ACOUSTIC VARIATION IN MODERN HEBREW FRICATIVES*

MAYA L. BARZILAI
Georgetown University

1 Introduction

This paper presents an acoustic study of fricatives produced by speakers of Modern Hebrew (MH). Intervocalic fricative productions from three male and three female speakers are analyzed with respect to four distinct acoustic cues. Though some results of the current study suggest a change in progress, led by female speakers, in which the voiced fricative /v/ is lenited to a more approximant-like realization, the data also show a wide range of interspeaker variation in MH fricative productions. Results have implications not only for the type of interspeaker variation that exists in production of fricatives, but also for speech perception. While there are several possible cues to contrasts among MH fricatives, different speakers employ unique subsets of these cues when conveying phonological contrasts.

2 Background

2.1 Variation

Research has revealed several social factors conditioning acoustic variation across speakers. One of these factors is the well-attested phenomenon of gender variation. Despite the few but significant physiological differences in the vocal tract across sexes, most acoustic differences between male and female speech across languages cannot be explained by physiology. Therefore, they are necessarily learned sociophonetic behaviors (Johnson, 2006). Such gendered use of features is specific to a given variety of a language, whether defined geographically or by other social factors. For example, Foulkes et al. (2010) show that a speaker's gender can be cued

* I am grateful to the audience at IATL 33 at Tel Aviv University, as well as to Jennifer Nycz, Shannon Mooney, Marcos Rohena-Madrazo, and Georgetown University's PhonLab, for their invaluable help and feedback on this project.

by a specific phonetic variant, but only for listeners who are familiar with the dialect in which that variant is linked to a gender difference. Therefore, in the process of speech perception, the listener relies on social information, such as the speaker's gender, in addition to the raw acoustic information transmitted.

Speech perception is further complicated by the fact that there are often multiple acoustic cues to a given phonological distinction. Denes (1995) shows that in order to perceive word-final obstruent voicing contrasts in English, the listener must rely on the relative duration of the previous vowel, and not merely on the presence of audible voicing in the obstruent. Lisker (1986) presents more extreme findings, showing that at least sixteen acoustic cues are relied upon, to different extents, by American English listeners perceiving the obstruent voicing contrast in 'rapid' and 'rabid'. In cases of multiple acoustic cues participating in the perception of a single contrast, these cues often exist in a trading relation, whereby the absence or reduction of one acoustic cue to a contrast is compensated for by the clear presence of another (Fitch et al., 1980; Best et al., 1981). Upon perceiving multiple cues, listeners have been shown to assign different perceptual weights to acoustic cues that are otherwise interchangeable, with some listeners preferring to attend to one cue and other listeners favoring another (Idemaru et al., 2012; Schultz et al., 2012). Furthermore, the acoustic cue that is assigned the greatest perceptual weight by a given listener is not necessarily the most reliable cue to the contrast in that listener's own acoustic production. These results reveal variability in both speech perception and speech production across speakers. The apparent complexity of the speech perception mechanism implies that speakers are at liberty to use one of many possible acoustic configurations when producing a given contrast.

2.2 Modern Hebrew Fricatives

MH fricatives provide an appropriate opportunity for the investigation of variation in the acoustic signal. Different articulatory considerations about fricative production make conflicting predictions concerning the types of fricatives that are most likely to surface. The high oral pressure required for the production of a fricative is at odds with the low oral pressure required to produce voicing, therefore making voiced fricatives more difficult to produce than voiceless fricatives. On the other hand, when a voiceless fricative appears intervocalically, any gestural misalignment results in at least partial voicing of the fricative (Rohena-Madrazo, 2011).

The extent of voicing, as well as other spectral properties, can be determined using specific acoustic measures. These measures include intensity, harmonics-to-noise ratio (HNR), and center of gravity (CoG), all of which provide different information about the acoustics of fricatives (Gradoville, 2011). HNR represents the amount of periodicity relative to the amount of frication across the duration of the fricative, such that higher values correspond to more voicing. CoG is a weighted average of the frequencies present in a given sound, providing the main cue to the place of articulation of voiceless fricatives. It also indirectly reveals the extent of voicing in a fricative: The frequency of voicing is lower than that of any frication. Therefore, the presence of voicing lowers the average frequency of voiced fricatives, and as a result voiced fricatives have lower CoG values than their voiceless counterparts (ibid.).

MH has a relatively rich fricative inventory, with fricatives at five places of articulation, and voicing contrasts at four of these places, as shown in Table 1.

Labiodental		Alveolar		Postalveolar		Uvular		Glottal
f	v	s	z	ʃ	ʒ	χ	ʁ	h

Table 1. MH (Modern Hebrew) fricative inventory (Bolozky, 1997)

Of the fricatives in Table 1, /ʒ/ appears in MH only in loan words. /h/ often surfaces as a glottal stop or is elided altogether in MH. Though /ʁ/ may surface as a fricative, it has variable realizations and patterns phonologically as an approximant (Bolozky, 1997). Therefore, the present study is restricted to the subset of MH fricatives presented in Table 2.

Labiodental		Alveolar		Postalveolar	Uvular
f	v	s	z	ʃ	χ

Table 2. MH (Modern Hebrew) restricted fricative inventory

Even with the omission of three fricatives from Table 1, the inventory of MH fricatives is substantial, with four places of articulation and two voicing contrasts. Therefore, it provides an ideal test case for variation in acoustic production.

2.3 The Present Study

The purpose of this study is to investigate the nature and extent of acoustic variation in the production of MH fricatives. It is hypothesized that while there may be some level of variation between female and male speakers, variation will also be observed in the production of the different fricatives between individual speakers regardless of gender.

3 Methods

The data analyzed in this study are excerpts from episodes of *Israel Story* (in Hebrew: *sipur isra'eli* 'Israeli story'), a podcast available online that features "everyday stories, told by, and about, regular Israelis".[1] Episodes were downloaded from the podcast website and converted from mp3 to WAV files. The samples selected for analysis from the *Israel Story* corpus were those with high audio quality and minimal background noise. The resulting set of recordings contained speech by three women and three men (aged 33-46), all native speakers of Hebrew with reported normal speech and hearing.

Due to phonological processes attested in MH, such as anticipatory voicing assimilation across consonant clusters (Bolozky, 1997), only intervocalic tokens of fricatives were examined in this study. This resulted in 910 tokens across the six contrastive fricatives and six speakers. Table 3 shows the number of tokens of each segment that were produced by each speaker.

[1] http://israelstory.org/en/about-us/.

	SEGMENT					
SPEAKER	/v/	/z/	/f/	/s/	/ʃ/	/χ/
1M	44	21	10	23	69	47
2M	35	13	10	26	51	44
3M	11	7	6	10	13	9
1F	37	23	8	27	43	26
2F	35	33	9	15	63	55
3F	14	3	8	7	22	33

Table 3. Number of tokens per speaker and segment

TextGrids for each speech sample were segmented manually in Praat (Boersma and Weenink, 2017), by a proficient MH heritage speaker, using visible frication and reduction of the wave's amplitude as acoustic landmarks for fricative segmentation, as seen in Figure 1. Periodicity, the main acoustic cue to voicing, was not used as a guide to segmenting fricatives from their surrounding vowels, as voicing during fricative production was one of the dependent variables measured.

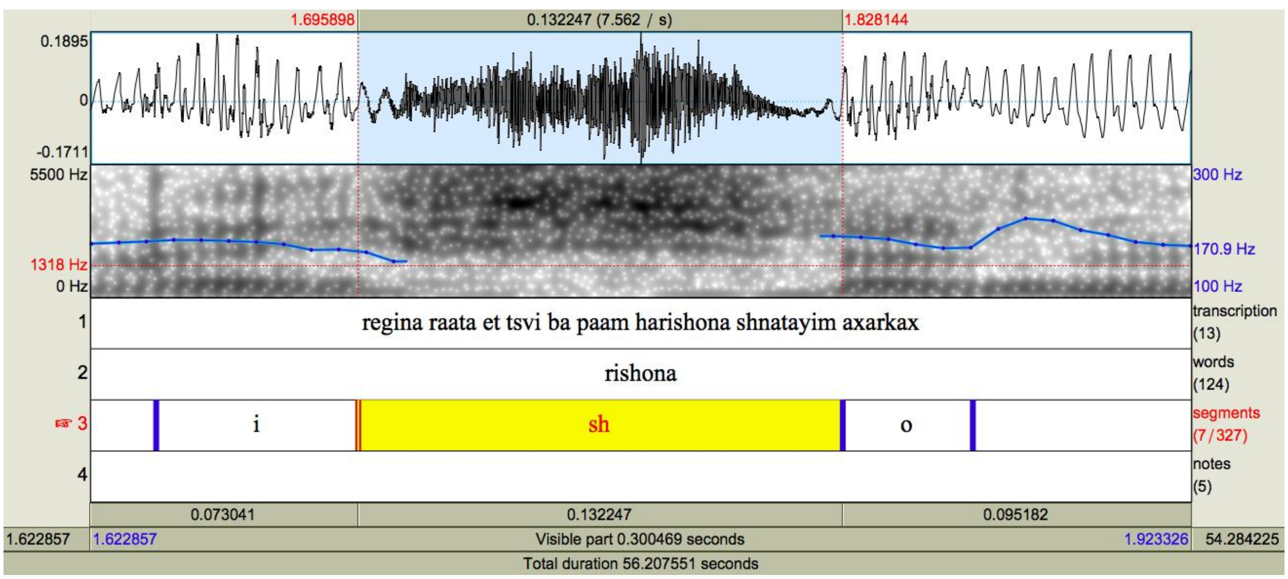

Figure 1. Segmentation of intervocalic fricative based on visible frication and changes in amplitude

For each fricative token, measures of duration, intensity, HNR, and CoG were extracted. Measures of F0 were taken at 21 evenly-spaced points across each token. Following the methodology of Rohena-Madrazo (2011), the number of these points, at which F0 was detected, was then divided by 21. The resulting measure, percent voicing, reflects the percentage of the fricative that was produced with voicing.

A preliminary examination of the data suggested that duration was not a cue to phonological contrasts in any of the fricatives, for any of the speakers. Therefore, only the remaining four measures were analyzed in this study:

1. Percent voicing
2. Intensity
3. HNR
4. CoG

For each fricative, a t-test revealed whether acoustic cues differed by speaker gender. Analysis of variance (ANOVA) tests with post-hoc Tukey HSD tests were used to compare fricative productions across the six speakers.

4 Results

4.1 Gender Variation

A preliminary scan of the acoustic data suggested that, compared to the voiced fricative productions of male speakers, voiced fricatives produced by female speakers showed low levels of frication, with high periodicity during the fricative, as well as a smaller change in intensity between the vowel and fricative portions of the waveform. A t-test on tokens of /v/ across genders confirmed this trend, revealing that the difference in mean HNR values for female and male productions of /v/ was highly significant ($t(168.6) = 4.48$, $p < 0.001$). The mean HNR for female /v/ tokens was approximately 2dB higher than the mean HNR for male /v/ tokens, corresponding to more periodicity and less frication during the fricative productions of female speakers than during those of male speakers. Furthermore, there was a significant difference in intensity between /v/ tokens produced by females and those produced by males ($t(173.87) = 2.41$, $p = 0.017$). The average intensity of female /v/-token productions was higher than that of male /v/ productions, an artifact of fricative productions with high periodicity. The HNR and intensity measures of /v/ productions by male and female speakers are presented in Figure 2.

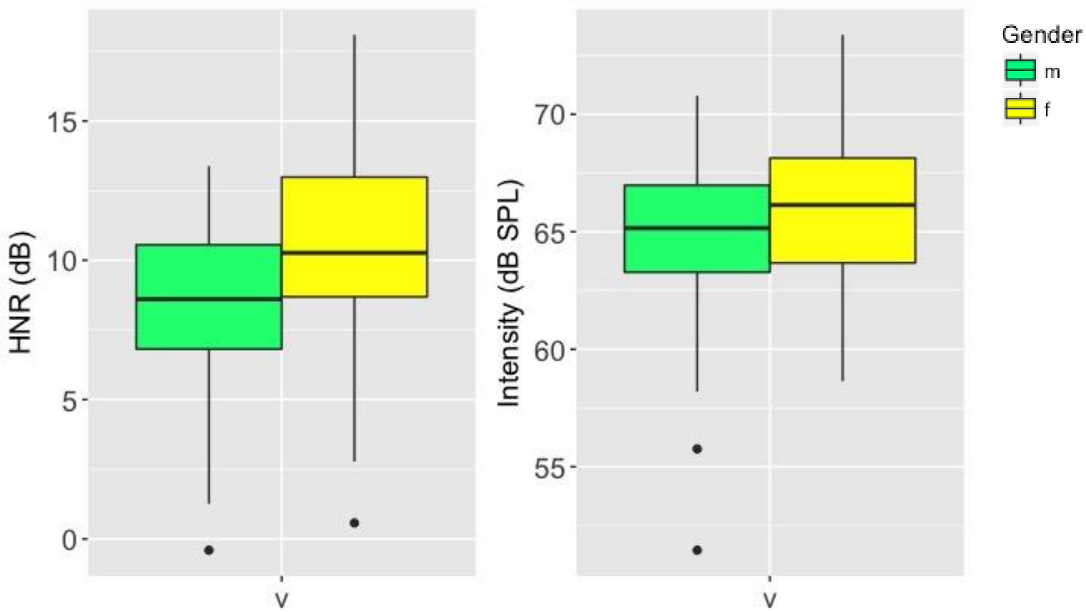

Figure 2. HNR (harmonics-to-noise ratio) and intensity of /v/ tokens by gender. m = male; f = female.

A t-test test revealed a highly significant difference in HNR values for female productions of voiceless fricatives /f s ʃ χ/, compared to male productions, similarly to the results reported for /v/ ($t(605.12) = 6.65$, $p < 0.001$). For each of the voiceless fricatives, female productions showed a higher average HNR value than male productions. This pattern is presented in Figure 3.

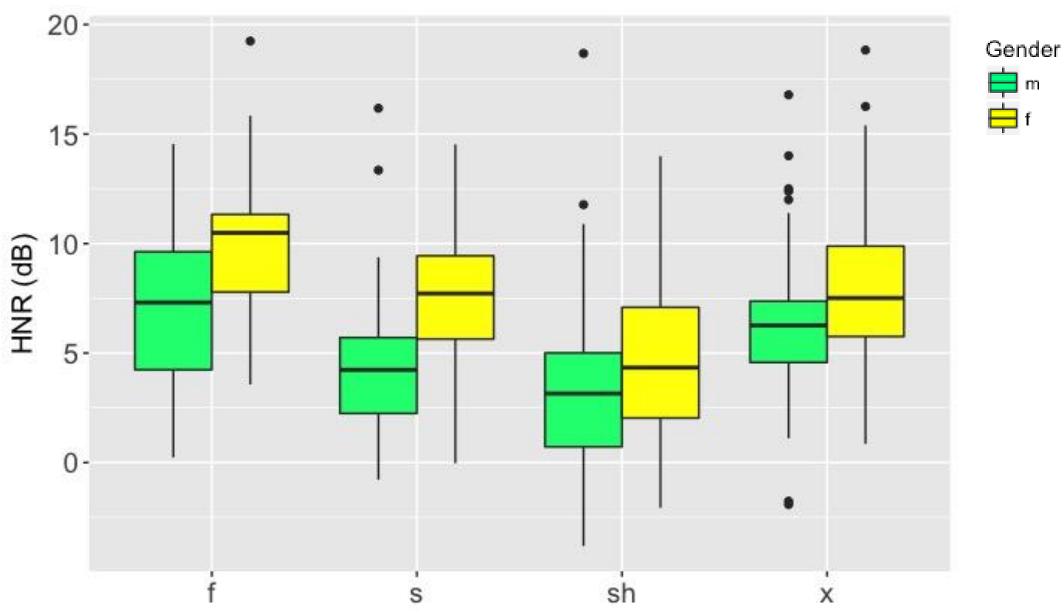

Figure 3. HNR (harmonics-to-noise ratio) of voiceless fricative tokens by gender. sh = ʃ; x = χ; m = male; f = female.

However, a closer look at each speaker's production of fricatives suggests that the trends presented in Figure 3 may in fact falsely emerge as a result of individual variation, which is obscured in a comparison by genders. Figure 4 shows the HNR values for voiceless fricatives by speaker, revealing that the trends observed with respect to gender variation may be coincidental. In fact, the data show wide-ranging interspeaker variation with no clear gender-related patterns.

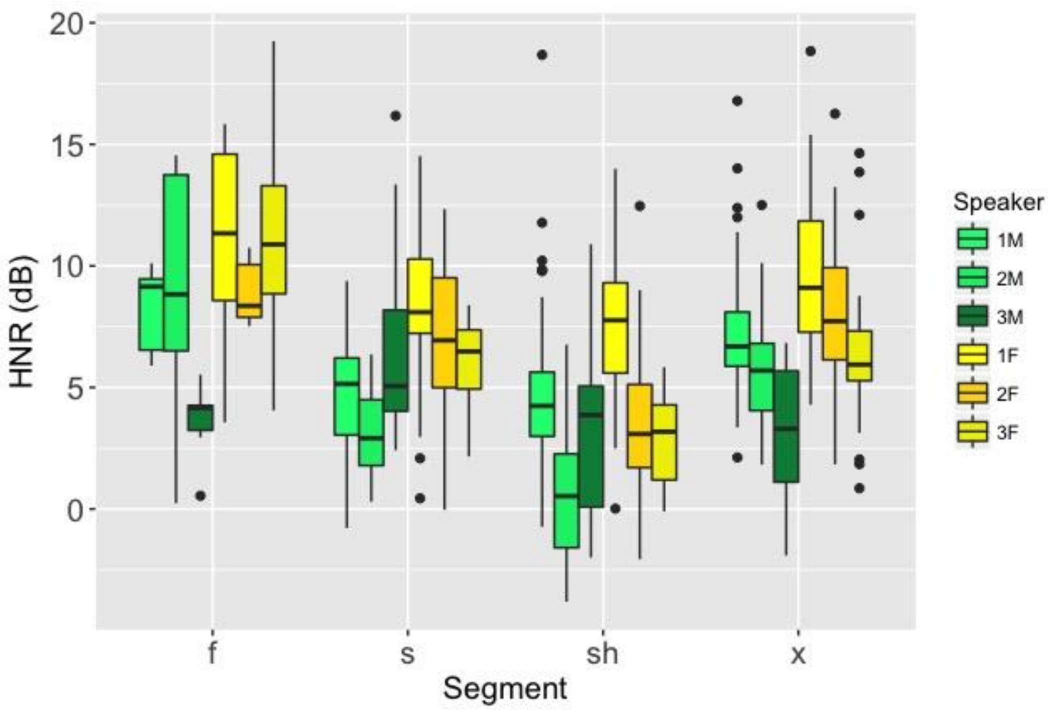

Figure 4. HNR (harmonics-to-noise ratio) of voiceless fricative tokens by speaker. sh = ʃ; x = χ.

4.2 Individual Variation

A closer investigation of individual variation reveals that different sets of cues are responsible for the various fricative contrasts conveyed by different speakers. Table 4 shows the acoustic cues used by each speaker to convey fricative contrasts. Dark green cells represent pairs of acoustic cues in which a significant difference is expected between the pair members based on the phonological contrast between them; light green cells represent phonological contrasts in which a significant difference is neither necessarily expected nor utterly unexpected between the pair members. For the place contrast between the voiced fricatives /v/ ~ /z/ (first column on the left), the only cue that is expected to convey the contrast is CoG, as this measure reveals acoustic differences that result from different places of articulation. Though the presence of voicing is expected to yield an overall low CoG value for both /v/ and /z/, the different place of articulation of each of the two phonemes is nonetheless likely to result in a significant difference in CoG. The next two columns show the cues for the two fricative voicing contrasts in MH: /v/ ~ /f/ and /z/ ~ /s/. Percent voicing, HNR, and CoG are all expected to serve as strong cues to this contrast, as they provide direct evidence of voicing. Intensity, though not inherently correlated with voicing, is not an unlikely cue to this contrast, as increased periodicity in the voiced segments may lead to higher intensity. The six rightmost columns show the cues that convey place of articulation contrasts for pairs of voiceless fricatives. Here, CoG is expected to reach significance in all cases, as it should provide the primary cue to place of articulation of voiced fricatives. Intensity may turn out to be an additional cue, simply because different voiceless fricatives differ in their inherent intensities (Gradoville, 2011).

		Contrast								
Speaker	Cue	/v/-/z/	/v/-/f/	/z/-/s/	/f/-/s/	/f/-/ʃ/	/f/-/χ/	/s/-/ʃ/	/s/-/χ/	/ʃ/ -/χ/
1M	Perc. Voi.	—	*	*	—	—	*	—	*	—
	Intensity	—	*	*	*	*	*	—	—	—
	HNR	—	—	*	*	*	—	—	*	*
	CoG	*	*	*	*	*	*	*	*	—
2M	Perc. Voi.	*	*	*	—	—	*	—	*	*
	Intensity	—	—	—	*	—	—	—	*	—
	HNR	—	—	*	*	*	*	*	*	*
	CoG	*	*	—	—	*	*	*	*	*
3M	Perc. Voi.	—	*	*	—	—	—	—	—	—
	Intensity	—	*	—	—	*	—	—	—	—
	HNR	—	—	—	—	—	—	—	—	—
	CoG	*	*	—	*	*	*	*	*	—
1F	Perc. Voi.	*	*	*	—	—	—	—	—	—
	Intensity	*	*	—	*	*	*	—	—	—
	HNR	—	—	*	—	—	—	—	—	—
	CoG	*	*	—	*	*	*	*	*	*
2F	Perc. Voi.	—	*	*	—	—	*	—	*	—
	Intensity	—	*	*	*	*	*	—	*	—
	HNR	*	—	*	—	*	—	*	—	*
	CoG	*	*	—	*	*	*	*	*	*
3F	Perc. Voi.	—	*	*	—	—	—	—	—	—
	Intensity	—	*	—	*	*	—	*	—	*
	HNR	—	—	—	*	*	*	—	—	*
	CoG	*	*	—	*	—	*	*	*	*

Table 4. Cues to fricative contrasts by speaker

Cells in Table 4 display an asterisk where ANOVA and post-hoc Tukey HSD tests reveal a significant difference in a cue to a given contrast ($p < 0.05$ in all cases); dashes signify that the difference in cue did not reach statistical significance. In general, the differences in cues to phonological contrasts represent acoustic interspeaker variation in fricative production. Though many of the green cells contain check marks, indicating an active cue where expected, there are cases in which expected cues to contrast are not realized. For instance, the only significant cue to the place contrast in the fricative pair /ʃ/ ~ /χ/ for speaker 1M was HNR, suggesting that this speaker produced a different degree of periodicity for each of the two voiceless fricatives. However, this is not directly entailed by the phonological contrast between the two fricatives. More surprising than the difference in HNR is the finding that the two fricatives were not distinguished by CoG in this speaker's productions, though CoG is typically expected to be the main cue to place of articulation contrasts among between fricatives. Similarly, speaker 3F produced differences in intensity and HNR, but not in CoG for the /f/ ~ /ʃ/ contrast; again, CoG was expected to be the main cue. The most extreme case of unexpected acoustic cues to a phonological contrast was observed in speaker 3M's realization of /ʃ/ versus /χ/, for which none of the four measures was significant. On the whole, Table 4 reveals a high degree of interspeaker variation in the acoustic realization of fricative contrasts, and, in several cases, contrasts that were not conveyed by the expected cue.

5 Discussion

5.1 Evidence of Female-led Change

The data reveal one major difference in the fricative production of male and female MH speakers: Female speakers tend to produce /v/ with higher HNR and intensity values than male speakers. The two measures suggest that the female speakers' production is more similar to a labiodental approximant than to a canonical fricative. This trend is not observed in the other voiced fricative in MH, /z/. Nonetheless, given the generalization that female speakers are often the leaders of sociolinguistic change (Labov, 2011), these findings could indicate the beginning of a change in progress.

A sound change such as this could signal the start of a larger pull chain, in which the lenition of /v/ to /β/ would create a gap in the MH phoneme inventory, and in turn would allow for the gradual voicing of /f/ to /v/ (Gordon, 2002). Though the data in this study did not present conclusive evidence of the second change, a larger data set would be able to offer greater insight into patterns of /v/ lenition and possible subsequent changes.

5.2 Interspeaker Variation

The data show a great deal of interspeaker variation, such that speakers produce different combinations of acoustic cues to convey the same phonological contrast. In many instances, the cues to a given contrast are not limited to those that are expected based on the phonological features involved. Conversely, the data reported in this study demonstrate that it is not the case that expected cues will necessarily be used by a speaker in order to convey a certain phonological contrast: Some of the expected cues did not reach significance. Aside from the production of /v/, as outlined above, interspeaker variation does not appear to align with speaker gender. There also do not appear to be any cue trading relations among sets of cues, as has been previously attested in other languages (e.g., Fitch et al., 1980; Best et al., 1981); there are no clear patterns showing that the absence of one cue is consistently compensated for by the presence of another.

The variation observed in this study has wide-ranging implications for speech perception of MH fricatives. Based on the results obtained here, MH listeners must be able to perceive phonological contrasts among fricatives despite unpredictable interspeaker acoustic variation. They must do so by way of a grammar that contains all the possible cues to the phonological contrasts, but that is perceptually flexible enough to accept any subset thereof in accommodating to a given speaker.

5.3 Directions for Future Research

Acoustic studies using data sets comprised of a larger number of speakers would serve to corroborate the results of this study. The size of this data set did not allow for an analysis that controls for either the identity of the vowels surrounding the fricative, or the position of the fricative with respect to stress; these phonological factors may contribute to differences in fricative production that could not be captured here. Given that this study was restricted to the investigation of intervocalic fricatives, any variation in the acoustic properties of fricatives in other prosodic positions remains open to further research. In addition, all of the speech samples studied here were taken from podcast recordings and are therefore of similar register and formality. The investigation of more speech contexts would help determine whether there is in fact a change in progress, as proposed by the preliminary results reported in the current study. Additionally, a closer investigation of intraspeaker variation may reveal that some speakers produce a higher level of variation in fricative productions, while other speakers are more consistent in general.[2]

The cells in Table 4 in which either the expected cues to the contrast did not reach significance, or in which there were no significant cues at all, suggest that other factors may play a role. For instance, in the case of some speakers, spectral cues such as formant transitions may provide the most informative cue to the identity of a fricative. Frequency of a given token and bigram frequency of pairs of segments may also govern the variation seen here to a certain extent; it is possible that words or bigrams with higher frequencies in MH exhibit more variation as a result of this frequency, without any phonetic or phonological explanation.[3]

The questions about speech perception that are raised in this study also call for further investigation. For instance, though female productions of /v/ are more approximant-like than male productions, it remains to be seen whether MH listeners are able to perceive the gender of a given speaker solely on the basis of their /v/ productions. In addition, the extent of interspeaker variation across several acoustic cues allows for a study of cue weight in speech perception, such

[2] Though the type of data examined in this study could be well suited to an analysis of intraspeaker variation, the amount of data did not allow for this type of investigation. Furthermore, a full picture of intraspeaker variation would require the investigation of data from varied speech contexts, which the recordings studied here do not provide.

[3] The degree to which a given phonological contrast is conveyed may also be in part determined by the functional load of the relevant phonological feature for the particular word; in the absence of close lexical neighbors, there is less need for the contrast to be as clearly articulated, as the chance for confusion of misperception is relatively low. Similarly to possible frequency effects, any effect of functional load would be independent of phonetic or phonological factors.

that some listeners may be sensitive to one subset of cues and other listeners – to another subset (Idemaru et al., 2012; Schultz et al., 2012). Future studies investigating perception by MH speakers will be able to confirm whether all of the cues assumed here are in fact active during perception. While each cue measured is attended to in some contrast of some language, it may be the case that, as far as MH is concerned, listeners do not attend to each of them in the perception of fricatives.

6 Conclusion

This study presents evidence of variation in the production of fricatives in MH. Whereas the lenited production of /v/ by female speakers may indicate the beginning of a broader phonological change in progress, the high degree of interspeaker variation found here points to a situation in which different speakers convey the same phonological contrast using distinct subsets of acoustic cues. Evidence is provided that there is not a direct relationship between the features involved in phonological contrasts and the phonetic cues used to convey them. Given these findings, the production of MH fricatives may encode information about speaker gender. The findings also suggest that the acoustic signal contains enough interspeaker variation that the MH listener must accept any combination of several possible cues in accommodating to a given speaker and perceiving the intended phonological contrast.

References

Best, Catherine T., Barbara Morrongiello, and Rick Robson. 1981. Perceptual equivalence of acoustic cues in speech and nonspeech perception. *Perception & Psychophysics* 29.3:191–211.

Boersma, Paul, and David Weenink. 2017. Praat: doing phonetics by computer. *http://www.praat.org/*

Bolozky, Shmuel. 1997. Israeli Hebrew phonology. *Phonologies of Asia and Africa* 1:287–311.

Denes, Peter. 1955. Effect of duration on the perception of voicing. *The Journal of the Acoustical Society of America* 27.4:761–764.

Fitch, Hollis L., Terry Halwes, Donna M. Erickson, and Alvin M. Liberman. 1980. Perceptual equivalence of two acoustic cues for stop-consonant manner. *Perception & Psychophysics* 27.4:343–350.

Foulkes, Paul, James M. Scobbie, and Dominic Watt. 2010. Sociophonetics. *The Handbook of Phonetic Sciences, Second Edition*, ed. William J. Hardcastle, John Laver, and Fiona E. Gibbon, 703–754. John Wiley & Sons.

Gordon, Matthew J. 2002. Investigating chain shifts and mergers. *The Handbook of Language Variation and Change*, ed. J. K. Chambers, Peter Trudgill, and Natalie Schilling-Estes, 244–266. Blackwell Publishing.

Gradoville, Michael Stephen. 2011. Validity in measurements of fricative voicing: Evidence from Argentine Spanish. In *Selected proceedings of the 5th Conference on Laboratory Approaches to Romance Phonology*, ed. Lauren B. Schmidt, Erik W. Willis, and Scott M. Alvord, 59–74. Cascadilla Proceedings Project.

Idemaru, Kaori, Lori L. Holt, and Howard Seltman. 2012. Individual differences in cue weights are stable across time: The case of Japanese stop lengths. *The Journal of the Acoustical Society of America* 132.6:3950–3964.

Johnson, Keith. 2006. Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics* 34.4:485–499.

Labov, William. 2011. *Principles of Linguistic Change, Volume 3: Cognitive and Cultural Factors*. Wiley Blackwell.

Lisker, Leigh. 1986. "Voicing" in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech* 29.1:3–11.

Rohena-Madrazo, Marcos. 2011. *Sociophonetic variation in the production and perception of obstruent voicing in Buenos Aires Spanish*. Doctoral Dissertation, New York University.

Shultz, Amanda A., Alexander L. Francis, and Fernando Llanos. 2012. Differential cue weighting in perception and production of consonant voicing. *The Journal of the Acoustical Society of America* 132.2:EL95–EL101.