REPORT ON VOCAL FRY IN INTERACTIONAL CONTEXTS
CREAKY VOICE AND PITCH AS AFFECTED BY AGE AND GENDER OF SPEAKER

SOPHIA SIX

University of Colorado Boulder

Previous research on vocal fry has been largely one-sided, investigating its role as a vocal phenomenon for young women specifically, and focusing on countering or confirming harmful stereotypes surrounding vocal fry as perpetuated by the media. This study attempts to expand an understanding of vocal fry by investigating its production by multiple genders and ages, as well as its role in an interactional context. In this study, participants of four age and gender groups (older women, older men, younger women, and younger men) were asked to conduct a conversation with a member of each identity in order to determine the effect of identity in an interactional role on the occurrence of vocal fry. While prior studies have implied that the use of vocal fry is most heavily dependent on the identity of the speaker, this study concludes that the identity of the participant may actually play a more important role in its occurrence. It also contextualizes the prominence of female vocal fry in relation to the frequency of vocal fry by other speakers.

Keywords: vocal fry, creaky voice, interaction, sociolinguistics, gender, age, sexism, phonetics

1. INTRODUCTION

Vocal fry has long been vilified by experts from every field and casual listeners alike as sounding “less competent, less educated, less trustworthy, less attractive, and less hirable”, or even “vulgar”, “repulsive”, “mindless”, and “really annoying” (Anderson & Klofstad 2014; Garfield 2013). Many critics also uphold the persistent claim that it is damaging to one’s vocal cords. Speech pathologist Susan Sankin infamously claimed in 2015 that “if [vocal fry] is a repetitive habit that you use over a long term … the vocal cords will show some sort of fatigue” (Gross). One of the early studies on vocal fry published in the Journal of Voice in 2011 by Wolk and Abdelli-Beruh also claimed that “the habitual use of fry is atypical and possibly a form of vocal abuse”. It is not difficult to notice that young women are almost exclusively the target of these criticisms, vocal fry being associated with women like Kim Kardashian and Britney Spears, and therefore, by extension, with the negative stereotypes surrounding the type of women who use it (Anderson & Klofstad 2014; Chao & Bursten 2021). In reality, while the phenomenon does occur most frequently in young women, nearly every demographic uses vocal fry, regardless of age and gender. A 2015 study by Callier and Podesva of Stanford University found that strong creak patterns also occur with young men and older women, and older men use some creak but with a different pattern of intonation. A 2016 study from the Journal of the Acoustical Society of America by Sarah Irons and Jessica Alexander found that young men may even use vocal fry more often than young women. Regardless, young people, especially women, are overwhelmingly the ones criticized for it. After the Journal of Voice published its pivotal paper “Habitual Use of Vocal Fry in Young Adult Female Speakers” in 2011, vocal fry suddenly became a viral household term. A flurry of controversial and angry thinkpieces attacking vocal fry as an unprofessional, intolerable, and damaging “fashion choice” unique to young women overpowered the media, far too many in number to begin to list here, and the issue remains controversial to this day (Reynolds 2015). In an ironic catch-22, while women with a naturally higher pitch of voice are seen as less competent and less authoritative, women who lower their voice using vocal fry are also seen as being less confident and less educated (Anderson & Klofstad 2014).

Contrary to popular belief, vocal fry is not damaging to the vocal cords, as can be proven by the fact that creaky voice is used as a contrastive linguistic feature in several languages, such as Danish, Burmese, Kwakw’ala, and Jalapa Mazatec, in which a word can have a different semantic meaning depending on its vocal quality (Gordon & Ladefoged 2001; Keating et al. 2015). For example, Jalapa Mazatec contrasts between modal (normal) voice, creaky voice, and breathy voice, já (modal) meaning “tree”, já̤ (breathy) meaning “he wears”, and já̰ (creaky) meaning “he carries” (Gordon & Ladefoged 2001). If there are entire cultures that use creak every day to contrast semantic meaning, it cannot be unhealthy or damaging, and far less a speech disorder, as some would call it. Research on the occurrence of vocal fry, like this study, can serve to dispel harmful stereotypes about the phenomenon, reduce relevant sexist biases and increase its normalcy, and further understand its role in linguistics.

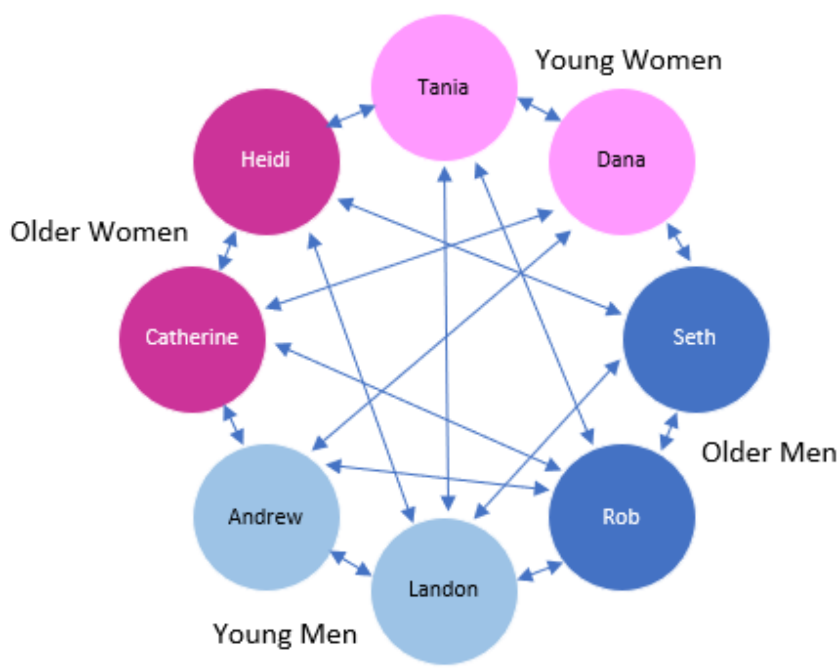
Vocal fry is a type of creaky voice, or sometimes used as another term for vocal creak. It is produced when the glottal closure is loose and relaxed at a very low pitch, allowing air to vibrate through slowly and irregularly, with audible pulses that sound like a rattle or croak. In this study, I will be measuring vocal fry by measuring the harmonics and pitch of vowels. Vocal fry should correlate with a greater difference in amplitude between the first and second harmonics, and a lower fundamental frequency (Keating et al. 2015). This experiment will study the effect of age and gender on vocal fry in an interactional context. I have designed interviews pairing speakers of different age and gender demographics (older women, younger women, older men, and younger men) to determine whether a speaker's use of vocal fry is affected by the age and gender of their conversational partner, as well as whether their own age and gender affects their vocal quality.

My first hypothesis is that fry will be more common among young speakers. This is supported by existing research that finds it to be a relatively new phenomenon among the younger generation (Anderson & Klofstad 2014; Wolk & Abdelli-Beruh 2011). Because young women are the primary users of vocal fry, and older generations are particularly offended by its usage as an undesirable feminine trait, I suspect that there will be a difference in which gender uses creak more between generations (Anderson & Klofstad 2014). My second hypothesis is that among young people, women will use vocal fry more, and among older people, men will use vocal fry more, as older women and younger women will draw a larger distinction between themselves in this matter, with differing definitions of a desirable vocal quality. I believe, however, that age will be the more important factor, and thus young females will use the greatest amount of vocal fry, but will be followed by young men before older men. Due to recent studies indicating a strong use of vocal fry among young men, I do not believe there will be an exceptionally large difference between young men and young women in their use of vocal fry, and there will be a bigger difference between older men and older women; older women using hardly any (Irons & Alexander 2016). My third hypothesis is that age of the conversational partner will be the biggest factor in determining an individual speaker's use of vocal fry. Because vocal fry has become increasingly prevalent as a social phenomenon in recent years, I hypothesize that all speakers will use more vocal fry when speaking with young partners. My fourth hypothesis is that I believe that differences due to a partner's gender will experience a shift across generations. Older speakers will use more vocal fry when speaking to women, and younger speakers will use more vocal fry when speaking to men. Thus, young women speaking to young men will use the most vocal fry, and older men speaking to other older men will use the least. This inference comes from a suspicion (and vague implication made by some of the research) that older generations view vocal fry in women as unattractive due to a lower pitch being seen as more masculine, and younger generations seeing a lower (or more "masculine") pitch as a positive trait holding more gravitas (Anderson & Klofstad 2015; Chao & Bursten 2021). Thus, an older man would perhaps be more likely to use vocal fry to confirm his masculinity in distinguishing himself from the more "effeminately" higher-pitched opposite sex (and therefore an older woman would do the opposite with a male partner, but may relax around another woman), while a younger female may use it to increase her status among male peers (with young males also conducting themselves oppositely). I also expect there to be some differences in the amount of vocal fry used based on the position of the speech sound within an utterance, or over time in an interaction. My fifth hypothesis is that I believe that vocal fry will occur at higher levels and a greater rate at the end of a phrase or utterance, and that it will increase slightly over the course of an interaction. I will be measuring vowels, but I do not expect there to be a difference in creak dependent on vowel quality, except perhaps for more creak in mid/central vowels (schwa), as I expect vocal fry to very frequently occur on the filler word "uh", as speakers relax their vocal cords.

2. METHODS

In order to collect data for this project, I arranged interviews between subjects of varying ages and genders. I solicited eight participants, half of them male and half female. Half were under 25 and half were over 40, in order to establish a noticeable age gap. I have changed participants' names here in order to protect their privacy. Each participant met for an interview with four other participants, one from each age and gender demographic. See Figure 1.

FIGURE 1. PARTICIPANTS & CONVERSATION PAIRING



Most of the participants did not previously know one another. Due to logistics issues, I did have to make two pairings between people that had established relationships. Heidi is the mother-in-law of Tania's sister, so the two are friendly, but have not spent enough time together to have a very close relationship. Catherine and Rob are a married couple, however. I was hoping to avoid pairings between people with a close personal relationship, but unfortunately I could not pair them with alternate participants for scheduling reasons. None of the other participants had previously met. All speakers are native speakers of American English and live in Colorado. All identify as Caucasian, but Landon also has some Latino background and Dana has some Cherokee background. Tania is from Washington, Dana is from Tennessee, Seth is from Maine, Rob and Catherine are from New York, Landon is from Texas, and Andrew and Heidi are from California. Each of their idiolects reflects a mix of features reflective of their home state as well as a standard Coloradan accent.

Participants met over Zoom. I supervised and recorded sessions but remained muted during the interviews. I asked participants to turn on their cameras during the interviews. I gave participants a script ahead of time to interview one another with. See Figure 2. The reason for a scripted set of interview questions was in order to a) maintain topic consistency so that I would have words or phrases to examine that would serve as minimal pairs repeated by a speaker in every session, and b) to combat the awkward silences inevitable when meeting a stranger for the first time, and personal differences in talkativeness and conversational productivity. I chose questions that would be mostly surface level and easy to talk about, but would allow speakers to briefly monologue during their answers. I instructed participants to take turns completing each overarching topic before switching roles. For example, Participant 1 would ask Participant 2 questions 1, 1a, 1b, and 1c, then Participant 2 would ask Participant 2 the same questions before they moved to the second set of questions. I encouraged participants to speak naturally and comfortably and to interject, react, or go off topic whenever it felt appropriate, and I warned them not to script or write down their answers. Meetings usually lasted 15-20 minutes, although a couple of the more talkative participants went off topic and lasted 30-40 minutes.

FIGURE 2. INTERVIEW QUESTIONS

1. What do you do for work? *(alternatively: a previous job/dream job)*
    a. What does your job entail?
    b. What is the hardest part of your job? Why?
    c. What is the most fulfilling part of your job? Why?
2. Did you go to college? *(If not, would you have liked to? Answer the rest hypothetically, with what you would like to study if you went to college)*
    a. Where did you go?
    b. What did you study?
    c. What did you find the most interesting about your studies/field of interest?
3. If salary/education/skills were not an issue, what would your dream career be? Why?
    a. Where in the world would you live if you had your dream job?
    b. What would you do in your free time?
4. Talk about your pets! *(alternatively: a favorite past pet or dream pet)*
    a. Name, species, breeds, gender, age, etc.
    b. What is a silly habit your pet has?
    c. (Feel free to share photos or model them on camera if you want to!)
5. Wrapping up: Got any fun plans for the rest of the day?

With eight participants each meeting with four of the others, this culminated in 16 meetings. I recorded audio separately for each partner, so the result was 32 sound files. In order to take a small sample of the data, I segmented 13 sections of each recording, roughly equivalent to how many relevant turns they took to answer the given questions. Then I began to measure utterance-initial and -final vowels. In order to avoid the complication of segmenting utterances in a given turn, I simply marked an entire multiple-utterance answer and measured the beginning and end of each answer. In this way, I was able to ensure that I was not making any errors mismarking the beginning or end of an utterance.

I marked 3 vowels at the beginning and 3 at the end of each section. This added up to 78 vowels in each recording, or 2,496 total measurements. I generally marked the first or last three vowels, in order, but often skipped highly reduced vowels in favor of more clear vowels with a longer duration. I chose to include common filler and grammar words as long as they were not exceptionally reduced. Most speakers used the same few words at the edges of utterances, such as “um,” “probably,” “yeah,” “so,” “that,” “I,” “like,” and “would.” In spite of this, I was able to collect a reasonable distribution of a variety of vowels, albeit with a higher degree of front vowels and slightly more high vowels. The vowel distribution for each speaker did not seem notably unequal. See Figures 3-5.

FIGURE 3. VOWEL DISTRIBUTION

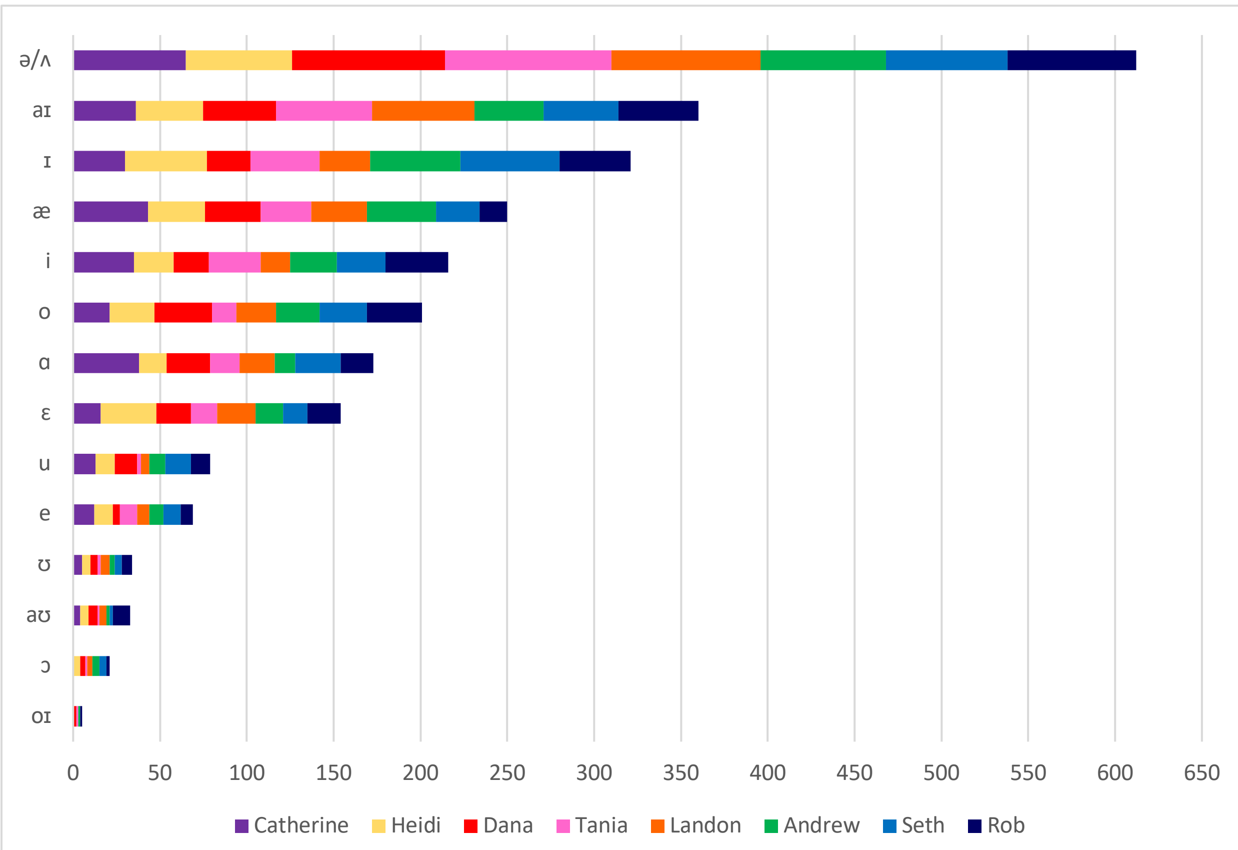


FIGURE 4. FRONTING

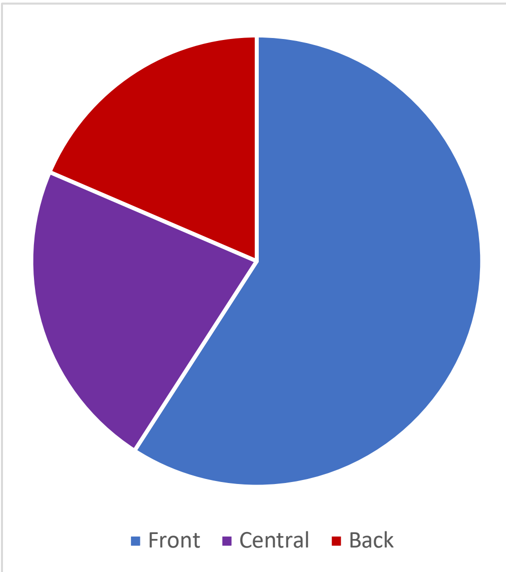
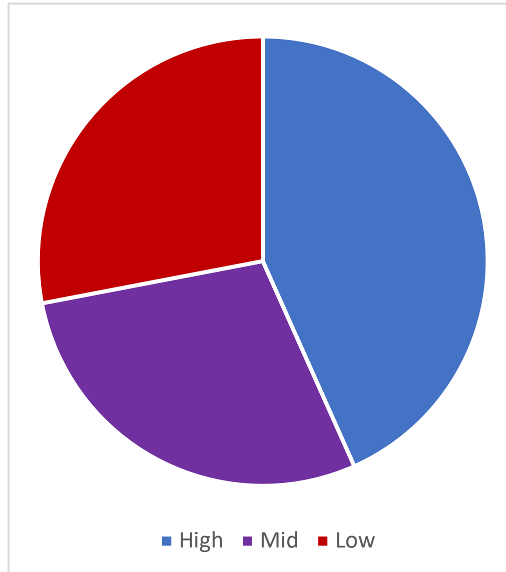


FIGURE 5. HEIGHT



I ran a nasality script (Styler, Scarborough, Johnstone, et al. 2014) on the annotated sound files in order to collect harmonic and formant amplitudes for comparison. The values I pulled from the script were the amplitude of harmonic 1 (H1), harmonic 2 (H2), formant 1 (A1), formant 3 (A3), and the fundamental frequency for pitch (f0). I subtracted H2, A1, and A3 from H1 respectively, expecting that more negative values would correlate with a higher degree of creak, and when it occurred with a low f0, may indicate vocal fry. I ended up focusing primarily on the H1-H2 value along with f0, but the H1-A1 and H1-A3 values served as verification of my methods, as they seemed to follow a similar pattern. I also divided each recording into three sections based on the time point of the first and final measurement of each speaker, and recorded whether the sound occurred early, mid, or late in the session.

3. RESULTS

Women had a lower mean H1-H2 value than men, at -8.409 (sd = 13.07) for women as compared to -4.716 (sd = 10.67) for men. With statistical significance being regarded as $p < 0.05$, this difference was found to be statistically significant at $p < 0.001$. Women had a higher mean f0 value than men, at 196.305 (sd = 44.63) as compared to 126.835 (sd = 39.24). This was statistically significant at $p < 0.001$. Thus, women showed more evidence of creak, but men displayed a lower pitch. Although one would expect creak and low pitch to occur together in vocal fry, this is not surprising, as men have a biological precedent for lower pitch overall. See Table 1 and Figures 6-7.

TABLE 1. GENDER

Speaker	**H1-H2**		**f0**	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**
Female	-8.409	13.066	196.305	44.631
Male	-4.716	10.672	126.835	39.236

FIGURE 6. H1-H2 BY GENDER

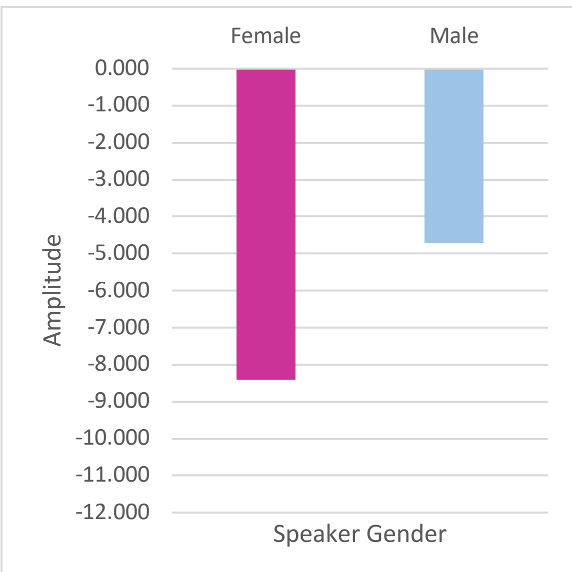
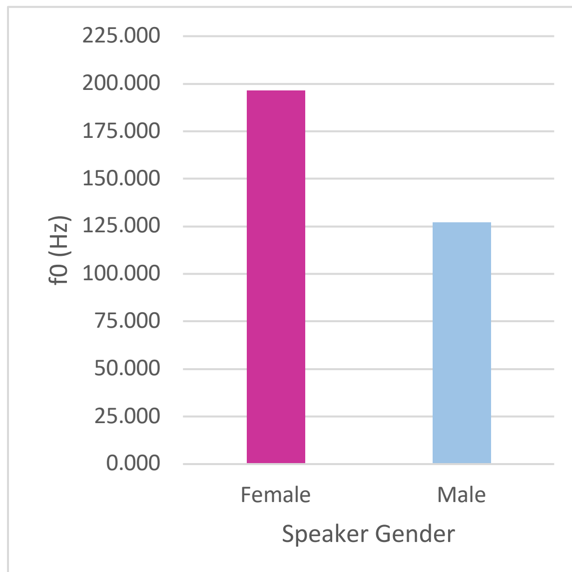


FIGURE 7. PITCH BY GENDER



Younger speakers showed a slightly lower mean H1-H2 (-6.714, sd = 12.36) as compared to the older group (-6.427, sd = 11.78) and a lower mean f0 value (159.752 Hz, sd = 56.23 vs 163.659 Hz, sd = 52.72). With statistical significance being regarded as $p < 0.05$, the difference in H1-H2 between age groups was not found to be significant at $p = 0.55$. However, there was a statistical difference between groups in f0 at $p = 0.02$. Thus, younger speakers can be said to have a lower pitch than older speakers, but there is no significant difference in creak. See Table 2 and Figures 8-9.

TABLE 2. AGE

Speaker	H1-H2		f0	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**
Older	-6.427	11.780	163.659	52.715
Younger	-6.714	12.364	159.752	56.234

FIGURE 8. H1-H2 BY AGE

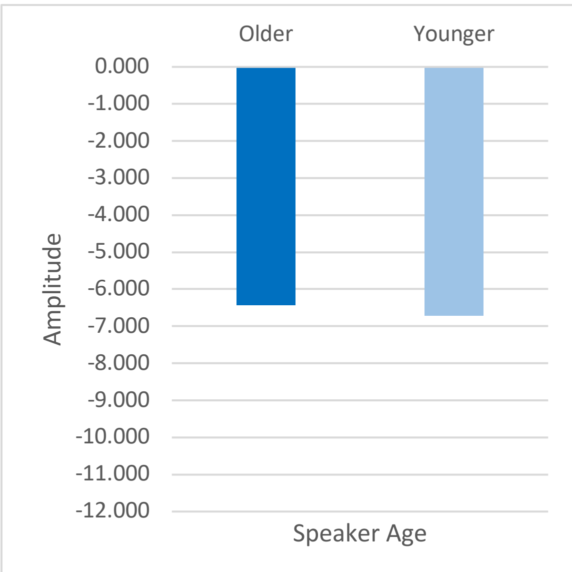
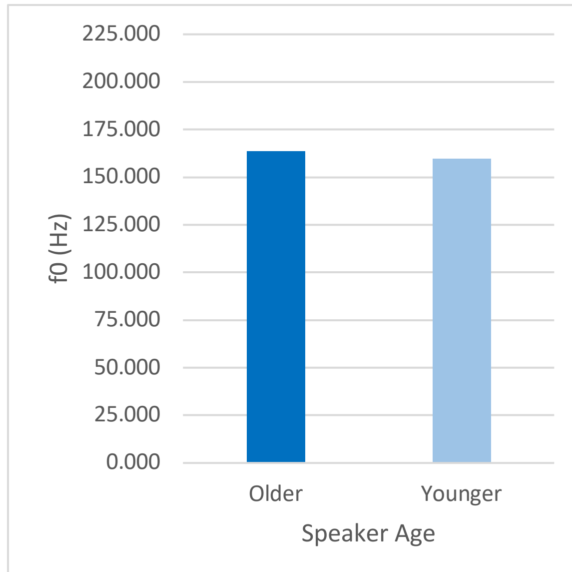


FIGURE 9. PITCH BY AGE



Young women had the lowest mean H1-H2 value (-10.530, sd = 14.84), and young men had the highest (-2.886, sd = 7.51). There was little difference between men (-6.55, sd = 12.84) and women (-6.306, sd = 10.63) in the older category, with men being slightly lower. With statistical significance being regarded as $p < 0.05$, differences between almost all values were found to be statistically significant at $p < 0.001$, except between older women and older men, in which case $p = 0.71$. Thus, young females can be said to have a higher level of creak than older speakers and young males, and young males show less creak than either group, but there is no significant difference in creak between genders in the older group. Gender again played a bigger role in f0 values, with men showing the lowest mean f0 values and women showing the highest. Age was also consistent, with younger speakers showing the lower mean f0 value in their gender category. Therefore, young men had the lowest value at 123.942 Hz (sd = 38.34), then older men at 129.73 Hz (sd = 39.93), younger women at 195.45 Hz (sd = 47.93), and older women having the highest value at 197.16 Hz (sd = 41.12). All differences crossing gender boundaries were found to be statistically significant at $p < 0.001$. The difference between young men and older men was also found to be significant at $p = 0.003$, but there was not found to be a significant difference between young women and older women ($p = 0.53$). Thus, young men can be said to have the lowest pitch, followed by older men, and women have a higher pitch with no difference dependent on age. See Table 3 and Figures 10-11.

TABLE 3. GENDER AND AGE

Speaker	**H1-H2**		**f0**	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**
Older Female	-6.306	10.634	197.156	41.116
Older Male	-6.550	12.844	129.732	39.932
Younger Female	-10.530	14.836	195.448	47.933
Younger Male	-2.886	7.506	123.942	38.342

FIGURE 10. H1-H2 BY AGE AND GENDER

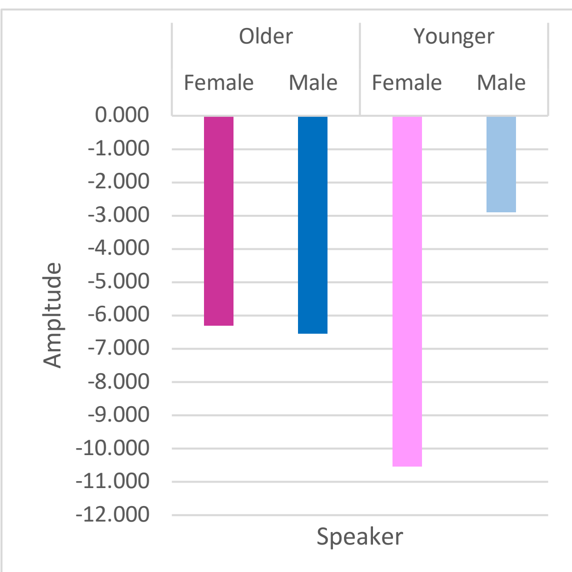
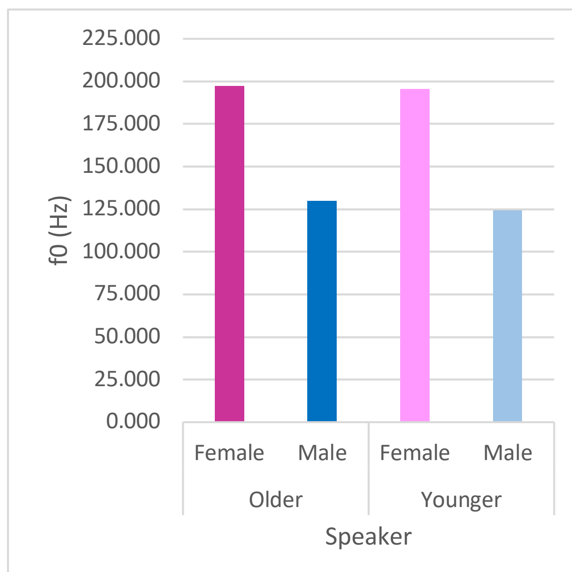


FIGURE 11. PITCH BY AGE AND GENDER



Among older women, those speaking to those in the same age group showed the lowest mean H1-H2 values. Values were lower when speaking to women than to men. When partnered with an older female the mean H1-H2 value was the lowest at -8.031 (sd = 10.93), followed by older male partners at -6.536 (sd = 9.21), younger female partners at -5.589 (sd = 12.02), and younger male partners resulted in the highest mean at -5.114 (sd = 10.06). Mean f0 values did not follow a particular pattern, with older female partners resulting in the lowest mean f0 value at 193.526 Hz (sd = 44.09), followed by younger male partners at 194.099 Hz (sd = 44.03), younger female partners at 198.833 Hz (sd = 39.26), and the highest mean f0 value occurring with older male partners at 202.282 Hz (sd = 36.18). With statistical significance being regarded as $p < 0.05$, differences between partner demographics were mostly insignificant. The only statistically significant differences occurred when comparing H1-H2 of older female and young male partners, for which $p = 0.02$, and when comparing f0 of older female and older male partners, for which $p = 0.03$. Thus, it can be said that partner demographic does not affect vocal fry in older women, except that there is more creak when the speaker is paired with the same demographic (older women) as compared to the opposite demographic (younger men), and gender seems to have an effect on pitch among same-age partners, causing lower pitch when speaking with older women than with older men. See Tables 4-5 and Figures 12-13.

TABLE 4. OLDER FEMALES

Partner	**H1-H2**		**f0**	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**
Older Female	-8.031	10.926	193.526	44.091
Older Male	-6.536	9.210	202.282	36.180
Younger Female	-5.589	12.020	198.833	39.255
Younger Male	-5.114	10.064	194.099	44.032

TABLE 5. PAIRED TWO-TAILED T-TEST COMPARING CONVERSATIONAL PARTNERS AMONG OLDER WOMEN SPEAKERS

	Older female vs young male	**Older female vs young female**	**Older female vs older male**	**Young female vs older male**	**Young female vs young male**	**Older male vs young male**
H1-H2	p = 0.07	p = 0.02	p = 0.16	p = 0.41	p = 0.72	p = 0.17
f0	p = 0.21	p = 0.98	p = 0.03	p = 0.36	p = 0.27	p = 0.07

FIGURE 12. H1-H2 IN OLDER WOMEN

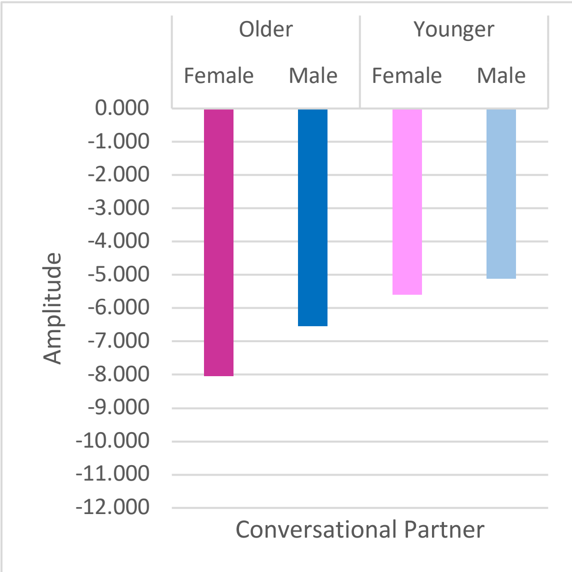
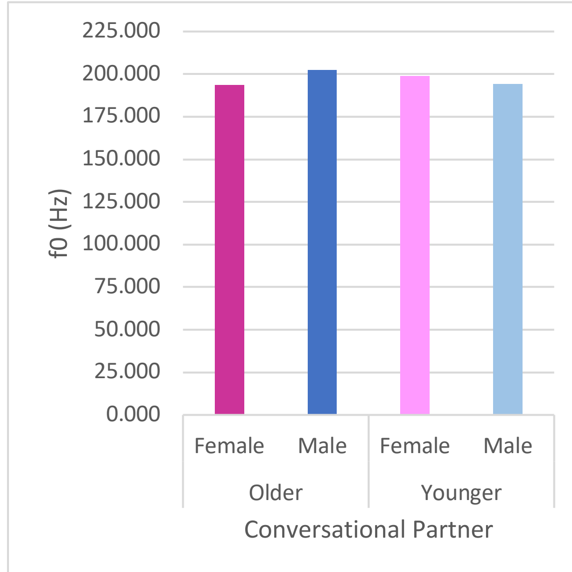


FIGURE 13. PITCH IN OLDER WOMEN



Among younger women, there was no particular pattern in H1-H2 values. When partnered with an older female the mean H1-H2 value was the lowest at -11.869 (sd = 14.64), followed by younger male partners at -10.325 (sd = 14.67), younger female partners at -10.263 (sd = 14.69), and older male partners resulted in the highest mean at -9.664 (sd = 15.385). Mean f0 values were lowest when speaking to another female, especially a younger female, and highest with a male, especially an older male. Young female partners resulted in the lowest mean f0 value at 191.255 Hz (sd = 49.42), followed by older female partners at 193.25 Hz (sd = 45.07), young male partners at 195.942 Hz (sd = 46.94), and the highest mean f0 value occurring with older male partners at 201.372 Hz (sd = 49.98). With statistical significance being regarded as $p < 0.05$, differences between partner demographics were mostly insignificant. The only statistically significant differences occurred when comparing f0 between young female and older male partners, for which $p = 0.047$. Thus, it can be said that partner demographic does not affect vocal fry in young women, except that they use a lower pitch with the same demographic (young women) as compared with the opposite demographic (older men). See Tables 6-7 and Figures 14-15.

TABLE 6. YOUNGER FEMALES

Partner	H1-H2		f0	
	Mean	Std. Dev.	Mean	Std. Dev.
Older Female	-11.869	14.643	193.250	45.071
Older Male	-9.664	15.385	201.372	49.982
Younger Female	-10.263	14.694	191.255	49.424
Younger Male	-10.325	14.665	195.942	46.938

TABLE 7. PAIRED TWO-TAILED T-TEST COMPARING CONVERSATIONAL PARTNERS AMONG YOUNG WOMEN SPEAKERS

	Older female vs young male	Older female vs young female	Older female vs older male	Young female vs older male	Young female vs young male	Older male vs young male
H1-H2	p = 0.29	p = 0.26	p = 0.09	p = 0.59	p = 0.98	p = 0.63
f0	p = 0.65	p = 0.59	p = 0.09	p = 0.047	p = 0.39	p = 0.34

FIGURE 14. H1-H2 IN YOUNGER WOMEN

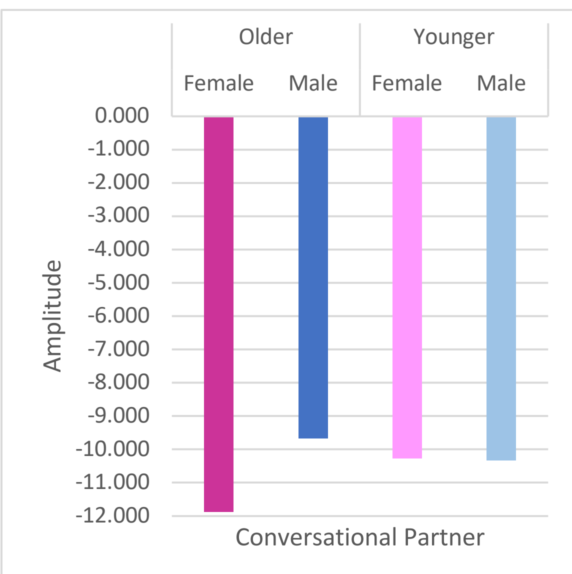
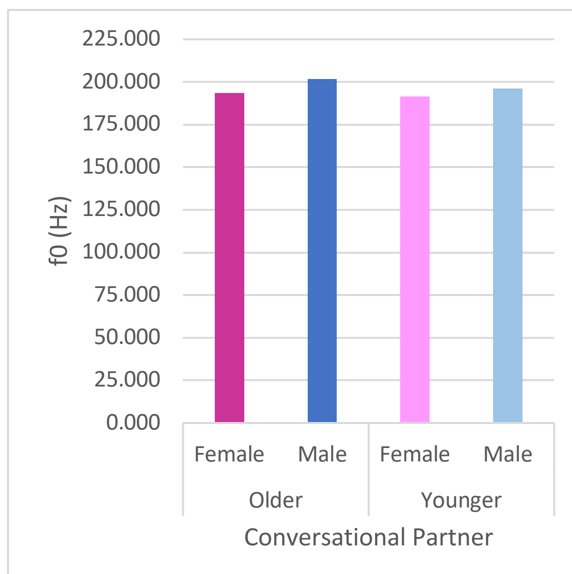


FIGURE 15. PITCH IN YOUNGER WOMEN



Among younger men, H1-H2 values were lowest when speaking with another male. When speaking to a male, H1-H2 was lower with young men than older men, and when speaking to a female, it was lower with older women than younger women. When partnered with a young male the mean H1-H2 value was the lowest at -3.872 (sd = 7.4), followed by older male partners at -2.813 (sd = 6.35), older female partners at -2.752 (sd = 8.03), and young female partners resulted in the highest mean at -2.101 (sd = 8.08). Mean f0 values were lowest when speaking to a young partner, and highest with older partners. Young male partners resulted in a lower f0 value than young female partners, and older female partners resulted in a lower f0 value than older male partners. Young male partners resulted in the lowest mean f0 value at 118.058 Hz (sd = 34.24), followed by young female partners at 122.665 Hz (sd = 39.32), older female partners at 125.244 Hz (sd = 35.61), and the highest mean f0 value occurring with older male partners at 129.795 Hz (sd = 43.02). With statistical significance being regarded as $p < 0.05$, differences between partner demographics were mostly insignificant. The only statistically significant differences occurred when comparing f0 between older female and young male partners, for which p = 0.03, and between young males and older males, for which p = 0.001, and when comparing H1-H2 between young males and young females, for which p = 0.047. Thus, it can be said that partner demographic only affects vocal fry for young men when paired with the same demographic (young men), which produces more creak, as compared to young women, and pitch is lowered when partnered with the same demographic as compared to older partners. See Tables 8-9 and Figures 15-16.

TABLE 8. YOUNG MALES

Partner	**H1-H2**		**f0**	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**
Older Female	-2.752	8.025	125.244	35.609
Older Male	-2.813	6.354	129.795	43.021
Younger Female	-2.101	8.081	122.665	39.317
Younger Male	-3.872	7.403	118.058	34.237

TABLE 9. PAIRED TWO-TAILED T-TEST COMPARING CONVERSATIONAL PARTNERS AMONG YOUNG MALE SPEAKERS

	Older female vs young male	**Older female vs young female**	**Older female vs older male**	**Young female vs older male**	**Young female vs young male**	**Older male vs young male**
H1-H2	p = 0.39	p = 0.23	p = 0.94	p = 0.39	p = 0.05	p = 0.15
f0	p = 0.52	p = 0.03	p = 0.22	p = 0.10	p = 0.27	p = 0.01

FIGURE 15. H1-H2 IN YOUNG MEN

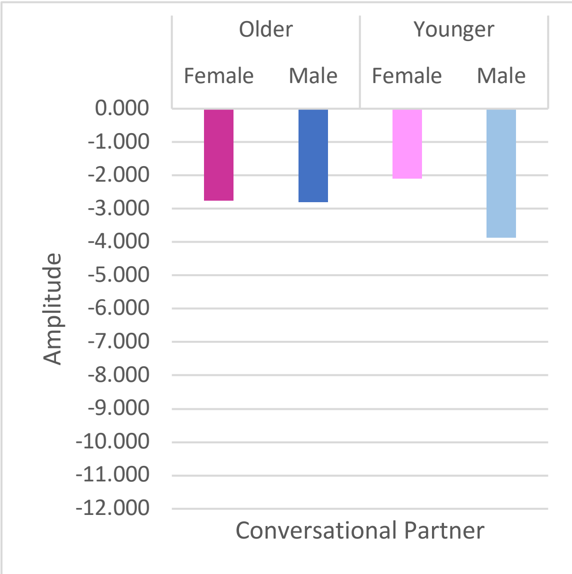
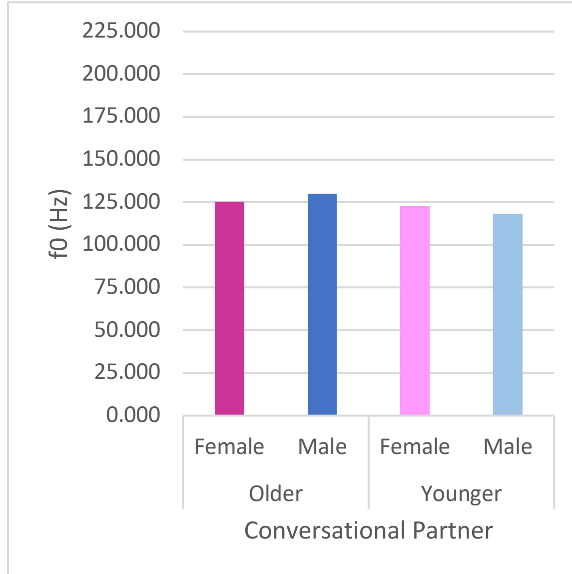


FIGURE 16. PITCH IN YOUNG MEN



Among older men, there was no particular pattern in H1-H2 values. When partnered with a young male the mean H1-H2 value was the lowest at -7.123 (sd = 12.56), followed by older female partners at -6.598 (sd = 13.51), older male partners at -6.421 (sd = 12.34), and young female partners resulted in the highest mean at -6.05 (sd = 13.04). Mean f0 values were lowest when speaking to a female partner, and lower among young partners than older partners. Young female partners resulted in the lowest mean f0 value at 121.994 Hz (sd = 43.34), followed by older female partners at 128.936 Hz (sd = 40.45), young male partners at 130.06 Hz (sd = 40.02), and the highest mean f0 value occurring with older male partners at 137.833 Hz (sd = 34.21). With statistical significance being regarded as $p < 0.05$, H1-H2 differences between partner demographics were all insignificant. However, most f0 comparisons did result in statistically significant results. Older men used a significantly lower pitch with older females than with older males ($p = 0.01$), with young females as compared to older ($p < 0.001$) or younger ($p = 0.03$) males, and with young males as compared to older males ($p = 0.02$). Thus, it can be said that partner demographics do not affect creak in older men, but there is a significant effect on pitch. Older men use the lowest pitch when speaking to women as compared to men, and the highest pitch when speaking to the same demographic (older men) as compared to older women or young men. See Tables 10-11 and Figures 17-18.

TABLE 10. OLDER MALES

Partner	**H1-H2**		**f0**	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**
Older Female	-6.598	13.511	128.936	40.453
Older Male	-6.421	12.342	137.833	34.209
Younger Female	-6.050	13.036	121.994	43.337
Younger Male	-7.123	12.560	130.064	40.021

TABLE 11. PAIRED TWO-TAILED T-TEST COMPARING CONVERSATIONAL PARTNERS AMONG OLDER MALE SPEAKERS

	Older female vs young male	**Older female vs young female**	**Older female vs older male**	**Young female vs older male**	**Young female vs young male**	**Older male vs young male**
H1-H2	p = 0.56	p = 0.61	p = 0.85	p = 0.76	p = 0.25	p = 0.48
f0	p = 0.08	p = 0.71	p = 0.01	p < 0.001	p = 0.03	p = 0.02

FIGURE 17. H1-H2 IN OLDER MEN

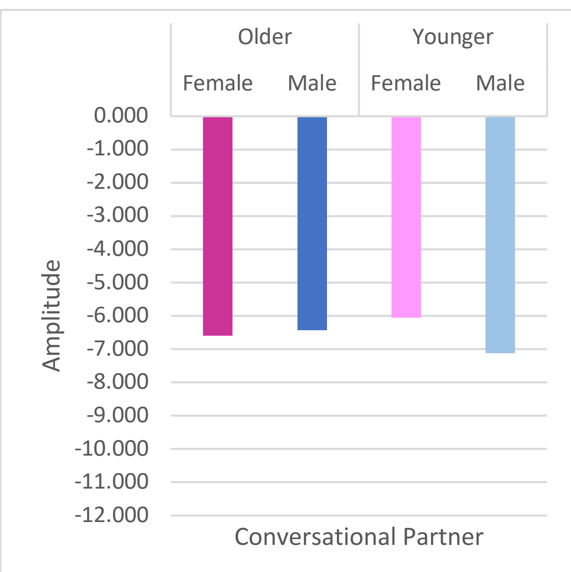
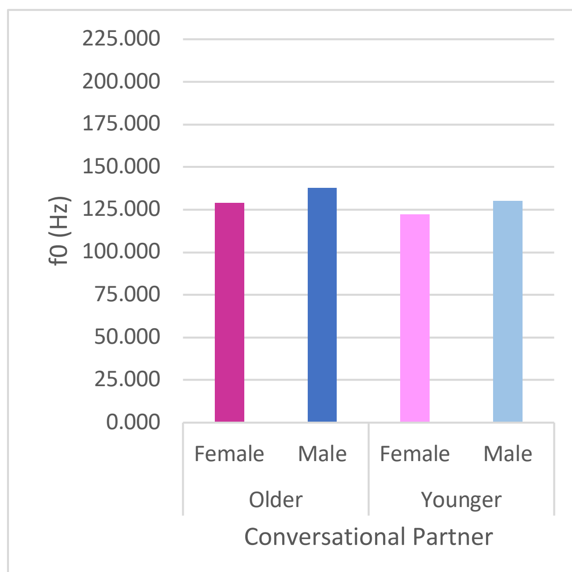


FIGURE 18. PITCH IN OLDER MEN



I also measured the effect of a sound's position in an utterance on H1-H2 and f0. With statistical significance being regarded as $p < 0.05$, an early position has a significantly lower ($p < 0.001$)

mean H1-H2 value at -8.082 Hz (sd = 13.19), as compared to a late position, measuring at -5.058 Hz (sd = 10.63). A late position, however, has a significantly lower ($p < 0.001$) mean f0 value, at 151.832 Hz (sd = 53.62), as compared to an early position, measuring at 171.586 Hz (sd = 53.64). Thus, it can be said that an early position in an utterance indicates more creak, but higher pitch, whereas a late position in an utterance indicates less creak but a lower pitch. See Table 12 and Figures 19-20.

TABLE 12. UTTERANCE POSITION

Position	**H1-H2**		**f0**	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**
Early	-8.082	13.191	171.586	53.642
Late	-5.058	10.634	151.832	53.623

FIGURE 19. H1-H2 BY UTTERANCE POSITION

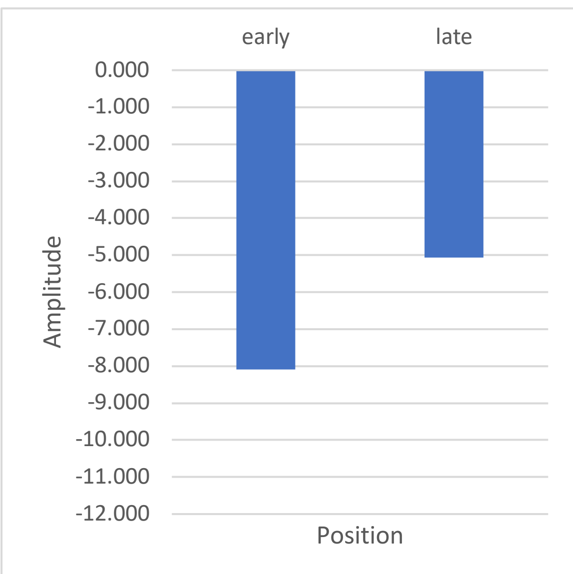
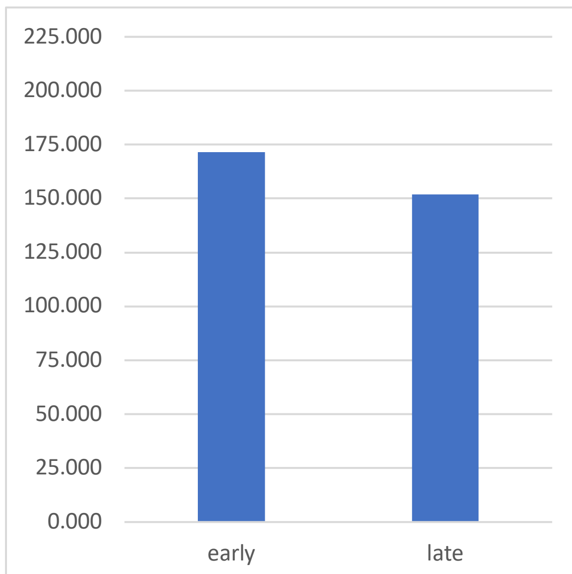


FIGURE 20. PITCH BY UTTERANCE POSITION



I also measured the effect of a sound's position in a recorded session on H1-H2 and f0. With statistical significance being regarded as $p < 0.05$, an early position has a significantly lower ($p < 0.001$) mean H1-H2 value at -7.287 Hz (sd = 11.47), as compared to a late position, measuring at -5.376 Hz (sd = 12.47). There is not a significant difference in H1-H2 between an early and mid position (p = 0.62), but a mid position has a significantly lower (p = 0.005) mean H1-H2 value at -6.984 (sd = 12.283) as compared to a late position. There is also a significant difference in f0 values dependent on session position, with an early position having a significantly lower (p = 0.01) mean f0 value at 160.106 (sd = 53.89) as compared to a late position, measuring at 164.294 (sd = 56.05). There is not a significant difference in f0 between an early and mid position (p = 0.15), but a mid position has a significantly lower (p < 0.001) mean f0 value than a late position. Thus, it can be said that creak decreases and pitch increases over the course of a recorded session. See Table 13 and Figures 21-22.

TABLE 13. SESSION POSITION

Position	**H1-H2**		**f0**	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**
Early	-7.287	11.468	160.106	53.892
Mid	-6.984	12.283	160.874	53.580
Late	-5.376	12.469	164.294	56.048

FIGURE 21. H1-H2 BY SESSION POSITION

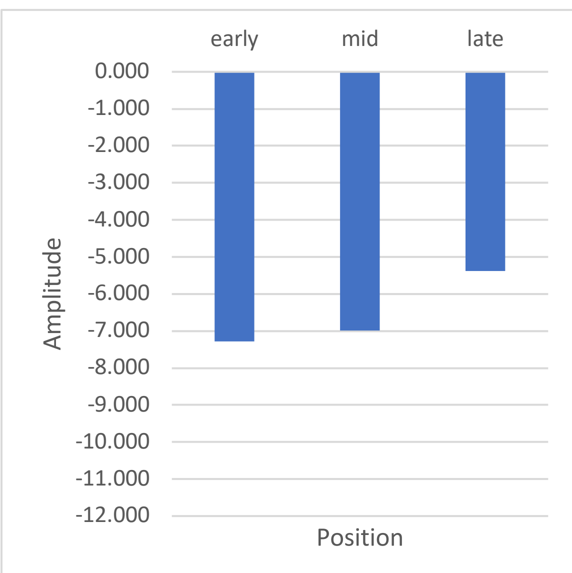
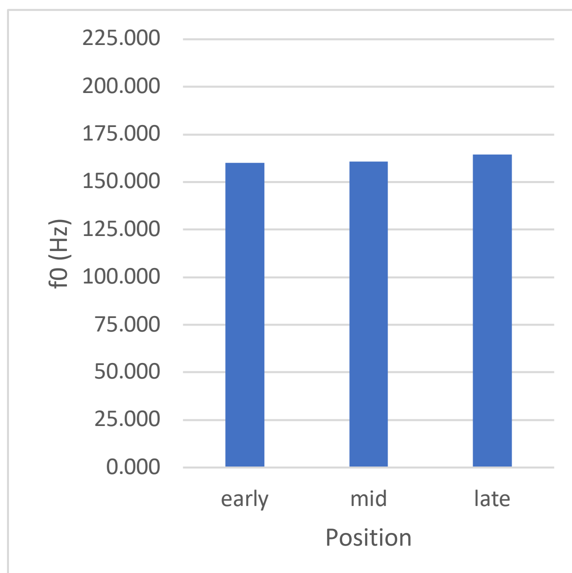


FIGURE 22. PITCH BY SESSION POSITION



When measuring the difference in H1-H2 according to vowel quality, mean values decrease with backing and raising of the vowel. Back vowels have the lowest mean H1-H2 value when comparing vowel fronting at -7.749 (sd = 11.04), followed by central vowels at -6.918 (sd = 12.52) and the highest value occurring in front vowels at -6.04 (sd = 11.76). High vowels have the lowest mean H1-H2 value when comparing vowel height at -7.454 (sd = 11.69), followed by mid vowels at -6.925 (sd = 12.53) and the highest value occurring in low vowels at -4.612 (sd = 10.92). When measuring f0 values according to vowel fronting, there was not a particular pattern obvious in the data. Central vowels had the lowest mean f0 at 124.551 Hz (sd = 35.942), followed by front vowels at 125.584 Hz (sd = 37.68), and back vowels having the highest f0 values at 129.952 Hz (sd = 45.38). When comparing height, mean f0 values lowered with lowering of the vowel. Low vowels had the lowest mean f0 value at 123.588 Hz (sd = 40.55), followed by mid vowels at 124.166 (sd = 35.89), and high vowels at 126.512 (sd = 41.42). When regarding statistical significance as $p < 0.05$, there was not found to be any significant difference in f0 changes across vowel quality. However, the difference in H1-H2 between back vowels and front vowels was found to be significant at $p = 0.02$, as well as between mid vowels and back vowels at $p = 0.001$ and between high vowels and low vowels at $p < 0.001$. Thus, it can be said that while vowel quality has no effect on pitch, creak increases in back vowels as compared to front vowels, and significantly increases as a vowel is raised. See Tables 14-16 and Figures 23-26.

TABLE 14. VOWEL MEASUREMENTS BY FRONTING

Fronting	**H1-H2 Mean**	**Std. Dev.**	**f0 Mean**	**Std. Dev.**
Front	-6.040	11.756	125.584	37.681
Central	-6.918	12.527	124.551	35.942
Back	-7.749	11.035	129.952	45.376

TABLE 15. VOWEL MEASUREMENTS BY HEIGHT

Height	**H1-H2 Mean**	**Std. Dev.**	**f0 Mean**	**Std. Dev.**
High	-7.454	11.685	126.512	41.416
Mid	-6.925	12.529	124.166	35.891
Low	-4.612	10.920	123.588	40.546

FIGURE 23. H1-H2 BY VOWEL FRONTING

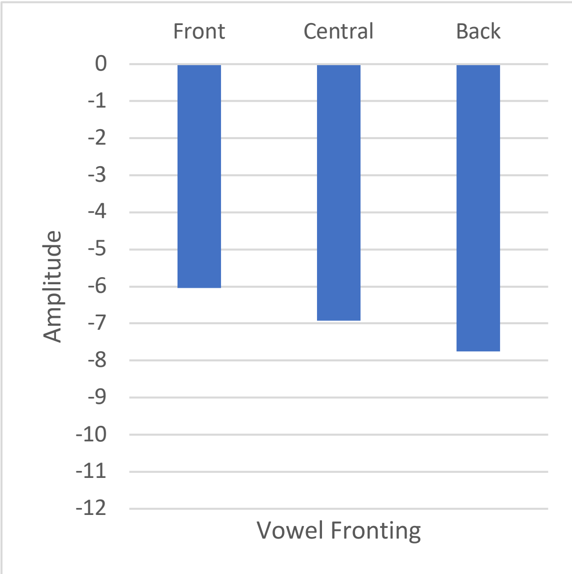


FIGURE 24. F0 BY VOWEL FRONTING

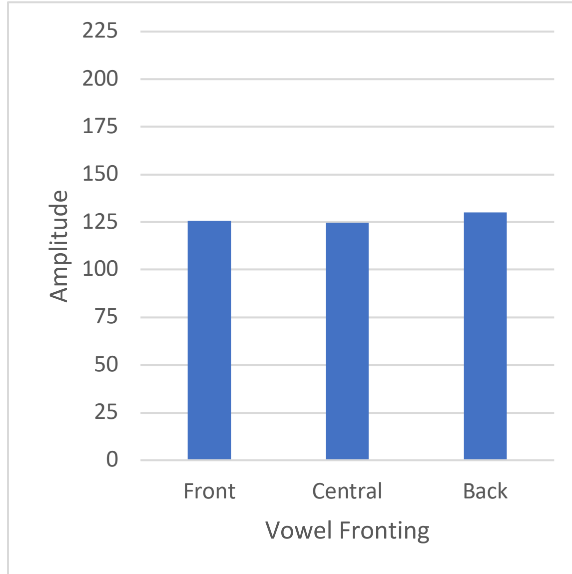


FIGURE 25. H1-H2 BY VOWEL HEIGHT

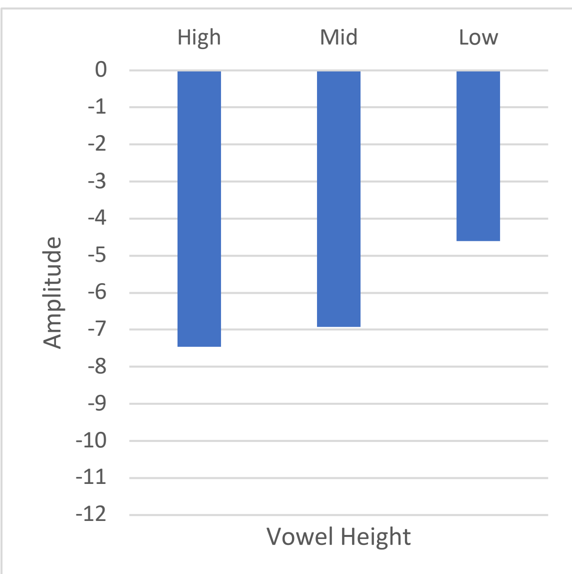


FIGURE 26. F0 BY VOWEL HEIGHT

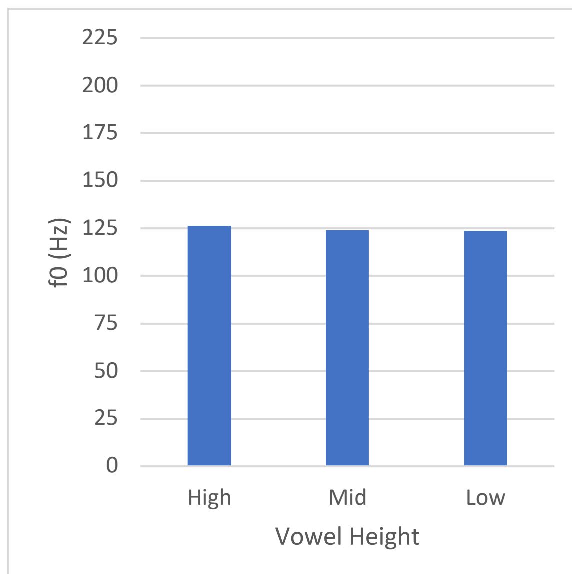


TABLE 16. PAIRED TWO-TAILED T-TEST COMPARING H1-H2 AND F0 BETWEEN VOWEL QUALITIES

	Front vs Central	**Central vs Back**	**Front vs Back**	**High vs Mid**	**Mid vs Low**	**High vs Low**
H1-H2	$p = 0.26$	$p = 0.26$	$p = 0.02$	$p = 0.46$	$p < 0.001$	$p < 0.001$
f0	$p = 0.74$	$p = 0.14$	$p = 0.23$	$p = 0.44$	$p = 0.86$	$p = 0.40$

4. DISCUSSION

I hypothesized that young speakers would use more vocal fry overall, and while they used significantly lower pitch than older speakers, the harmonics measure did not show a difference between the two groups, so I cannot confirm my first hypothesis and say that the pitch change was due to vocal fry. Comparing gender differences in speakers, women used more creak than men, and men had a lower pitch overall. I hypothesized that young people of both genders would use vocal fry more than older people of both genders, and that among young people, women would use more, while among older people, men would use it more. My hypothesis that young women would use the most vocal fry was correct, but that was the only part of my second hypothesis that was correct, as young men actually used the least vocal fry, so age was not a good indicator. There was no significant difference between genders in the older group. Young women used significantly more vocal fry than all older people, and older people used more than young men. While unexpected, this is a very interesting finding, suggesting that gender has recently emerged as a predictor of creak for younger generations. Perhaps young women have increased their level of creak in comparison to the older baseline, and young men have decreased theirs in order to distinguish themselves from their young female peers (or vice versa). Young women used a lower pitch than older women, and young men a lower pitch than older men, which fits the pattern of increased creak in younger groups, while accounting for biological differences.

Overall, possibly due to my small sample size of speakers, many of my findings related to age and gender in an interactional context did not point to the demographic of an interactional partner having an effect on the speaker's use of vocal fry. Most differences due to the demographic of a conversational partner were statistically insignificant, but those that were significant stood out as potentially being part of a pattern. Older women used significantly more creak when speaking to another older woman as compared to with a younger man, and they lowered their pitch when speaking to another older woman as compared to an older man. Young women lowered their pitch with other young women as compared to with older men. Young men lowered their pitch when speaking to other young men as compared to older women or men, and used more creak when speaking to another young man as compared to with a young woman. Older men experienced a lot of interaction between pitch and partner demographic, using the lowest pitch with young women, and the highest pitch with other older men.

I hypothesized that differences in use of vocal fry due to the partner demographic would be mostly consistent regardless of the speaker demographic, with all speakers using more vocal fry with young people, and all older people using it more with men while all young people used it more with women. While my findings were overall not overwhelming and lacking in statistical significance, probably due to a small sample size, the small differences that I did find point to an interesting interaction that is more dependent on the speaker demographic than the partner – opposite to my hypothesis. Most groups used the most creak or lowest pitch with the same demographic, and the least with the opposite demographic. Older women used the most creak and lowest pitch with other older women, and the least creak with young men. Young women used the lowest pitch with other older women, and the least creak and highest pitch with older men. Young men used the most creak and lowest pitch with other young men. Older men did not quite fit this pattern because of their unique pitch interactions, but they did use the least creak with young women. Older men had the opposite effect in pitch, using the highest pitch with other older men and the lowest with young women.

Other interesting interactions occurred in the data aside from those related to speaker and partner differences. I hypothesized that measurements taken at the end of an utterance would have more vocal fry than those at the beginning. Strangely, I found the opposite pattern in H1-H2 differences, showing significantly more creak on early words, yet a much lower pitch on late words. I would expect vocal fry to occur phrase-finally, and for low pitch and greater creak to correlate. I did take H1-A1 and H1-A3 measurements as well (calculating the difference between the first harmonic and first formant or third formant), which should also indicate creak, and while I did not run an analysis on these figures at this time, a brief glance at the results seemed to indicate that they generally followed the same pattern as H1-H2 values. An idea I had was that because I included filler words in my measurements, perhaps I was picking up a disproportionate amount of creak from the initial "um" that most people produced at the beginning of an utterance. However, after removing schwa measurements from the data, my results seemed to follow largely the same pattern, at least in early and late values. I am not sure what else would contribute to this incongruency in pitch and creak, but the numbers are very significant.

I also hypothesized that position in an overall conversation would have an effect on vocal fry as well. I expected vocal fry to increase over the course of a conversation, due to speakers warming up to one another and relaxing their voices. In fact, I again actually found the opposite result, and both creak and low pitch significantly decreased over the course of a conversation. Finally, I measured the difference in vocal fry due to vowel quality, expecting to find little difference between vowels, except higher levels of fry on schwa (due to initial "um"). I found that vocal fry increases as a vowel becomes more high and more back, so /u/, /ʊ/, and /o/ would have the greatest amount of creak. This is surprising to me, as I expected more fry on common words people used while thinking, such as "um," "probably," and "yeah." The examples I can think of would be "so," and "would," occurring in contexts like "so, yeah" and "I think I would…", but I am surprised that those would be the most creaky if that is the case.

While most of my hypotheses were not confirmed, I am intrigued by the patterns I have found here and am interested in doing more research on the topic. Due to the scope of my data, this experiment was limited in many ways. First, eight participants representing four demographics is likely far too small a data pool to get conclusive results, so further research would need to be done with much larger groups of participants. Participants were also not particularly diverse, being almost entirely Caucasian, and while everyone was from different regions of the US, individual dialect influences seemed somewhat neutralized by a Colorado accent in most speakers. Diversity of sexual orientation could also help strengthen the data, as most of my participants were heterosexual, and where they were not, my experiment design did not account for differences in vocal quality due to socialized norms of other sexual orientations or any expressions of gender outside the norm or the binary. Also, being but one person doing annotations and measurements, 78 vowels from each of the 32 recordings was the extent of what time permitted me to sample for this initial study, but considering the recordings are 15-30 minutes long, there is a vast amount of data left that has not been measured. In the future, I would like to have these recordings completely annotated and measured for a deeper foray into this set of data. I would also like to examine "middle" measurements in addition to those at the beginning and end of utterances. In those samples I did take, my vowel selection process could have been more consistent, as I did not always choose exactly the first and last three vowels to measure, but the "clearest" vowels that were closest to those. I often skipped vowels that were particularly reduced or in some way unclear. I also noticed that almost all the words I measured included common filler words such as "um", "probably", "so", and "yeah", so measurements of a greater variety of words with greater semantic content could likely improve this study as well. I was also not confident in all of the vowel markings I made, as some distinctions were hard to make between vowel and consonant when the vowel was followed by approximants with strong formant patterns. In many cases, I relied on marking by ear rather than by the spectrogram, but I would expect that some vowel measurements may have included consonantal content.

There were a few inconsistencies in my process. I tried to only pair people that had not met before, but I had to make two exceptions due to problems with scheduling meetings. Heidi is the mother-in-law of Tania's sister, and while they were friendly and familiar, I felt that they were not so familiar that they would not adhere to normal vocal patterns they used with others. Catherine and Rob, however, are romantic partners, but I could not find a way to avoid this pairing. They noticeably spoke to one another quite differently than they had with everyone else, so I don't think this set of data was particularly strong in the study. There were other minor inconsistencies in the data collection, such as speakers using a different device to log into a session, a camera not working in one session, and my partner supervising a couple meetings while I was unavailable.

I felt that what I saw and heard in the data did not necessarily always match the results. I felt, for instance, that Andrew and Landon had strong final vocal fry, possibly more than the older women group and certainly more than the older men group, but the results show them as the group with the least overall fry. I also felt that the fact that there was more creak measured on early words in an utterance was extremely strange, as I certainly saw and heard very strong creak on phrase-final words in many of the speakers and not on early words, but no speaker showed a greater level of creak at the end according to the data. Tania, for instance, seemed to have extremely strong vocal fry at the end of her utterances, and I heard her initial utterances as very high pitched and modal, but her results pointed to a particularly strong early creak and much lower final creak. This leads me to wonder if varying methods of measuring creak, or measuring for different types of creak as defined by Keating et al., may have been beneficial for these particular speakers.

Vocal fry is a phenomenon that is largely misunderstood and misrepresented. Additional research will help to build understanding and acceptance of it as a legitimate linguistic tool, and I hope that this study has done its part to take one small step in that direction.

REFERENCES

Anderson, Rindy & Casey Klofstad. 2014. Vocal Fry May Undermine the Success of Young Women in the Labor Market. *PLoS ONE* 9.5. doi:10.1371/journal.pone.0097506

Callier, Patrick & Robert Podesva. 2015. Multiple Realizations of Creaky Voice: Evidence for Phonetic and Sociolinguistic Change in Phonation. *New Ways of Analyzing Variation* 44. Online: https://stanford.edu/~podesva/documents/nwav2015-1up.pdf

Chao, Monika, & Julia Bursten. 2021. Girl Talk: Understanding Negative Reactions to Female Vocal Fry. *Hypatia* 36(1). 42-59. doi:10.1017/hyp.2020.55

Esposito, Christina. 2010. Variation in contrastive phonation in Santa Ana Del Valle Zapotec. *Journal of the International Phonetic Association* 40(2). 181-198. doi:10.1017/S0025100310000046

Keating, Patricia, Marc Garellek, & Jody Kreiman. 2015. Acoustic properties of different kinds of creaky voice. *Proceedings of the 18th International Congress of Phonetic Sciences.* Online: http://idiom.ucsd.edu/~mgarellek/files/Keating_etal_2015_ICPhS.pdf

Garfield, Bob. 2013. "Old Fart" Responds to the Great Vocal Fry Outcry of 2013. *Slate*. Online: https://slate.com/human-interest/2013/01/young-women-and-vocal-fry-slate-podcast-wars-continue.html

Gordon, Matthew, & Peter Ladefoged. 2001. Phonation types: a cross-linguistic overview. *Journal of Phonetics* 29. 386-406. Online: http://gordon.faculty.linguistics.ucsb.edu/phonation.pdf

Gross, Terry. (Host). 2015. Fresh Air Weekend: Weighing In On What It Means to 'Sound Gay'. [Radio broadcast episode]. Online: https://www.npr.org/2015/07/11/421470117/fresh-air-weekend-weighing-in-on-what-it-means-to-sound-gay

Higdon, Michael. 2016. Oral Advocacy and Vocal Fry: The Unseemly, Sexist Side of Nonverbal Persuasion. Legal Communications and Rhetoric. *JALWD* 13. 209-220. Online: https://heinonline.org/HOL/P?h=hein.journals/jalwd13&i=212

Irons, Sarah & Jessica Alexander. 2016. Vocal fry in realistic speech: Acoustic characteristics and perceptions of vocal fry in spontaneously produced and read speech. *The Journal of the Acoustical Society of America* 140.3397. doi:https://doi.org/10.1121/1.4970891

Reynolds, Eileen. 2015. What's the Big Deal About Vocal Fry? An NYU Linguist Weighs In. *New York University*. Online: https://www.nyu.edu/about/news-publications/news/2015/september/lisa-davidson-on-vocal-fry.html

Styler, Will, Rebecca Scarborough, Sarah Johnstone, et al. 2014. Automated Nasality Measurement Script Package. *CU Phonetics Lab*.

Wolk, Lesley & Nassima Abdelli-Beruh. 2011. Habitual Use of Vocal Fry in Young Adult Female Speakers. *Journal of Voice* 26.3. E111-E116. doi:https://doi.org/10.1016/j.jvoice.2011.04.007