the findings of Koffi and Bloch (2017) and Jongman et al. (2000). This outcome is interesting and warrants further examination as the anatomical differences between the two genders are expected to enable the male speakers to produce considerably louder fricatives.

Table 2 shows that the duration differences between the voiceless and voiced sibilants are significant, whereas voiceless consonants are longer in duration for both genders. Noise duration is a reliable cue for voicing discrimination with longer durations in voiceless fricatives (Klatt, 1976; Jongman et al., 2000). Sibilant [s] is around 50 ms longer than [z] in the male speakers' data. Also, [ʃ] is around 30 ms longer than [ʒ]. Similar patterns can be observed between the voiceless and voiced sibilants in the female speakers' data.

Comparing the two genders in terms of noise duration reveals that the female speakers produce significantly longer fricatives except for [ʒ]. Jongman et al. (2000) specified that [ʃ] is 178 ms, while Gordon et al.'s data (2000) indicated that [ʃ] is 171 ms. In this study, [ʃ] is 109 ms in the male speakers' data and 143 ms in the female speakers' data.

Regarding the COG values, [s] and [z] have higher concentrations of energy compared to other fricatives, similar to the results of Jongman et al (2000), Wallin and Koffi (2017) and Koffi and Lopez-Backstrom (2018). The level of energy concentration decreases as the place of articulation is further back in the vocal tract. As a result, the alveolar fricatives have higher COG compared to post-alveolar ones, as shown in Table 2.

As for gender-based variations, Jongman et al (2000) and Haley et al. (2010) indicated that COG values of male speakers were lower than that of female speakers. The findings of this study, as in Table 2, are in line with these studies where the female speakers produce fricatives with higher frequency than the male speakers. The frequency difference is ≥ 630 Hz which makes the two datasets distinct according to JND.

Unlike the male speakers, voicing differences do not influence the COG values in the data of the female speakers. As can be seen in Figure 7, the female speakers produce voiced and voiceless fricatives with similar energy values.

To summarize, the above discussion indicates that the distinction between the sibilants [s], [z], [ʃ] and [ʒ] could be made based on the spectral characteristics of these fricatives. Specifically, a combination of noise duration and COG helps distinguish the voicing and place of articulation of these fricatives, while intensity is the least informative cue. Notably, gender-based variations are observed especially in COG and noise durations. The examination using Praat could benefit from a larger pool of participants and fricatives in different word positions. In addition, comparison with other Arabic dialects is a point to consider in future research.

REFERENCES

Boersma, P. and Weenink, D. (2024) PRAAT: Doing Phonetics by Computer (Version 6.4.08) [Computer software]. http://www.fon.hum.uva.nl/praat/

Gordon, M., Barthmaier, P., & Sands, K. (2002). A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association*, *32*(2), 141-174.

Haley, K. L., Seelinger, E., Mandulak, K. C., & Zajac, D. J. (2010). Evaluating the spectral distinction between sibilant fricatives through a speaker-centered approach. *Journal of Phonetics*, 38(4), 548-554.

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America*, *108*(3), 1252-1263.

Klatt, Dennis H. (1976). Linguistic Uses of Segmental Duration in English: Acoustic and Perceptual Evidence. *Journal of the Acoustical Society of America*, 59 (5), 1208-1221.

Kochetov, A., & Lobanova, A. (2007). Komi-Permyak coronal obstruents: Acoustic contrasts and positional variation. *Journal of the International Phonetic Association*, *37*(1), 51-82.

Koffi, E., & Bloch, M. (2017). The acoustic correlates of the voiceless Palatal Fricative [ʃ] in Central Minnesota English. *Linguistic Portfolios*, *6*(1), 5.

Koffi, E., & Lopez-Backstrom, M. (2018). The Acoustic Correlates of Alveolar Fricatives in the Idiolect of a Northern Minnesota Female. *Linguistic Portfolios*, *7*(1), 5.

Ladefoged, P., and Ian M. 1(996). *The Sounds of World Languages*. Blackwell Publishers Ltd: Malden, MA.

Moore, B. C. (2008). Basic auditory processes involved in the analysis of speech sounds. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1493), 947-963

Stevens, K. N. (2000). *Acoustic phonetics* (Vol. 30). MIT press.

Thomas, Erik R. 2011. Sociophonetics: *An Introduction*. Basingstoke: Pelgrave Macmillan.

Wallin, J., & Koffi, E. (2017). The acoustic correlates of fricatives in whispered speech: an idiolectal analysis. *Linguistic Portfolios*, *6*(1), 9.