

GENDER VARIATION IN CREAKY VOICE AND FUNDAMENTAL FREQUENCY

Shannon Melvin and Cynthia G. Clopper

The Ohio State University
melvin.338@osu.edu, clopper.1@osu.edu

ABSTRACT

The current study examines creaky voice and the fundamental frequency associated with modal and creaky voicing across genders. We analyzed read speech data from five male and five female Midwestern Americans between the ages of 18 and 25. We found that female speakers use creaky voice more often than male speakers in this dialect. Additionally, the male speakers produced a much smaller difference in fundamental frequency between modal and creaky voicing than the female speakers.

We suggest that the difference in fundamental frequency ranges across genders may cause creaky voice to be more perceptually salient in the speech of women than men. As a result, we propose that women use creaky voice more often than men because it is more salient in their speech and may thus more easily carry social meaning.

Keywords: creaky voice, gender, fundamental frequency

1. INTRODUCTION

Creaky voice is a mode of voicing characterized by aperiodicity and a drop in F0 [6, 7]. Some languages, such as Jalapa Mazatec and Kashaya Pomo, employ creaky voice to create phonemic contrasts [3]. In English, however, creaky voice has been observed as a both a sociophonetic marker [3] and a prosodic marker [2, 8].

Creaky voice has previously been found to be a characteristic of masculine speech. In a study of two dialects of British English, experimenters found that creaky voice was more often used by male participants than by female participants, and thus proposed that creaky voice is more likely to be associated with the speech of men than women [3]. In a more recent study of young American women, however, experimenters found that creaky voice is also a characteristic of feminine speech [9]. Because Wolk et al. [9] only examined the speech of women, however, more research is necessary to determine the relationship between gender and creaky voice in American English. One such study examining the connection between creaky voice and gender was

conducted by Yuasa [10], who observed that female speakers of Californian English produce creaky voice more frequently than their male counterparts. Yuasa [10] also found that creaky voice was perceived as educated and professional, and suggested that women use creaky voice to project that particular social image. In the current study, we find additional evidence that women use creaky voice more frequently than men in a comparison of the frequency of creaky voicing in male and female Midwestern American English speech.

In the current study, we also examined the relationship between gender and speakers' ranges in fundamental frequency for both modal and creaky voicing. In a study examining the ranges in fundamental frequency employed by men and women across multiple phonation registers, Hollien and Michel [5] found that women have a wider and higher range in fundamental frequency during modal voicing than men, but that both genders have a very similar range in fundamental frequency during creaky voicing. These results suggest that the difference in fundamental frequency between modal and creaky voicing is larger for women than for men. However, in their study, participants were asked to imitate frequencies that were presented to them as pure tones and from these data, Hollien and Michel [5] described the possible ranges in fundamental frequency that can be produced by men and women in both modal and creaky voicing. However, because their study examined the imitation of tones rather than natural speech, the frequency ranges they observed may not be representative of normal speech production. Additionally, their study aimed to find the modes of voicing corresponding to specific frequencies, rather than the frequencies corresponding to identifiable modes of voicing, as our study does.

Hillenbrand and Houde [4] found that the most significant perceptual cue for creaky voice is the lowering of F0. We hypothesize that, because of the greater difference in F0 between modal and creaky voicing for women than men, the usage of creaky voice is more perceptually salient in the productions of women than men, leading women to produce creaky voice more frequently than men. Thus, in our study we examined the frequency ranges for men and women in modal and creaky voicing to draw

connections between those ranges and the social meanings associated with the frequency with which male and female speakers employ creaky voicing [10].

2. METHODS

2.1. Speakers

The speech data used in this study were collected from five males and five females ranging in age from 18-25 years old. All are native speakers of the Midland dialect of American English, which is spoken in the American Midwest. Each participant is a self-reported monolingual speaker of American English with no reported history of a speech or hearing impairment.

2.2. Passages

In this study, we examined speech production in five read passages, which are each approximately 60 seconds in duration. The speakers were asked to read the passages from a computer screen one at a time as if talking to a friend. They were recorded in a sound-attenuated booth using high-quality digital recording equipment.

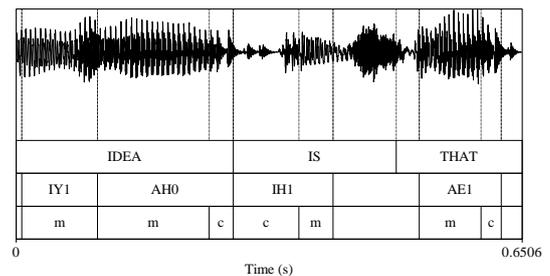
2.3. Voice quality analysis

Using Praat, each syllable nucleus in the recordings was labelled by hand as creaky, modal, neither creaky nor modal (i.e., another voice quality), deleted, or disfluent. We identified creaky voicing as speech with a perceptibly lower F0 and visibly aperiodic waveform. Non-creaky, non-modal voice quality was identified as speech with otherwise atypical quality but without the characteristics of creaky voice. Deleted nuclei were identified as nuclei which had become devoiced or otherwise unarticulated during production. Speech marked as disfluent included utterances during which the speaker coughed, laughed, mispronounced words, or did not read the passage accurately. If multiple modes of voicing were used for a single syllable nucleus, each mode of voicing was segmented and labelled appropriately, as illustrated in Figure 1.

In Figure 1, the waveform of the utterance is accompanied by three tiers of annotation: the words that were produced in standard orthography, the vowels in a plain-text phonetic transcription including stress, and the voice quality labels that we assigned to each vowel. The label “m” codes for modal voicing, while the label “c” codes for creaky voicing. Figure 1 illustrates a series of syllable nuclei with more than one mode of voicing. In the second syllable of “idea,” the vowel is mostly modal

with a small portion of creaky voice at the end. The following vowel in “is” begins with creaky voicing, but becomes modal for the ending portion. The vowel in “that” is similar to the final vowel in “idea”: it is mostly modal with a small portion of creaky voice at the end before a final glottal stop.

Figure 1: Example of a series of syllables labelled with more than one mode of voicing. Modal voicing is indicated by “m” and creaky voicing is indicated by “c” in the third annotation tier.



A total of 12,793 syllable nuclei were included in our final analysis, with 1237-1320 syllables per speaker. For each nucleus, the voice quality label(s), the mean F0, the start time, and the end time for each labelled segment were extracted. Syllables labelled as neither creaky nor modal, as deleted, or as disfluent were excluded from the analysis, as were those that contained any part labelled as neither creaky nor modal. The analysis is therefore based on syllables that were fully modal, fully creaky, or a combination of modal and creaky voiced.

Figure 2 shows productions of the word “greed” in modal and creaky voice for one of the female speakers. The vowel in the top panel is evidently modal, with a periodic waveform. In contrast, the bottom waveform illustrates creaky voicing, with the vowel characterized by several distinct glottal pulses.

Similarly, modal and creaky productions by one of the male speakers are shown in Figure 3. In the top panel, the waveform illustrates a modal production of the word “sea.” In contrast, the bottom panel shows the word “seas” produced by the same male speaker with creaky voicing. As in the comparison shown in Figure 2, the creaky production exhibits fewer, more distinct glottal pulses than the modal production, consistent with the lower F0 and aperiodicity characteristic of creaky voice.

Figure 2: Productions of “greed” with modal voicing (top) and creaky voicing (bottom) by the same female speaker.

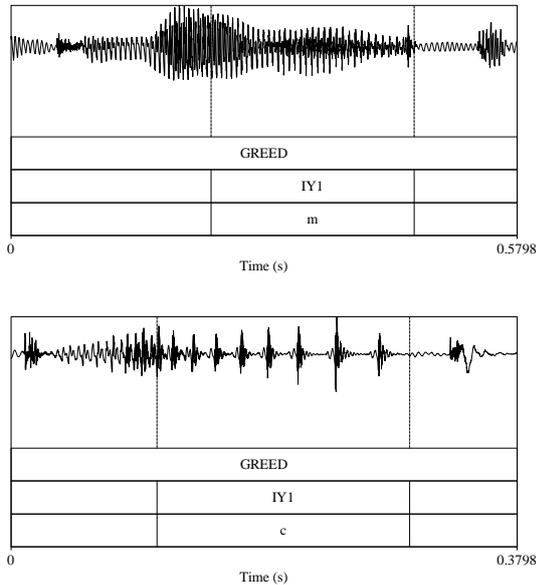
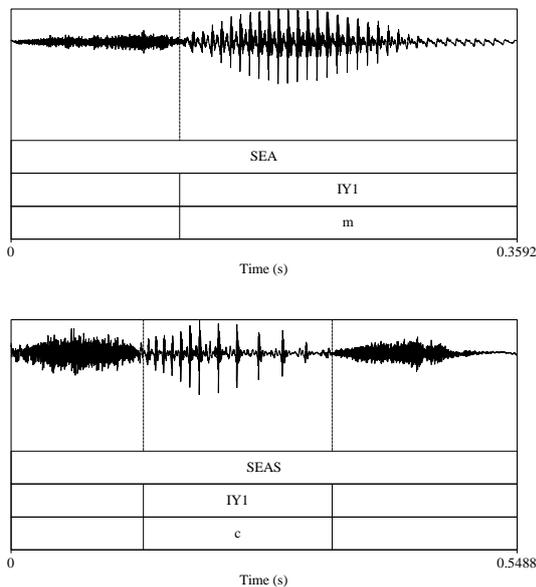


Figure 3: Productions of “sea” and “seas” in modal voicing (top) and creaky voicing (bottom) by the same male speaker.



3. RESULTS

To explore the effect of gender on the overall rate of creaky voicing, we calculated the proportion of syllables containing creaky voice out of the total number of modal and creaky syllables produced by each speaker. Syllables with periods of both modal and creaky voice were treated as creaky syllables in this analysis. Table 1 shows the mean proportion of syllables containing creaky voice for each gender. As shown in Table 1, the female speakers produced

a higher proportion of syllables with creaky voice than the male speakers. To confirm this interpretation of the data, we conducted a two-sample t-test comparing the proportion of syllables containing creaky voice across gender. We found that gender had a marginally significant effect on the proportion of syllables containing creaky voice ($t(8) = 2.04, p = 0.076$), suggesting that the women in our sample used creaky voicing somewhat more often than the men.

Table 1: Mean proportion of syllables containing creaky voice for male and female speakers. Standard deviations are shown in parentheses.

Gender	Proportion of creaky syllables
Male	0.124 (0.045)
Female	0.183 (0.046)

To explore the effect of gender on the implementation of creaky voice within the syllable, we examined the subset of syllables for each speaker that contained creaky voicing, including those syllables that were fully creaky and those that had more than one mode of voicing within a single nucleus (see Figure 1). For each syllable containing creaky voicing, we calculated the total duration of creaky voicing within the syllable and compared that measure to the total duration of the syllable to obtain a measure of the proportion of creaky voicing within the syllable. The mean proportion of creaky voicing within the syllable is shown in Table 2 for each gender. The male and female speakers produced almost identical proportions of creaky voice within the syllable. A two-sample t-test comparing the proportion of creaky voice within the syllable across gender confirmed that there is no significant difference in the proportion of creaky voice within creaky syllables across men and women ($t(8) = -0.078, p = 0.94$). Thus, in our sample the implementation of creaky voice did not differ across genders.

Table 2: Mean proportion of creaky voice within syllables for male and female speakers. Standard deviations are shown in parentheses.

Gender	Proportion of creaky voice within syllables
Male	0.907 (0.068)
Female	0.904 (0.065)

Finally, we examined the mean fundamental frequency difference between modal and creaky voicing for each gender. For each speaker we calculated the mean fundamental frequency during modal voicing, as well as the mean fundamental frequency during creaky voicing. We then subtracted the mean creaky F0 from the mean modal F0 to find

the mean difference in F0 between the two modes of voicing for each speaker. Table 3 shows the mean difference in F0 between the two voice qualities for male and female speakers. As shown in Table 3, there is a greater difference in F0 between the voicing modes for the female speakers than for the male speakers. This difference was confirmed by a two-sample t-test, which revealed a significant difference in mean F0 between creaky and modal voicing for men and women ($t(8) = 7.58, p < 0.001$).

Table 3: Mean difference in F0 between modal and creaky voicing for male and female speakers. Standard deviations are shown in parentheses.

Gender	F0 Difference (Hz)
Male	36.80 (13.93)
Female	89.80 (7.11)

4. DISCUSSION

Our results suggest that women produce creaky voice more frequently overall than men, although the implementation of creaky voice within the syllable does not differ across genders. These results suggest that Yuasa’s [10] observations of Californian American English extend and apply to Midland American English and that creaky voice may be a sociophonetic marker amongst women in the United States more generally. This result contrasts with the results obtained by Henton and Bladon [3], who found that creaky voice was a characteristic of the speech of men. The differences across studies may be explained by several factors. First, Henton and Bladon’s [3] study focused on two dialects of British English, while ours and Yuasa’s [10] examined dialects of American English. Thus, creaky voice may have different social meanings in the USA and the UK. Second, Henton and Bladon’s [3] study was conducted almost three decades ago. Although creaky voice had previously been associated with masculine speech, it is possible, as Yuasa [10] hypothesized, that women have come to produce creaky voice in imitation of masculine speech as a way to assert themselves in a male-dominated society. Thus, a change in the social meaning of creaky voice may have occurred over the last three decades, so that creaky voice is now a social marker of female speech.

Unlike Wolk et al. [9], we found that creaky voice is not a mode of speech exclusive to female speakers. In fact, we found only a modest difference in overall proportion of syllables containing creaky voicing for men and women, and the men and women in our sample used similar proportions of creaky voice within the syllables that they produced

with creaky voicing. That is, while women were observed to use creaky voice more often overall, the utterances containing creaky voice produced by the men did not differ in their implementation from those containing creaky voice that were produced by the women. Therefore, despite the modest difference in the overall amount of creaky voicing used by each gender, we suggest that the usage of creaky voice as a phonational register is similar within syllables across gender.

Finally, we found that the difference between the mean fundamental frequencies for modal voicing and creaky voicing is much larger for women than for men by approximately 53 Hz. This difference in fundamental frequency between the two types of voicing for male and female speakers could provide an additional explanation for the relationship between the frequency of usage of creaky voice and gender. As noted above, Hillenbrand and Houde [4] found that the most significant perceptual cue for creaky voice is the lowering of F0. The greater difference between modes for the female speakers than the male speakers may contribute to greater perceptual salience of creaky voice in the productions of women than men. This greater salience may allow women to use creaky voice to convey social meanings, such as those proposed by Yuasa [10], more easily than men, leading to the change in social meaning from a masculine [3] to a feminine [10] characteristic.

The speakers were recorded reading the same set of passages in this study to constrain the segmental and prosodic context in which we explored creaky voice. However, prosody is variable even in read speech [1]. Because prosody is also a documented predictive factor for the use of creaky voice [2, 8], the next stage of this research will involve considerations of prosodic structure in the realization of creaky voice across genders. Given that creaky voicing has been linked to particular prosodic environments [2, 8], it is possible that there are more prosodic environments in which women tend to use creaky voice than men, lending some additional explanation for why creaky voice is observed more in the speech of women than men overall. Further studies should examine the combined effects of creaky voicing and prosody in the speech of men and women to explore this possibility.

5. ACKNOWLEDGMENTS

This work was partially supported by an Undergraduate Research Scholarship from the College of Arts and Sciences at Ohio State

University and the National Science Foundation (BCS-1056409).

6. REFERENCES

- [1] Clopper, C.G., Smiljanic, R. 2011. Effects of gender and regional dialect on prosodic patterns in American English. *Journal of Phonetics* 39, 237-245.
- [2] Dilley, L., Shattuck-Hufnagel, S., Ostendorf, M. 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24, 423-444.
- [3] Henton, C., Bladon, A. 1988. Creak as a sociophonetic marker. In: Hyman, L.M., Li, C.N. (eds). *Language, Speech and Mind: Studies in Honour of Victoria A. Fromkin*. London: Routledge, 3-29.
- [4] Hillenbrand, J.M., Houde, R.A. 1996. Role of F0 and amplitude in the perception of intervocalic glottal stops. *Journal of Speech and Hearing Research* 39, 1182-1190.
- [5] Hollien, H., Michel, J.F. 1968. Vocal fry as a phonational register. *Journal of Speech and Hearing Research* 11, 600-604.
- [6] Gordon, M., Ladefoged, P. 2001. Phonation types: A cross-linguistic overview. *Journal of Phonetics* 29, 383-406.
- [7] Ladefoged, P., Maddieson, I. 1996. *The Sounds of the World's Languages*. Oxford: Blackwell.
- [8] Redi, L., Shattuck-Hufnagel, S. 2001. Variation in the realization of glottalization in normal speakers. *Journal of Phonetics* 29, 407-429.
- [9] Wolk, L., Abdelli-Beruh, N.B., Slavin, D. 2012. Habitual use of vocal fry in young adult female speakers. *Journal of Voice* 26, e111-e116.
- [10] Yuasa, I.P. 2010. Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? *American Speech* 85, 315-337.