Creaky, She Spoke: Examining f0, Vocal Creak, and Perceptions of Young Women's Professionalism

Katie A. Conner
The Ohio State University

Creaky, She Spoke: Examining f0, Vocal Creak, and Perceptions of Young Women's Professionalism

Abstract

Recent sociolinguistic work on creak (also known as vocal fry, vocal creak, and creaky voice) has generally focused on its gendered use and perceptions, and more specifically on creak and young women. While Yuasa (2010) found creaky young women to be perceived as "educated, professional, and upwardly mobile" (p.316) sounding, Anderson et al. (2014) and Gallino and Pinto (2021) found the opposite in that creaky young women were perceived to be "less competent, less educated, less trustworthy, less attractive, and less hireable" (p.5). However, all of these works consider creak in isolation, eschewing analyses that examine creak in tandem with other linguistic variables. As voices are perceived as a whole unit, and not singular linguistic features, this leads to an incomplete understanding of the ways in which listeners might perceive creak differently when it is paired with other features such as f0, prosody, and phonetic variation related to dialect, etc. Work examining creak interacting with other linguistic features will provide better understanding for how, when, and why creak is perceived in specific ways by listeners. This study interrogates whether there is an interaction between f0 and creak that might affect professionalism perceptions of speakers and, additionally, how a listeners' speech attitudes about women and creak might mediate f0 and creak's effects on professionalism ratings. 125 participants rated stimuli produced by five young white women from the Midland region sourced from the Nationwide Speech Project corpus (Clopper and Pisoni 2004) on six Likert scales: professional, attractive, friendly, feminine, educated, and authoritative. Additionally, participants completed a set of workplace and women's speech attitudes Likert scale measures (e.g., "to be successful in the workplace, young women should change how they speak to sound more professional"). The presence of creak in stimuli was determined via impressionistic listening and examining spectrograms in Praat. Factor analysis showed that ratings for the six traits patterned into two factors: "Competence" and "Warmth". These two factors were used as dependent measures in linear mixed effects regression models, with average ratings for both factors as the dependent variables. The Competence model included main effects for Workplace Sexism attitudes ($p<.001$) and f0 ($p<.01$), but failed to reproduce findings from the aforementioned previous studies with regards to creak. This study supports previous work by Parker and Borie (2017) that calls for more nuanced and complex analyses of creak moving forward, and the need to move beyond attempting to examine creak as a variable in isolation. It also points to the possibility that the indeterminacy of previous findings on creak and social meaning can be attributed, in part, to differences in experimental design.

Creaky, She Spoke: Examining f0, Vocal Creak, and Perceptions of Young Women's Professionalism

Katie Anne Conner*

1 Introduction

Recent sociolinguistic work has identified creak as a feature that typifies "young women's speech," and has examined creak's effects on listener perceptions of speakers. Much existing work has examined creak in isolation without considering other phonetic aspects such as f0, prosody, and tone. As creak cannot exist in complete isolation from other linguistic features, and especially as creak typically occurs on the lower end of a speaker's f0 range (Johnson 2011, Thomas 2011), studying creak without examining other features that might contribute to its perception gives an incomplete explanation for listener perceptions of creak. This study addresses the effects of possible relationships between f0, creak, and gender ideologies on the perception of young women speaking.

This study failed to reproduce an effect of creak on professionalism ratings, instead finding that only f0 and listener's sexist workplace ideologies had a significant effect on professionalism ratings. These results are evidence that accounting for other linguistic features in experiment and analysis design in relation to creak is vital. Additionally, the indeterminacy of the results in regard to the perceptual effects of creak, when examined in tandem with other previous work (Yuasa 2010, Anderson et al. 2014, Gallena and Pinto 2021, Parker and Borie 2018) raises questions around the ways experimental design (e.g., sample size, stimuli sourcing and creation, etc.) might be contributing to effects found between creak and perceptual ratings. Do studies examining creak need to have a larger sample size to help document a small effect? Does asking participants to creak versus identifying and using "natural" occurrences of creak affect listener perceptions? How does sentence content or speech act of the sample containing creak affect listener perceptions?

2 Motivating Literature

2.1 Creak from a Phonetic and Sociolinguistic Perspective

Vocal creak (i.e., creak, creaky voice, vocal fry, and laryngelization) refers to a mode of phonation in which glottal pulses of the vocal folds are irregular in both frequency and amplitude (Johnson 2011, Thomas 2011). Glottal pulses in creak are irregular in comparison to "modal" phonation, which is generally regarded as a "neutral" phonation (Laver 1980, Gobl and Ni Chasaide 2010). The term "creak" is an umbrella term which refers to several varieties of the phonation (further discussion of these varieties can be found in Keating, Garellek, and Kreiman 2015, and Davidson 2019), which might be perceived as a perceptually coherent group (Davidson 2019). Catford likens the sound of creak to, "a series of rapid taps, like a stick being run along a railing" (1964:32).

Phonetic literature also makes clear that while vocal creak in some individuals might be pathologically originating (due to age or illness generally), that creak is a regular sociolinguistic (Mendoza-Denton 2011, Henton and Bladen 1988, Wolk et al. 2012, Abdelli-Beruh et al. 2016, Szakay 2012), prosodic, segmental, and phonemic (Ladefoged and Maddieson 1996, Gordon and Ladefoged, Riad 2003) cue for many languages. In the context of American English, creak generally occurs utterance-finally and can occur across multiple syllables and words in utterance final positioning.

Sociolinguistically, previous studies suggested links between creak and regional language varieties, including: several British varieties (Catford 1964, Esling 1978, Henton & Bladon 1988, Trudgill 1974), Southern California Chicano English women's speech (Fought 2003), Northwest American English (Ingle 2005), and Māori English in New Zealand (Szakay 2012). Previous work has also shown that creak seems to carry affect meaning for speakers and listeners. These include

anger, sarcasm, and disgust (Murray and Arnott 1993), as well as resignation (Laver 1980), emotional containment or lack of engagement (Zimman 2015, Zimman 2017), and relaxation and intimacy (Ni Chasaide et al. 2004). There is additional evidence that creak might carry some race and/or ethnicity meaning as well (Mendoza-Denton 2011). Finally, some of the most extensive previous linguistic work on creak in English focuses on gender (to be discussed further in Section 2.3) and status/class. In that work, creak, when used by men, has been found to correlate with working class dialects in Norwich (Trudgill 1974), as well as greater social status or class in Edinburgh (Esling 1978).

2.2 F0 from a Phonetic and Sociolinguistic Perspective

Fundamental frequency (f0) is the rate at which the vocal folds vibrate during the production of speech (Johnson 2011, Thomas 2011). This measure is the acoustic correlate of pitch, a perceptual feature. Some previous work has found that lower pitched voices are generally perceived to be more dominant and strong in both men and women (Borokowska and Pawlowski 2011, Jones et al. 2010, Puts et al. 2007). Previous research has also shown that these positive effects can translate to hiring situations and workplaces, with some studies, such as those by Klofstad, Anderson, and Peters (2012) and O'Connor and Barclay (2017), stating that lower f0 voices are rated as more trustworthy and more leadership-worthy in some contexts. Finally, Ohala (1994) linked pitch to gender expression, naming this concept "frequency code," through which he points out the attribution of higher-pitched phonations like falsetto voices to femininity, and lower-pitched phonations, including creak, with masculinity.

2.3 Creak and Gender

Yuasa (2010), Wolk et al (2012), and Podesva (2013) have all found that female speakers are more common adopters of creaky voice than their male counterparts. However, in some of these studies creak was only auditorily coded, with no spectrograms used, meaning there could be a lack of continuity in what counted as creak or possible issues in non-audible creak, while being phonetically and acoustically present (see Davidson (2019) for discussion of this issue). In contrast with those previous studies, Melvin and Clopper (2015) found that females only exhibited 6% more creak than males in their study. Numbers range from study to study as to how much creak was used by men and women, but this variation suggests that the mode of evaluation of creak is an important consideration. Davidson (2019) also found that there seems to be a small measure of bias in the detection of creak between men and women's voices, with creak being more often detected in women's voices than men's in certain production environments.

Some early work identifies creak as a feature indexing masculinity, and/or a feature of "men's speech." In the United Kingdom, creak has been found more often utterance-finally in the speech of men who spoke Received Pronunciation (RP) (Catford 1964) and also those men who were considered upper class for that time and region. Some of this work has pointed to creak as a "hypermasculine" feature, in some cases used by young men or young women to index this hypermasculinity (Stuart-Smith 1999, Mendoza-Denton 2011:266).

More recently, in the United States, creak has been theorized and studied as indexing femininity both from a production/ perception standpoint and an ideological/metalinguistic standpoint. In the early 2000s, creak rose to prominence as the most recent in a line of "concerning features of young women's speech" (e.g., *like*, uptalk, tag questions, "filler words") that fed moral panic surrounding the intelligence and character of young women. This creak-caused-panic lead to many think pieces urging young women to "reclaim your strong female voice" (Wolf 2015), "stop undermining you and your message" (Van Edwards No Year), sound like a boss, or any other number of vague promises about people taking young women seriously if they'd just stop doing *that* (Vuolo 2013, Akst and Pietsch No Year, Gallena and Pinto 2021, Anderson et al. 2014, Fessenden 2011, and many others not cited here). Indeed, there has been work on creak focusing on the very specific pathological framing that ran rampant, discussing the characterization of creak as "catching," "contagious," "sickening," and "an epidemic" (Slobe 2016, Slobe 2019).

Previous work on creak and perceptions of young women's professionalism has yielded seemingly conflicted findings. Yuasa's work focused both on the production and perception of creak,

comparing American English and Japanese speakers, and additionally young American women and young American men. She found that creaky women were perceived to be more educated, more confident, and more "'professional' or 'upwardly mobile' and 'urban'" (Yuasa 2010:329) than the non-creaky counterparts in her study. She additionally found creaky women were perceived to be "slightly 'hesitant' and not so 'confident'" (Yuasa 2010:330). Anderson et al.'s (2014) study focused wholly on the perception of creak, asking participants to rate the speakers (young men and young women). They found that creaky young women were rated to be "less competent, less educated, less trustworthy, less attractive, and less hirable" (Anderson et al. 2014:5). Based on these results, they caution young women against using vocal fry to best "maximize labor market perceptions" (Anderson et al. 2014:5). Finally, Gallena and Pinto (2021) examined speech language pathologists' perceptions of creaky graduate students. The authors reported that the results supported that of Anderson et al. (2014), finding that vocal creak had a "negative impact" on the perception of students' abilities.

The lack of clarity across results in these studies in the social meaning of creak raises the question of whether, and how, the interaction of vocal creak and higher or lower f0 might alter listener perception and evaluation of vocal creak when it is present in a young woman's speech, and additionally, how language and gender ideologies might play into perception of these linguistic variables as well. While vocal creak and f0 have been studied as single variables of interest, and studied in tandem as scholars have sought to identify ways in which creak and f0 might interact and lead to differences in the perception and identification of creak, this study attempts to tease apart how a higher or lower f0 might affect the perception of creak. And based on this, the way f0 and creak's relationship may lead to perceptual differences relating to the professionalism, authority, and femininity of young women.

3 Methods

3.1 Stimuli Sourcing and Creation

Stimuli were taken from the Nationwide Speech Project corpus (Clopper and Pisoni 2006). The NSP corpus features speech from 60 speakers, equally distributed across regional dialects of American English (featuring West, South, Midland, North, New England, and Mid-Atlantic) and an even distribution of 5 men and 5 women from each region. There is little variation across speakers demographically otherwise (speakers were all 18 - 25 years old, white, native English speakers, with parents who were also native English speakers). The relative homogeneity of the samples in this corpus lent itself to this experiment allow for a narrower interpretation of perceptual differences. Reducing demographic noise allows for a closer examination of how creak and f0 affect listener perception and judgment more specifically than might otherwise be possible.

This study used the "high-predictability sentences" from the NSP. These were used due to the low chance that sentence content (such as that from the low predictability sentences) would confuse participants, and because read passages didn't contain enough naturally present creak for the needs of the experiment. Additionally, the read sentences contained minimal prosodic difference, as compared to read passages and interviews, again, for narrower interpretation of perceptual differences. All high-predictability sentences from women for the Midland region were downloaded, then measured and annotated in Praat. All stimuli were measured for their total length (in seconds), average f0 (Hz), f0 range (Hz, minimum to maximum), presence or absence of creak, location of creak in the utterance, total time length of creak, number of syllables with, average creak f0, creak minimum f0, and creak maximum f0. Additionally, any utterances containing creak were annotated to Praat TextGrid for where the creak occurred in the utterance, what words contained creak, and what syllables contained creak. Once measurements were taken, the stimuli were divided into six experimental conditions; two different levels for f0, and three different levels for position of creak. The six conditions in the experiment were as follows: lower f0 & no creak, lower f0 & utterance final creak, lower f0 & utterance final creak as well as creak elsewhere within the utterance, higher f0 & no creak, higher f0 & utterance final creak, and finally higher f0 & utterance final creak as well as creak elsewhere within the utterance. It was from these conditions that stimuli were chosen for the final experiment. Each stimulus condition was comprised of two stimuli that fit that condition's

parameters per speaker (for each of the 5 speakers), for a total of 10 stimuli per condition, and a total of 60 stimuli total in the experiment.

3.2 Experimental Design

The study was constructed and data collected during the spring/summer of 2020. Qualtrics was used for this study, and participants were recruited through the Prolific recruitment platform. The survey consisted of three parts: (1) a rating task, (2) a vocal creak attitudes measure, and (3) a demographic survey. Prior to beginning the survey participants were directed to "imagine that a young woman has just started as an intern at a large business," and were instructed further to consider how that young woman "should" speak in order to be successful both in her current position, as well as future, more senior, positions. They then began the experiment.

3.3 Speaker Rating Task and Attitudes Measures

In the speaker rating task, participants were presented with three stimuli from each of the six experimental stimuli conditions in the experiment, for a total of 18 stimuli per participant. The order of the experimental conditions and stimuli presented were randomized and balanced within Qualtrics. Participants were tasked with listening to and rating each stimulus on a Likert scale of 1 - 6 (1 being "does not sound this way" and 6 being "sounds extremely this way") to indicate how friendly, professional, feminine, educated, attractive, and authoritative the speaker sounded. The six rating traits were selected following from previous work such as Yuasa (2010) and Anderson et al. (2014), and balanced so that three focused around competence traits and three around warmth traits. Participants were unable to return to previous stimuli listened to once the individual stimuli on each page was rated and the rating submitted, such that participants could not alter previous ratings based on comparison between stimuli.

After completing the rating task, participants completed a short set of measures to gather information about their attitudes of women's speech in the workplace, as well as policing women's speech. First, participants were given a definition of vocal creak, and given sounds clips illustrating what vocal creak was, as several measures made explicit reference to vocal creak. Participants read the measure statements, and then gave their answers on a Likert scale of 1-6 (1 being "totally disagree", 6 being "completely agree"). These measures provided for a direct comparison between linguistic and gender ideologies and the perception of creak and f0.

The measures were adapted from several pre-existing scales dealing with various attitudes toward women, including the Attitudes Toward Female Professors Scale (Brant 1979), the Managerial Attitudes toward Women Executives Scales (Dubno 1985), the Sexist Attitudes Toward Women Scale (Benson and Vincent 1980), the Beliefs about Women Scale (Belk and Snell 1986), the Women as Managers Scale (Terborg et al. 1977), and the Women Managers Communication Competencies Scale (Wheeless and Berryman-Fink 1985). While some measures were pulled directly from these scales, others were changed in wording to reflect a more linguistic focus. Within the attitudes measures of this survey, 7 are reversed polarity while 8 were unaltered (i.e., 7 of the measures were stated in the negative and 8 were stated in the positive). This design serves two major purposes. First, as a quality control to make sure participants answered non-randomly, second, as a method of making the participant carefully read and engage with each statement, in order to keep them from possibly glossing over the set of measures without carefully considering their actual opinions and attitudes prior to answering based on what they assume is being asked.

3.4 Demographics

Participants were asked for their age, racial and ethnic identity, their gender identity, what languages were spoken at home, whether they considered American English a native language, what the highest level of education completed was, and what zip code they currently lived in.

Participants were all over 18 and self-selected into the study via Prolific. Participants self-identified as living in the United States, speaking English as once of their primary languages, and additionally as not having any known speech or hearing disorders.

Due to incomplete surveys, only 125 of the total 134 submissions of the survey created were used for the analysis that follows. Participants were mostly white, cis-gendered, and between the ages of 18 and 34 (mean age = 33, sd = 12). Most participants had completed a 4-year college degree by the time they participated in the study.

Age Groups	Race	Gender	Education
18 to 34 - 73	white - 87	cis man - 71	Some High School - 1
35 to 54 - 40	Asian - 11	cis woman - 48	High School - 18
55 to 74 - 12	Black - 10	trans woman - 2	Some College - 18
	Latinx - 7	other - 2	Associate Degree - 15
	Multiracial - 10		Bachelor's - 33
			Some Grad/Prof School - 13
			Grad/Prof Degree - 27

Table 1: Demographic information of the 125 participants used in the study analysis.

3.5 Analysis

As the initial step in the analysis, two factor analyses were conducted on results from the study to determine the best possible analysis path forward given all six traits that participants rated each stimulus on, in addition to the fifteen measures on the participants' speech attitudes on the workplace, women, and vocal creak. A confirmatory factor analysis was completed to determine possible groupings from participants of the 6 traits rated for each speaker per stimuli (professional, authoritative, educated, feminine, friendly, attractive). The *factanal()* function in R with the default *varimax* rotation was used for both factor analyses. In the confirmatory analysis, two factor groups of three traits were identified. These two factors were named Competence (representing the grouping of professional, authoritative, and educated) and Warmth (representing the grouping of feminine, friendly, and attractive) for the purposes of the analysis. Composite scores were calculated as the mean of the three items belonging to each factor (as opposed to utilizing Bartlett scores). This score makes up the Competence mean score and Warmth mean score used in the subsequent statistical hypothesis testing for the study.

An exploratory factor analysis was conducted on the fifteen measures on gender, creak, and workplace speech attitudes that the participants completed toward the end of the study procedure. In the data cleaning phase of the analysis, all measures had scores adjusted such that they all aligned across the Likert scales used, so higher scores indicated more positive/less judgmental responses, and lower scores reflected less positive/more judgmental responses. This factor analysis identified five appropriate factors. These were named and will be further identified in the paper as follows:

Factor 1 - Workplace Sexism
Factor 2 - Policing & Judging (Young) Women's Speech
Factor 3 - Professional Women & Creak in the Workplace
Factor 4 - Comparing Professional Men & Women
Factor 5 - Men are Concerned with Sounding Professional Too

Beyond this initial factor analysis, the hypothesis testing phase for this study centered on the possible interaction between creak and f0, and the ways in which gendered language and workplace attitudes might be mediating factors for the perception of creak in a speaker by listeners. Both the Competence and Warmth factor scores were used as dependent variables in the construction of a mixed effects model. For both the Competence and Warmth hypothesis testing models, fixed effects tested included production variables (the speaker's average f0 across each stimulus, the aggregate duration of creak across each stimulus), as well as variables from the survey/experimental perception data from participants (the participant's answer to whether or not they knew that creak has a negative social image, and the participant's factor scores for each of the five factors identified in the exploratory factor analysis of the 15 gender, workplace, and language attitudes measures). Random intercepts were included in both the Competence and Warmth models for participant and speaker, with a nested stimuli effect as well under the speaker random intercept. Maximal random slopes were utilized in these models, except for where models failed to converge.

4 Results

4.1 Competence Regression Results

A linear mixed effects model was fitted using a step-up process, with the Competence mean score as the dependent variable. Independent variables that were identified for model fitting as fixed effects included main effects for average f0 of the stimuli (a continuous numeric variable) and the aggregate amount of creak within the stimuli (a continuous numeric variable representing timed amount of creak in a given stimulus). Interactions in the hypothesis testing included all possible two-way interactions between creak duration and average f0, between creak duration and the participant's responses as to whether they were aware of creak's negative social connotations when used (a categorical variable coded as "yes", "maybe", and "no", with a Helmert contrast), and between creak and each of the 5 factors identified from the group of 15 measures on gender, creak, and professionalism that participants completed (a continuous numeric variable made up of the bartlett score from the exploratory factor analysis output). All relevant random slopes and intercepts were tested, with some being excluded due to model convergence errors. To avoid possible collinearity effects within the model, the measures of where creak occurred and number of syllables of creak that were collected during the acoustic analysis of stimuli were omitted from the hypothesis testing models.

	Estimate	Std. Error	t-value	p-value
(Intercept)	3.078	0.416	7.396	$p < .001$ ***
Workplace Sexism (Factor 1)	-0.180	0.056	-3.240	$p < .01$ **
Average f0	0.005	0.002	2.585	$p < .05$ *

Table 2: Model fixed effects output with Competence mean score as the DV.

Table 2 provides the model output for the best fit maximal model, which included significant main effects for both "Workplace Sexism" (factor 1) and the average f0 of the stimulus. The model of best fit included random intercepts for each of the data's grouping factors (speaker, participant, and stimuli, with stimuli being nested within speaker), and additionally included random slopes for the "Workplace Sexism" factor, for both speaker and stimuli. In the case of the Workplace Sexism factor, there is a negative relationship, such that participants who were "less judgmental" for the workplace sexism factor (which translates to having less sexist views) gave overall lower mean scores for Competence to the stimuli heard. In the case of the main effect for average f0 of a given stimulus, there is a positive relationship, such that an increase in the average f0 of a stimulus correlated with higher mean Competence scores awarded by the participants.

4.2 Warmth Regression Results

As with the Competence factor, a mixed effects linear regression model was fit to the Warmth factor (comprised of the feminine, friendly, and attractive averaged scores). For this set of hypothesis testing models, main effects tested included production variables (the speaker's average f0 across each stimulus, the aggregate duration of creak across each stimulus), as well as variables from the survey/experimental perception data from participants (the participant's answer to whether or not they knew that creak has a negative social image, and the participant's factor scores for each of the five factors identified in the exploratory factor analysis of the 15 gender, workplace, and language attitudes measures). All relevant random slopes and intercepts were tested, with some being excluded due to model convergence errors. In order to avoid possible collinearity effects within the model, the measures of where creak occurred and number of syllables of creak that were collected during the acoustic analysis of stimuli were not included within the hypothesis testing models.

	Estimate	Std. Error	t-value	p-value
(Intercept)	2.924	0.396	7.377	$p < .001$ ***
Workplace Sexism	-0.200	0.053	-3.780	$p < .001$ ***

(Factor 1)				
Prof. Women & Creak (Factor 3)	0.189	0.048	3.920	*p* < .001 ***
Average f0	0.006	0.002	3.749	*p* < .001 ***
Knows Creak is Negative: "Yes"	0.334	0.138	2.423	0.017 *
Knows Creak is Negative: "Maybe"	0.295	0.155	1.900	0.060
Men Concerned w/ Prof. Too	0.119	0.050	2.396	0.018*
Workplace Sexism (Factor 1) * Judging Women's Speech (Factor 2)	0.089	0.044	2.010	0.047 *
Prof. Women & Creak (Factor 3)* Comparing Prof. Men & Women (Factor 4)	0.097	0.041	2.368	0.020 *

Table 3: Model fixed effects output with Warmth factor score as the DV.

Table 3 provides the model output for the best fit maximal model, which included significant main effects for the average f0 of the stimulus, "Workplace Sexism" (factor 1), "Professional Woman and creak" (factor 3), "Men are worried about sounding professional too" (factor 5), and whether the participant knew creak had negative connotations. Significant interactions included "Workplace sexism" (factor 1) and "policing women's speech" (factor 2), and "Professional women and creak" (factor 3) and "Comparing professional men and women" (factor 4). The model of best fit included random intercepts for each of the data's grouping factors (speaker, participant, and stimuli, with stimuli being nested within speaker), and additionally included random slopes for the "Comparing professional men and women" factor, for both speaker and stimuli. The relationship between average f0 and the Warmth mean score is positive, such that higher average f0 within a given stimulus correlates with a higher Warmth mean score awarded by the participant for that stimulus. The relationship between "professional women and creak" (factor 3) and the Warmth factor score is positive, such that participants who expressed less judgmental, more "positive" ideologies surrounding professional women and vocal creak awarded higher mean scores for the Warmth factor to stimuli. And finally, the relationship between "men being concerned with sounding professional too" (factor 5) and the Warmth factor score was also positive, such that less judgmental views held by participants correlated with higher mean ratings being given for the Warmth factor.

5 Discussion

Focusing on the study's larger question of professionalism, per the results presented in Section 4.1, average f0 and workplace sexism attitudes had an effect on the Competence factor scores for speakers, but vocal creak did not, either as a main effect or in an interaction. Given that previous work (Yuasa 2010, Anderson et al. 2014, Gallena and Pinto 2021) *has* found support for creak having perceptual effects on professionalism, and given that there is a strong folk linguistic narrative of creak being something people are conscious of and negatively perceive, it was unexpected that this study did not find support for previous findings. Some possible answers to this can be found in the methodological differences between this study and previous work. However, it is also possible that the failure to reproduce this effect here is due to this study's accounting for f0 in models where Yuasa (2010), Anderson et al. (2014), and Gallena and Pinto (2021) did not. While these previous studies discuss creak in isolation, sans formal empirical experimentation and discussion of other variables within the stimuli in their papers, other linguistic work, in addition to the work presented

here, might help to untangle some of the tension and flux across the results of previous work. Questioning not only creak, but also other features in tandem, to see what effects or interactions can help to explain perceptual effects, will be paramount.

5.1 Experimental Power

One potential analysis of these results points to the possibility that there is an effect within the data, but that it is a particularly weak effect, and thus has not emerged with the limited participant pool of 125 used for the analysis within this current work. Yuasa's (2010) work and analysis utilized 175 listeners, Anderson et al.'s (2014) study utilized 800 listeners, and Gallena and Pinto's (2021) study had 150 listeners. The widest gap clearly lies between the Anderson et al. (2014) work and this study, but a replication using the same methods while collecting a much larger sample size may be in order to further probe whether increased power might reveal a weak effect hidden by an underpowered study.

5.2 Priming and Stimuli Sourcing/Creation

Another possible explanation for the indeterminacy of the results concerning creak across studies is that differences between the specific experimental protocols used here versus in previous experiments contributed to the lack of support for creak's effect on professionalism ratings. Anderson et al.'s (2014) work very specifically focused its participants on a workplace hiring situation, which included having speakers intentionally affect creak and non-creak stimulus readings of "Thank you for this opportunity." There are several differences here that might point toward contrasting results. First, the study carried out here utilized pre-existing elicited stimuli. Stimuli were chosen due to the naturalistic occurrence or non-occurrence of creak, in contrast with Anderson et al., (2014) where speakers were instructed to creak as they spoke. These stimuli were sensical, but unrelated to the workplace (e.g., "a round hole won't take a square peg", "put some butter on your bread"). It is possible that the Anderson et al. stimuli and instructions primed participants to evaluate the speakers differently than my instructions to "think about what it means to sound professional" alone.

Another possibility lies in the possible differences between "natural" and "affected" creak. As previously stated in this paper, creak does not exist alone in the speech signal, but alongside multiple other language features. It is possible that speakers are doing something intonationally, with pitch, vowel length, or any other number of features when asked to creak "on purpose" that does not occur in the same way when creak occurs "naturally". This could mean more attention to paid to creak by listeners, or could possibly mean listeners are picking up on creak as well as other features, possibly leading to some of the effects previously attested to in studies.

6 Conclusion

The current study does not lend additional support to previous findings that the presence of vocal creak in young women's speech correlates with lower professionalism ratings from listeners. While is it unclear if this indeterminacy is due to the experimental design, due to this study's accounting for the effect of f0 on perception in its inferential modelling, or possibly some other yet undiscovered explanation, it does raise further questions surrounding how language scientists pursue these lines of inquiry, and how the questions asked and variables chosen for study directly shape the available set of null and alternative hypotheses.

Methodological questions aside, the results of this study *do* provide support for pushing back against the popular narrative that vocal creak is point-blank a "problem" in the speech of young women. The results demonstrate that attempting to account for negative perceptions through creak alone is an oversimplification of the complexity of how speech patterns are perceived. To assign all weight of professionalism (or educatedness, authority, etc.) perceptions to a single feature is to ignore the contributions of many other speech features, language attitudes, gender ideologies, and beyond that might be playing a role in the way a listener makes social sense and meaning of a speaker's utterances. And by that token, it is an overreach to claim that young women can avoid sexist linguistic profiling, gender discrimination or sexism in hiring practices, simply by "being cognizant" or wary of using creak, as Anderson et al. (2014) and Gallena and Pinto (2021) suggest.

References

Abdelli-Beruh, Nassima B., Thomas Drugman, and RH Red Owl. 2016. Occurrence frequencies of acoustic patterns of vocal fry in American English speakers. *Journal of Voice,* 30(6).

Akst, Lee and Christine M. Pietsch. No Year. Is Vocal Fry Ruining My Voice? *Johns Hopkins Medicine.* URL https://www.hopkinsmedicine.org/health/conditions-and-diseases/is-vocal-fry-ruining-my-voice

Anderson, Rindy C., Casey A. Klofstad, William J. Mayew, and Mohan Venkatachalam. 2014. Vocal Fry May Undermine the Success of Young Women in the Labor Market. *PLoS ONE*, 9(5).

Belk, Sharyn. S., and William E. Snell 1986. Beliefs about Women: Components and Correlates. *Personality and Social Psychology Bulletin*, 12(4):403–413.

Benson, Peter L., and Steven Vincent. 1980. Development and Validation of the Sexist Attitudes Toward Women Scale (SATWS). *Psychology of Women Quarterly*, 5(2):276–291.

Borokowska, Barbara, and Boguslaw Pawlowski. 2011. Female voice frequency in the context of dominance and attractiveness perception. *Animal Behavior* 82:55–59.

Brant, William D. 1979. Attitudes Toward Female Professors Scale. *Psychological Reports,* 44(3, Pt 2):1310-1310.

Catford, John C. 1964. *Phonation Types: The Classification of Some Laryngeal Components of Speech Production: In Honour of Daniel Jones.* London: Longmans.

Clopper, Cynthia G., and David B. Pisoni. 2006. The Nationwide Speech Project: A new corpus of American English dialects. *Speech Communication,* 48:633-644.

Davidson, Lisa. 2019. Perceptual coherence of creaky voice qualities. Conference proceedings presented at the International Congress of Phonetic Sciences ICPhS 2019.

Davidson, Lisa. 2019. The effects of pitch, gender, and prosodic context on the identification of creaky voice. *Phonetica.*

Dubno, Peter 1985. Attitudes toward women executives: A longitudinal approach. *Academy of Management Journal*, 28(1):235-239.

Esling, J. 1978. The identification of features of voice quality in social groups. *Journal of the International Phonetic Association*, 8(1-2):18-23.

Fessenden, Marissa. 2011. 'Vocal Fry' Creeping Into U.S. Speech: Young women end sentence with a gravelly buzz in what could be a catching trend. *Science*. URL https://www.science.org/content/article/vocal-fry-creeping-us-speech

Fought, Carmen. 2003. *Chicano English in Context*. New York: Palgrave Macmillan.

Gallena, Sally K. and James A. Pinto. 2021. How Graduate Students With Vocal Fry Are Perceived by Speech-Language Pathologists. *SIG 10 Issues in Higher Ed* 6:1554-1565.

Gobl, Christer, and Ailbhe Ní Chasaide. 2000. Testing Affective Correlates of Voice Quality through Analysis and resynthesis. in *Proceedings of the ISCA Workshop on Speech and Emotion: A Conceptual Framework for Research*, ed. Ellen Douglas-Cowie, Roddy Cowie, and Marc Schröder, 178–83.

Gordon, Matthew, and Peter Ladefoged. 2001. Phonation types: A cross-linguistic overview. *Journal of Phonetics*, 29(4):383–406.

Henton, Caroline, and Anthony Bladon. 1988. Creak as a Sociophonetic Marker. in *Language, Speech and Mind: Studies in Honour of Victoria A. Fromkin*, ed. Larry M. Hyman and Charles N. Li, 3–29.

Ingle, Jennifer. 2005. Pacific Northwest Vowels: A Seattle Neighborhood Dialect Study. Paper presented at the 2005 Annual Meeting of the Acoustical Society of America, Vancouver, B.C..

Johnson, Keith. 2011. *Acoustic and Auditory Phonetics*. Hoboken, NJ: Wiley Blackwell.

Jones, Benedict C., David R. Feinberg, Lisa M. DeBruine, Anthony C. Little, and Jovana Vukovic. 2010. A domain specific opposite-sex bias in human preferences for manipulated voice pitch. *Animal Behavior* 79:57–62.

Keating, Patricia A., Marc Garellek, and Jody Kreiman. 2015. Acoustic properties of different kinds of creaky voice. In *ICPhS*, 2015(1):2-7.

Klofstad, Casey A., Rindy C. Anderson, and Susan Peters. 2012. Sounds like a winner: voice pitch influences perception of leadership capacity in both men and women. *Proc. R. Soc. B,* 297:2698–2704.

Laver, John. 1980. *The Phonetic Descriptions of Voice Quality.* Cambridge: Cambridge univ. press

Ladefoged, Peter, and Ian Maddieson. 1996. *The Sounds of the World's Languages*. Oxford: Blackwell.

Melvin, Shannon, and Cynthia G. Clopper. 2015. Gender variation in creaky voice and fundamental frequency; ISCfI,(ed.) 18th International Congress of Phonetic Sciences. Edinburgh: Scotland.

Mendoza-Denton, Norma. 2011. The Semiotic Hitchhiker's Guide to Creaky Voice: Circulation and Gendered Hardcore in a Chicano/a Persona. *Journal of Linguistic Anthropology* 21:261–280.

Murray, Iain R., and John L. Arnott. 1993. Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *The Journal of the Acoustical Society of America*, 93(2):1097-1108.

Ni Chasaide, Ailbhe, and Christer Gobl. 2004. Phonatory quality and affect. In W.J. Hardcastle and J.M. Beck. (eds.), *A figure of speech: A Festschrift for John Laver*, 323–346.

O'Connor, Jillian J. M., and Pat Barclay. 2017. The influence of voice pitch on perceptions of trustworthiness across social contexts. *Evolution and Human Behavior*, 38:506–512.

Ohala, John J. 1994. The frequency code underlies the sound-symbolic use of voice pitch. *Sound symbolism*, 2:325-347.

Parker, Michelle A., and Stephanie A. Borrie. 2018. Judgments of intelligence and likability of young adult female speakers of American English: The influence of vocal fry and the surrounding acoustic-prosodic context. *Journal of Voice*, 32(5):538-545.

Podesva, Robert. 2013. Gender and the Social Meaning of non-modal Phonation Types. *Proceedings of the 37th Annual Meeting of the Berkeley Linguistics Society*, 427-448.

Puts, David Andrew, Carolyn R. Hodges, Rodrigo A. Cardenas, and Steven J.C. Gaulin. 2007. Men's voices as dominance signals: vocal fundamental and formant frequencies influence dominance attributions among men. *Evolution and Human Behavior* 28:340–344.

Riad, Tomas. 2003. The origin of Danish stød. In Aditi Lahiri (ed.). *Analogy, Levelling, Markedness: Principles of Change in Phonology and Morphology*. Walter de Gruyter. 261–300.

Stuart-Smith, Jane. 1999. Glasgow: Accent and voice quality. *Urban voices: Accent studies in the British Isles*, 203-222.

Slobe, Tyanna. 2019. Creaky Voice as Embodied Contagion in U.S. Media. Paper presented at Language in the Media Conference at Universidade Federal do Rio de Janeiro, Rio de Janeiro.

Slobe, Tyanna. 2016. This American Creak: Metaphors of Virus, Infection, and Contagion in Girls' Social Networks. Paper presented at the 2016 American Anthropological Association Annual Meeting, Minneapolis.

Szakay, Anita. 2012. Voice quality as a marker of ethnicity in New Zealand: From acoustics to perception. *Journal of Sociolinguistics*, 16(3):382-397.

Terborg, James R., Lawrence H. Peters, Daniel R. Ilgen, and Frank Smith. 1977. Organizational and personal correlates of attitudes toward women as managers. *Academy of Management Journal*, *20*(1):89-100.

Thomas, E. R. 2011. *Sociophonetics: An Introduction*. New York, NY: Palgrave Macmillan.

Trudgill, Peter. 1974. *The Social Differentiation of English in Norwich*. Cambridge: Cambridge univ. press

Van Edwards, Vanessa. No Year. Vocal Fry: What it is and how to get rid of it. *Science of People*. URL https://www.scienceofpeople.com/vocal-fry/

Vuolo, Mike. 2013. Lexicon Valley Episode No. 24: Get Your Creak On. *Lexicon Valley Podcast*. URL http://www.slate.com/articles/podcasts/lexicon_valley/2013/01/lexicon_valley_on_creaky_voice_or_vocal_fry_in_young_american_women.html

Wheeless, Virginia E., and Cynthia Berryman-Fink. 1985. Perceptions of women managers and their communicator competencies. *Communication Quarterly*, *33*(2):137-148.

Wolf, Naomi. 2015. Young women, give up the vocal fry and reclaim your strong female voice. *The Guardian*. URL https://www.theguardian.com/commentisfree/2015/jul/24/vocal-fry-strong-female-voice

Wolk, Lesley, Nassima Abdelli-Beruh, and Dianne Slavin. 2012. Habitual use of vocal fry in young adult female speakers. *Journal of Voice*, 26(3):111-116.

Yuasa, Ikuko P. 2010. Creaky voice: a new feminine voice quality for young urban oriented, upwardly-mobile American women?. *American Speech*, 85(3):315- 337

Zimman, Lal. 2017. Operationalizing stance: Affective stance in a pervasively creaky transgender speaker. Paper presented at the 2017 Annual Meeting of the LSA, Austin.

Zimman, Lal. 2015. Creak as disengagement: Gender, affect, and the iconization of voice quality. Paper presented at NWAV 44, University of Toronto.

Department of Linguistics
The Ohio State University
Columbus, OH 43210
conner.280@osu.edu