



8-1-2021

## Just a little respect: Authority and competency in women's speech

Bridget Anne Carroll

[How does access to this work benefit you? Let us know!](#)

Follow this and additional works at: <https://commons.und.edu/theses>



Part of the [Linguistics Commons](#)

---

### Recommended Citation

Carroll, Bridget Anne, "Just a little respect: Authority and competency in women's speech" (2021). *Theses and Dissertations*. 4066.

<https://commons.und.edu/theses/4066>

This Thesis is brought to you for free and open access by the Theses, Dissertations, and Senior Projects at UND Scholarly Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of UND Scholarly Commons. For more information, please contact [und.commons@library.und.edu](mailto:und.commons@library.und.edu).

JUST A LITTLE RESPECT:  
AUTHORITY AND COMPETENCY IN WOMEN'S SPEECH

by

Bridget Anne Carroll  
Bachelor of Arts, Wichita State University, 2017

A Thesis  
Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Master of Arts

Grand Forks, North Dakota  
August  
2021

This thesis, submitted by Bridget Anne Carroll in partial fulfillment of the requirements for the Degree of Master of Arts from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

---

Dr. Mark E. Karan, Chair

---

Dr. Katrina E. Tang

---

Dr. Kristine M. Trammell

This thesis is being submitted by the appointed advisory committee as having met all of the requirements of the School of Graduate Studies at the University of North Dakota and is hereby approved.

---

Chris Nelson, Associate Dean  
School of Graduate Studies

---

June 3, 2021

## PERMISSION

Title            Just a Little Respect: Authority and Competency in Women's Speech  
Department    Linguistics  
Degree         Master of Arts

In presenting this thesis in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my thesis work or, in his absence, by the Chairperson of the department or the Dean of the School of Graduate Studies. It is understood that any copying or publication or other use of this thesis or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my thesis.

Bridget Anne Carroll  
June 3, 2021

## TABLE OF CONTENTS

LIST OF FIGURES	vi
ACKNOWLEDGEMENTS	ix
ABSTRACT	x
CHAPTER 1 INTRODUCTION	1
1.1 Women’s Speech	3
1.1.1 Two dialects	3
1.1.2 Lexical distinctions	3
1.1.3 Tag-questions	4
1.1.4 Politeness in performing gender	5
1.1.5 Distinctions in conversation	6
1.1.6 Gender inequality	7
1.2 Pitch	8
1.2.1 Vocal tract and average pitch	8
1.2.2 Lower voices signal dominance and leadership	11
1.2.3 Attractiveness	12
1.2.4 Context-based accommodations	13
1.3 Uptalk	15
1.3.1 Definition of uptalk	16
1.3.2 Origins of uptalk	17
1.3.3 Societal views on uptalk	18
1.3.4 Questioning intonation	19
1.3.5 Contrast with continuation rise	20
1.3.6 Uptalk users	20
1.3.7 Meaning and motivation	21
1.4 Fry	23
1.4.1 Definition of fry	23
1.4.2 Frog voice	25
1.4.3 Perception	26
1.5 Conflicting motivations in young women’s speech	26
1.6 Videoconferencing	27
CHAPTER 2 METHODOLOGY	31
2.1 Stage one: Production	31
2.1.1 Anomalies	33
2.1.2 Recording preparation	33
2.1.3 Analysis of recordings	35
2.1.4 Pitch measurements	35
2.1.5 Fry measurements	37
2.1.6 Uptalk	37
2.2 Stage two: Perception	38
2.2.1 Selection of recordings	38
2.2.2 Surveys	41
CHAPTER 3 RESULTS	43
3.1 Overview of analysis	43
3.1.1 Number of phrases	43
3.1.2 Statistical methods	45
3.2 Pitch results	46
3.2.1 Statistical tests for pitch	47
3.2.2 Speaker variations	49
3.3 Uptalk results	53
3.3.1 Uptalk description	54
3.3.2 Statistical results for uptalk	58
3.3.3 Interview	59

3.3.4	Casual	61
3.3.5	Presentation	62
3.4	Fry results	64
3.4.1	Statistical tests for fry	66
3.5	Perception test results	70
3.5.1	General trends	73
3.5.2	Perceptions of pitch: Speaker 1	73
3.5.3	Perceptions of uptalk: Speaker 13	76
3.5.4	Perceptions of fry: Speaker 7	78
3.5.5	Gender	80
3.5.6	Peer group perception	82
CHAPTER 4	DISCUSSION	86
4.1	Summary of findings	86
4.1.1	Pitch	86
4.1.2	Uptalk	87
4.1.3	Fry	88
4.2	Limitations of this research	89
4.2.1	Number of speakers and participants	89
4.2.2	Race	89
4.2.3	Videoconferencing	89
4.2.4	Naturalness and uniformity as competing values	90
4.3	Range of pitch	91
4.4	Attitudes about uptalk	92
4.5	Further research	93
4.6	Conclusion	94
APPENDICES		96
APPENDIX A	.....	97
APPENDIX B	.....	99
APPENDIX C	.....	100
APPENDIX D	.....	102
REFERENCES		108

## LIST OF FIGURES

Figure 1. Average pitch of each speaker with range from Klofstad .....	9
Figure 2. Average pitch of each speaker with range from Shevchenko .....	11
Figure 3. Spectrogram of uptalk .....	17
Figure 4. Fry with low perceived pitch.....	24
Figure 5. Fry with high perceived pitch.....	25
Figure 6. Example from perception survey .....	41
Figure 7. Final questions on perceptions survey .....	42
Figure 8. Average pitch by style .....	47
Figure 9. Average pitch (in Hz) for each speaker by style.....	50
Figure 10. Average pitch (in Hz) for each speaker in casual speech .....	52
Figure 11. Change in pitch during uptalk rise vs duration of rise .....	55
Figure 12. Length of phrases across speech styles.....	57
Figure 13. Percentages of phrases with uptalk.....	60
Figure 14. Text for presentation speech.....	63
Figure 15. Age of participants grouped by gender .....	72
Figure 16. Survey results for Speaker 1 .....	74
Figure 17. Survey results for Speaker 13 .....	77
Figure 18. Survey results for Speaker 7 .....	79
Figure 19. Responses for Speaker 13 by gender.....	81
Figure 20. Results for Speaker 13 by leadership preference.....	82
Figure 21. Survey responses for Speaker 13: Interview by peer group.....	83
Figure 22. Responses for Speaker 1: Interview by peer group.....	84
Figure 23. Responses for Speaker 1: Presentation by peer group.....	85

## LIST OF TABLES

Table 1. Average speaking pitch by gender (Klofstad et al. 2016) .....	9
Table 2. Average speaking pitch by gender and class (Shevchenko 1999).....	10
Table 3. Audacity settings for export.....	34
Table 4. Text grid tiers — labels and descriptions .....	35
Table 5. Amount of fry in each phrase.....	37
Table 6. Measurements for Speaker 1 .....	39
Table 7. Measurements for Speaker 7 .....	39
Table 8. Measurements for Speaker 13 .....	40
Table 9. Number of phrases in each speech style by speaker .....	44
Table 10. Average pitch for each style.....	46
Table 11. Statistical results for pitch and style .....	48
Table 12. Average pitch (in Hz) for each speaker by style.....	49
Table 13. Amount of fry for Speaker 3 in interview and casual speech .....	50
Table 14. Types of fry for Speaker 3 in interview and casual speech.....	51
Table 15. Percentage of fry for Speakers 9, 15, 16, and 17.....	53
Table 16. Percentages of phrases with uptalk by speech style .....	53
Table 17. Increase in F0 during uptalk rises .....	56
Table 18. Uptalk rises that increased in F0 by at least 100%.....	56
Table 19. Average duration of uptalk rise.....	57
Table 20. Statistical results for uptalk .....	58
Table 21. Use of uptalk by Speaker 10 .....	61
Table 22. Uptalk use for Speaker 6 compared with other speakers.....	64
Table 23. Rankings of percentages of fry by speech style .....	65
Table 24. Statistical results for fry .....	66
Table 25. Location of fry in the phrase.....	67



Table 26. Statistical results for phrase-initial fry and speech style.....	68
Table 27. Statistical results for phrase-initial fry and education .....	68
Table 28. Statistical results of phrase-medial fry and style .....	69
Table 29. Statistical results of phrase-final fry and style.....	69
Table 30. Gender of survey participants .....	70
Table 31. Age of survey participants .....	71
Table 32. Gender preferences in leadership by participant gender .....	73

## ACKNOWLEDGEMENTS

In group projects, one person does all the work and everyone else gets the credit for it. In this project, the insight, expertise, and encouragement of others has been indispensable, but, unfortunately, it is only my name on the paper. I am grateful to my committee chair, Dr. Mark Karan, for his patience and mentorship and to Dr. Katie Tang and Dr. Kristine Trammell for strengthening my research and writing.

This thesis could not have happened without the efforts of David Padgett, the 18 women who added yet another Zoom call to their busy schedules, and Mark and Dr. Bethany Shockley who taught me everything I now know about statistics (all errors are mine). I am grateful for everyone who took time to respond to the survey and recruit other people to participate.

My thanks to the SIL-UND community for their commitment to excellence and collaboration. Erin, Kate, Becky, Cory,  $\sigma$ , and the thesisers all contributed to this research through advice and encouragement.

Thank you to my family — Dave, Margaret, William, Nathan, Emily and especially Isabelle, who not only proof-read the first draft but also inspired the whole idea. Thank you to the random lady who made fun of her voice in a Zoom call. I owe you one.

## ABSTRACT

Young women have conflicting motivations directing how they use pitch, vocal fry, and uptalk intonation. High pitch and uptalk may emphasize their femininity, but low pitch and vocal fry are associated with better leadership. Thus, it is difficult to predict how young women will speak in a particular situation. This thesis measures how 16 young women used pitch, vocal fry, and uptalk in three different speech styles collected through videoconferencing calls. Surveys determined how the changes in speech affected the listener's judgments of the speaker. The lowest average pitch was in interview style speech and the largest range of pitch in casual style speech. The young women used more uptalk in interview style speech than in presentation or casual speech. The highest amount of fry was in presentation style speech. Male participants were more likely than female participants to judge a speaker using uptalk as less competent.

## CHAPTER 1 INTRODUCTION

The speech of women, especially young women, has been under particular scrutiny since 1973 when Lakoff's seminal work "Language and woman's place" was published. Women's speech patterns are often given as a reason why women are perceived as less assertive, forthright, and competent, particularly in a professional environment (Klofstad et al. 2012). Dissimilarities between men and women's speech continue to exist. Differences themselves are not an issue, but discrimination based on these differences is. In response, many women in the professional environment effectively code-switch, leaving behind "lady-like language" in favor of neutral or even masculine qualities (Lakoff 1973). However, this choice creates the potential for derision for trying to talk more like their male counterparts in arenas where authority and competence are essential qualities (Lakoff 1973).

Lakoff's work was published nearly 50 years ago. Since then, more research has been done on differences in vocabulary and syntax between men and women and possible changes that take place in female voices in different environments and with diverse audiences. As Holmes (1990) has pointed out, the context and use of these differences, not merely their presence, needs to be understood. Thus, there are still questions to be asked: what changes do young women make in their speech when they want to appear more competent and professional? Do these speech adjustments affect how the speakers are perceived?

This paper examines what female speakers actually do in three different contexts during a videoconferencing call and how their speech changes are perceived by their

audience. I will look at three speech characteristics, pitch, uptalk, and vocal fry, occurring in three speech styles, casual narrative, presentation, and interview. I argue that the context, presence of an interlocutor, and formality level of the speech setting will impact a young women's pitch and use of uptalk and fry. My original hypothesis was that in talking with a male interviewer in a videoconferencing call, women would use lower speaking pitch than in a casual conversation with a peer. However, there are conflicting motivations for raising or lowering pitch. Listeners may judge a speaker with lower pitch to be more authoritative (Klofstad et al. 2012). Still, lower pitch may negatively affect perceptions of attractiveness (Collins & Missing 2003) which presents a quandary.

Because vocal fry usually occurs with low fundamental frequency (F0), I hypothesized that speakers would use more vocal fry in the speech style with the lowest pitch. I further suspected that listeners judge female speakers to be *more* competent when they have lower pitch but *less* competent when they have higher use of vocal fry which creates a second quandary.

The motivations to use or avoid uptalk are also varied. On the one hand, uptalk can be used to "hold the floor" and avoid interruptions, so a young woman might use it when she is concerned that she will be interrupted by a superior or to mitigate her remarks if she thinks her superior will be displeased with what she is saying. On the other hand, the use of uptalk is often seen as uncertainty or overly deferential and may be avoided by a speaker wanting to appear confident. These conflicting interpretations of uptalk present a third quandary for young women. Without uptalk, they may appear assertive but risk sounding rude. If they use uptalk, they may sound more polite but seem unsure of themselves.

So what do women actually do? Do they increase their use of uptalk in formal situations or decrease it? Specifically, in the context of videoconferencing, what changes

do women speaking American English make in average vocal pitch and instances of uptalk and vocal fry in different formality contexts? How are these speech changes perceived? This study offers answers to these questions.

## 1.1 Women's Speech

### 1.1.1 *Two dialects*

Lakoff puts forward the idea that women, especially young educated women, have to be fluent in two different dialects of English: lady-like speech and neutral speech (1973: 48)<sup>1</sup>. Yet having the use of two dialects means she must choose when to use which one.

If she refuses to talk like a lady, she is ridiculed and subjected to criticism as unfeminine; if she does learn, she is ridiculed as unable to think clearly, unable to take part in a serious discussion: in some sense, as less than fully human.

These two choices which a woman has — to be less than a woman or less than a person — are highly painful. (Lakoff 1973: 48)

Lakoff suggests this uncertainty in correct language choice may be why women participate less than men in college class discussions (1973: 48). But what is lady-like language?

### 1.1.2 *Lexical distinctions*

Female English speakers have a different acceptable vocabulary than male English speakers. (Lakoff 1973: 49). Lakoff gives the example of color terms. Women are allowed to use words like *mauve* and *beige*, but men generally are not (Lakoff 1973: 49). Though Lakoff did not have quantitative evidence for her claim, it seems to have been

---

<sup>1</sup> Lakoff's analysis has been criticized for its contrived data and reinforcement of stereotypes (Holmes 1990). However, her work led the way for discussions about the linguistic realization of differences in power and role in society. The stereotypes themselves are useful for my research (see Section 1.1.4) as I attempt to uncover, through data of natural speech, how women actually talk in the 21st century.

corroborated by further research. In one study, men were more likely to use basic color terms than women even when provided with a list of more exact color terms to choose from (Nowaczyk 1982).

Conversely, men typically are allowed to use stronger expletives than women (Lakoff 1973: 50). In de Klerk's study of cursing in adolescents, she found that while both boys and girls used stronger swear words with same-sex peers, boys were more willing to use forceful expletives in all social situations (de Klerk 1991). There were strong inhibitions against using swear words with any people who had a higher social status or more power (de Klerk 1991). Notably, girls who attended private school were more likely to use forceful expletives than boys who attend government schools in some contexts (de Klerk 1991). These results seem to indicate the influence of self-perceived status in the freedom to use swear words.

Lakoff also notes that some women are able to adopt men's language, for example the growing use of profanity by women. Whereas, men rarely use vocabulary from the female lexicon unless they are intentionally subverting expectations of masculinity (Lakoff 1973: 50).

### *1.1.3 Tag-questions*

The differences between male and female speech extend beyond vocabulary. Tag-questions such as "isn't it?" or "right?" are appended to statements to mitigate their forcefulness or allow the listener to disagree (1973: 54)). While both men and women use tag-questions, they seem to use them for different reasons (Holmes 1990). Tag-questions can be used both when the speaker is sure of the veracity of their statement or to express uncertainty, but are generally not used when the speaker is the only one that could verify the information (Lakoff 1973: 55), as in example (1) below.

(1) \*I am tired, aren't I?

As Lakoff asserts that the presence of tag-questions show that “expression of uncertainty is favored” in female speech even if the speaker has a high degree of certainty about their statement (1973: 45). Holmes set out to examine Lakoff’s assertions about tag-questions by analyzing a large corpus of data for pragmatic particles used as either a hedge or an intensifier (Holmes 1990). She attests that the functions of tag-questions are to request confirmation from the listener, facilitate the listener’s participation, mitigate or soften statements, or to challenge the listener (Holmes 1990). The last function is not one that Lakoff included in her claims. Holmes’ research found that while men and women both used tags, women used them more often to include the listener and men used them more often to express uncertainty (Holmes 1990). Only men used tag-questions to challenge the listener (Holmes 1990).

While similar in function and motivation to uptalk, tag-questions contrast with uptalk in two ways. First, uptalk is marked by intonation and not syntax, and second, uptalk can be used when the speaker is the only one that can confirm the information (Lakoff 1973: 55). Uptalk will be discussed further in Section 1.3.

#### *1.1.4 Politeness in performing gender*

Though most of Mill’s work, *Gender and politeness*, is concerned with politeness at the discourse level not the acoustic properties, she offers a nuanced look at what has been said previously about politeness in female speech and warns against false interpretations of data (Mills 2003). She argues that most of Lakoff’s and other’s examinations of female speech are based on stereotypes and not grounded in actual data. However, Mills acknowledges that stereotypes are still valuable as “those structural inequalities, and the stereotypes that we hypothesize on the basis of our knowledge of these inequalities, do play a role in the way that the interaction takes shape.” (Mills 2003)



One of the roles of stereotypes is in gender performance. Performance of gender involves behaving in a particular way to emphasize one's gender, usually through the use of stereotypes that may or may not be used naturally by people of that gender (Eckert & McConnell-Ginet 2003). Lexical, syntactic, and intonation norms in women's speech are used to perform gender. Women can perform their gender linguistically by choosing to speak in ways that are stereotypically associated with women such as using higher speaking pitch or more deference strategies. What is important is that these choices reflect what is believed to be true about women; it is not necessary for these choices to reflect what women actually do.

It is significant that the linguistic characteristics described above are categorized in the literature as "women's" speech, but gender may not actually be the determining factor that divides male and female speech styles. Lakoff suggests that these contrasts in speech patterns between men and women are not really due to differences in gender but differences in power and role in society, leading to higher use of deference strategies (1973: 53). Mills suggests that these contrasts may be a function of in-group, out-group identity (Mills 2003). People may use more deferent language when talking with people who do not belong to their in-group. Or they may set aside deferent language and use more familiarity when attempting to identify themselves as a member of the in-group (Mills 2003). However, the power discrepancy between men and women and distinction in their societal roles and identity groups persists. Consequently, it is possible to continue to use the label of "women's speech" until more research is done to determine what other communities use these same mitigating strategies. Then, the label "women's speech" can be refined.

### *1.1.5 Distinctions in conversation*

In male-female conversations, women use more strategies to promote interaction than men do (Fishman 1983). In Fishman's study, women asked more than double the

number of questions men did and gave more active listening feedback such as “mm”(1983: 94–95). The listening feedback women gave was done simultaneously with the male’s speech to demonstrate their participation in what he was saying. If the man gave listening feedback, it was at the end of the woman’s speech and discouraged interaction (Fishman 1983: 96). In female to female speech, conversation tends to focus on building rapport and interpersonal relationships rather than purely sharing information (Talbot 2020: 93). However, even if men and women tend to fill different conversational roles, they are not confined to them (Talbot 2020: 105).

### *1.1.6 Gender inequality*

In the 2020 *Gender Gap Report*, the World Economic Forum ranked the United States 53rd out of 153 countries in progress toward gender equality (Schwab et al. 2020). The most significant disparities between men and women worldwide exist in the areas of political empowerment and economic participation (Schwab et al. 2020). In the American workplace, there are overwhelmingly more men than women in high-profile positions. The Fortune 500 list for 2020 includes the highest number of female CEO’s yet — 37 (Catalyst 2020). Thirty-seven female CEO’s is a large increase from the 1970s and early 1980s when there were only two women on the Fortune 1000 list, but is still only 7.4% of all CEOs in Fortune 500 companies (Catalyst 2020). While there are many reasons for this, some that have nothing to do with gender inequality, it remains true that there are overwhelmingly more men than women in visible leadership positions in the workplace.

Because women were historically barred from political or economic positions of power, the women who join the workforce or get elected to office are entering into traditionally masculine communities. As they do so, they must also adopt new practices that may not be seen as masculine today but have their roots in masculine behavior. “Practices of masculinist communities have become naturalized as simply professional

practices.” (Talbot 2020: 195) Whether women are trying to perform or downplay their gender, show deference or assume power, or distance or align themselves with a particular group, the way they talk signals their identity to the world around them.

## **1.2 Pitch**

It is a truth universally acknowledged that it is not necessarily what you say that matters but how you say it. Kramer (1963) observed that vocal stereotypes abound and that people judge others based on their voices. These judgments range from how the speaker feels in the moment to physical characteristics to their personality and capability (Kramer 1963).

### *1.2.1 Vocal tract and average pitch*

The size of the larynx and vocal folds and length of the vocal tract regulates the possible frequencies a person can produce (Klofstad et al. 2016). The size of the larynx is a secondary sex characteristic (Lewis 2002). Typically, cisgender males have much larger larynxes than cisgender females due in part to higher testosterone levels, but there is wide variation between individuals (Klofstad et al. 2016). Each person can shorten and lengthen their vocal folds to produce different pitches, but these changes will always be limited by physiology (Klofstad et al. 2016). While men on average are larger than women and have lower voices, the size of an individual does not necessarily affect their average speaking pitch (Lass & Brown 1978).

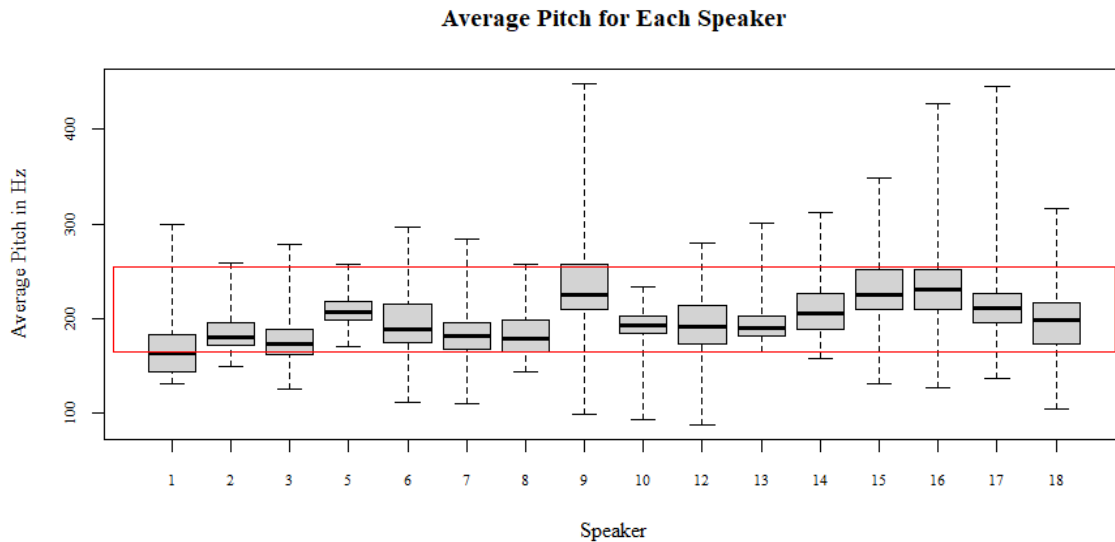
Klofstad (2016) proposes average speaking ranges for both men and women as shown in Table 1 below.

**Table 1. Average speaking pitch by gender (Klofstad et al. 2016)**

Gender	Average Pitch
Male	85-180 Hz
Female	165-255 Hz

My findings on the average speaking pitch over all speech styles were consistent with Klofstad et al. The red rectangle in Figure 1 below shows the range quoted by Klofstad overlaid on the results from my research. The average of all phrases from all three speech styles is shown.

**Figure 1. Average pitch of each speaker with range from Klofstad**



While some speakers, such as Speaker 9 and 16, had a large range, the middle 50% of phrases from all speakers fell within the parameters except Speaker 1 who had a lower average pitch than expected.

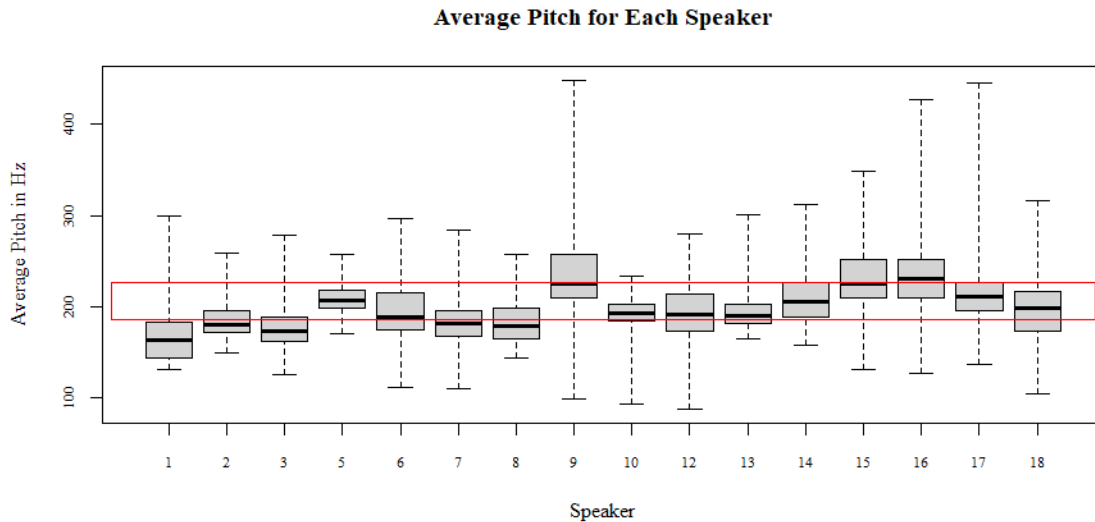
Shevchenko's (1999) study on the relationship between fundamental frequency in American English and class, gives average pitch for each class for both men and women shown in Table 2 below.

**Table 2. Average speaking pitch by gender and class (Shevchenko 1999)**

<b>Class</b>	<b>Men</b>	<b>Women</b>
<b>Upper</b>	102 Hz	185 Hz
<b>Middle</b>	111 Hz	227 Hz
<b>Lower</b>	126 Hz	198 Hz

Note that the lowest speaking pitch was found in the upper-class samples for both men and women. Lower speaking pitch from some women could be an attempt to imitate the speech styles of the upper class, but that is outside the scope of this project. Women had a larger pitch range than men and slightly higher variation within groups (Shevchenko 1999). Because I did not categorize the speakers in my study by class, I used Shevchenko's averages for all three classes to create a range that should encompass the mean pitch of speakers from any social class. Shevchenko's findings, marked by the red rectangle, are compared with my data in Figure 2 below.

**Figure 2. Average pitch of each speaker with range from Shevchenko**



As shown in Figure 2, Shevchenko’s averages do not match as well with my data as Klofstad’s. This difference may be because Shevchenko’s research was done over twenty years ago, and the average speaking pitch is lower now.

### *1.2.2 Lower voices signal dominance and leadership*

Apple et al. (1979) used recordings artificially manipulated for speech rate and pitch in their study and asked participants to judge if the speakers, all men, in the recordings were telling the truth. They found that lower pitch, even artificially controlled, correlated with higher truthfulness ratings. Further, “Men speaking in higher pitched voices were perceived as less potent (smaller, thinner, slower)” than those with lower pitched voices (Apple et al. 1979: 722). This perception may be because high testosterone levels are associated with lower speaking pitch in men (Dabbs Jr. & Mallinger 1999). Testosterone and other androgens are necessary for fertility in men (Dohle et al. 2003), and testosterone has been linked to dominance and violence (Dabbs

et al. 1995). This finding corresponds with the implication from Apple et al. that men with lower voices are perceived as more “potent” (1979).

The effects of testosterone and other pubescent hormones on the pitch of male voices have long been acknowledged. It was common practice until the late 18th century to castrate boys before puberty to preserve the high range of their singing voice (Britannica 2013). These singers, known as castrati, took the place of female performers in Roman Catholic church music since women were expected to be silent in church (Abitbol et al. 1999).

The study by Klofstad et al. (2012) surveyed people’s responses to the changes in pitch for the phrase “I urge you to vote for me this November.” The sentence was spoken by both men and women, and the recordings altered to produce a higher and lower pitched version from each speaker. Both men and women had overall preferences for lower pitched voices in both male and female speakers. Specifically, they found female speakers with lower pitch to be perceived as stronger, more competent, and more trustworthy (Klofstad et al. 2012: 2701).

### *1.2.3 Attractiveness*

It is difficult to isolate the effects of pitch on listener perception since every utterance is encoded with layers of information. Collins and Missing (2003) attempted to do this in their study on pitch and attractiveness by using only vowels as their speech data. People listened to recordings of vowels spoken by women and rated the speaker for attractiveness. The participants were shown pictures of the speakers, independent of their recordings, which they also rated for attractiveness and made a guess as to the speaker’s age. They found that women with higher speaking voices were judged as more attractive based on both their voice recordings and the pictures of their faces, and the participants guessed the more attractive speakers were younger than the other speakers (Collins & Missing 2003). Collins and Missing propose that these findings are connected

because high pitch is associated with youth and young people are deemed more attractive (reasons for this will be discussed below). However, age was not a significant factor in their tests because their sample had a small age range (Collins & Missing 2003: 1002).

Estrogen and progesterone are hormones that contribute to female fertility (Abitbol et al. 1999). These hormones seem to influence average speaking pitch. One study measured the speaking pitch of women during high-fertility and low-fertility times of their menstrual cycle and found that speaking pitch was higher in the high-fertility samples (Bryant & Haselton 2009).

It seems that evaluations of a woman as attractive are closely tied with subconscious evaluations of her fertility. When evaluated by their speaking voice, women are rated as more attractive when they seem to be young (Collins & Missing 2003) and when they are in high-fertility stages of their menstrual cycle (Pipitone & Gallup 2008). Since high pitch is associated with both youth and high-fertility, it follows that a woman will be rated as more attractive if she has a higher speaking voice. Therefore, women with higher voices are perceived as more attractive (Collins & Missing 2003) but also less competent (Montepare & Zebrowitz-McArthur 1987; Klofstad et al. 2012). These perceptions lead to conflicting motivations for pitch manipulation.

#### *1.2.4 Context-based accommodations*

The context of speech acts often determines the content of the speech, even acoustic characteristics. People may change their speech depending on the formality level of the situation or based on who is present in the conversation. Two studies, in particular, have looked at how people change their pitch to accommodate for the context or the interlocutors.



Kennard examined the speech of female drill instructors in training and compared their speech with male drill instructors in training (2006). She looked at vowel duration, syllable peak pitch, mean discourse pitch, standard deviation of pitch, pitch range, and speaking rate. Kennard found that male and female speakers patterned similarly in their use of pitch. The average speaking pitch for all participants was highest in the most authoritative speech style (Kennard 2006: 133). This finding is surprising given the perception of lower voices as more competent (Klofstad et al. 2012). Kennard suggests that the high pitch for women in the most authoritative speech style has to do with the interaction between pitch and amplitude (2006: 144). As the women are projecting their voices, they also raise their pitch. This explanation may be true of female drill instructors but does not necessarily explain why the male drill instructors also had higher pitch in this speech style.

The range of pitch was also similar for male and female drill instructors. The two types of authoritative speaking styles had larger variation of pitch than the two non-authoritative speaking styles (Kennard 2006).

Lewis' study focused on pitch changes in different interview styles prompted by the interlocutor — the person taking part in the conversation who is not being studied. As Lewis (2002) found, the kind of interlocutor present affects the person speaking as people accommodate their speech to the interlocutor in two ways: changing their speech to be what they think the interlocutor wants to hear or mimicking the interlocutor's speech (Lewis 2002: 28).

These accommodations can be contradictory. If a woman is speaking with a male interlocutor and wants to appear more feminine, raising pitch may be a way to perform gender (Talbot 2020) (Section 1.1.4). However, if she is accommodating through mimicry, her pitch would naturally be lower since men have lower voices than women on average (Section 1.2.1). If a woman is speaking with a female interlocutor and wants

to highlight camaraderie, she may use higher pitches (Lewis 2002: 31). If she wants to assert dominance, she may need to lower her pitch (Apple et al. 1979).

Lewis found that women had larger pitch ranges when speaking with female interlocutors (Lewis 2002: 59). When comparing data for all 12 female speakers in her study, there was no statistical difference in mean or median pitch in speech with male or female interlocutors, nor were there statistical differences in speech depending on the status of the interlocutors (Lewis 2002). However, six speakers had higher median pitches and larger ranges with female interlocutors to a significant level when examined as a subgroup (Lewis 2002). Additionally, four speakers had significant differences in their pitch based on the status of the interlocutors, but Lewis did not have a large enough sample with consistent results to measure statistical significance for this subgroup (Lewis 2002). These different accommodations based on the interlocutor suggests that women do make changes in their speech to attempt to control how they are perceived. It is important to note that Lewis' study looked only at pitch fluctuations and did not measure other kinds of changes or the effectiveness of those speech changes.

### **1.3 Uptalk**

Though it is often called an emerging bad habit of young female English speakers (Davis 2010), uptalk is neither new (Ching 1982), nor solely used by women (Armstrong et al. 2015), nor unique to English (Abe 1955).

Uptalk, while having similar intonation to a yes-no question, has a metapositional meaning instead of a propositional meaning (Guy et al. 1986). It may be that uptalk has evolved from tag-questions such as “you know?” or “right?” (Lakoff 1973) since it, at times, has a similar function.

### *1.3.1 Definition of uptalk*

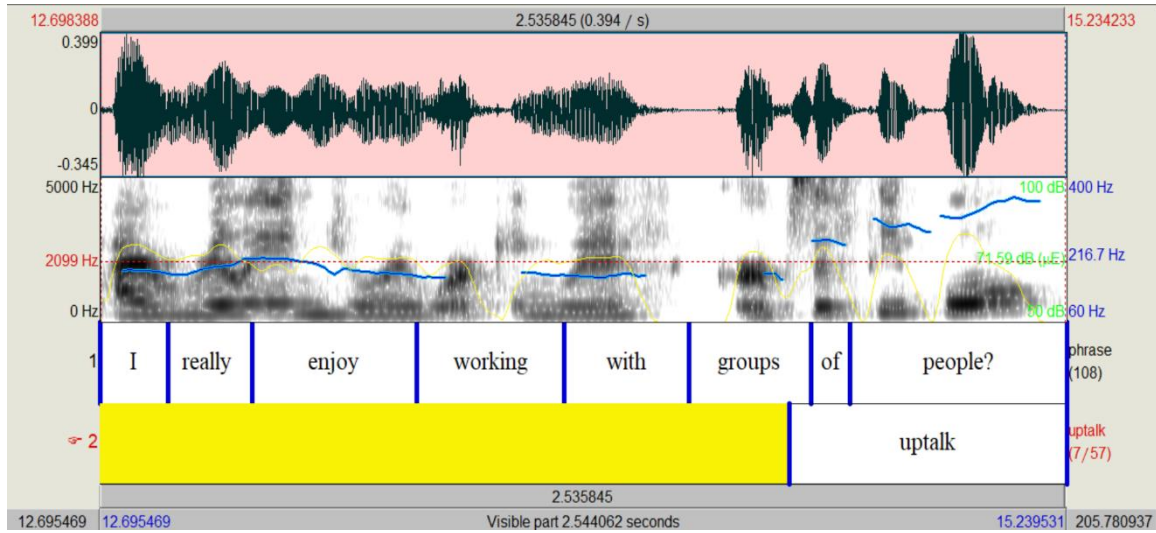
Uptalk is another speech characteristic evaluated in this research. Most basically, uptalk is rising pitch at the end of a phrase which may sound like the speaker is asking a question instead of making a statement (Di Gioacchino & Crook Jessop 2011). It is sometimes labeled high-rising terminal (HRT) intonation or some variation (Warren 2016: 5).

For this thesis, I have chosen Warren's definition of uptalk. Uptalk is "a marked rising intonation pattern found at the ends of intonation units realised on declarative utterances, and which serves primarily to check comprehension or to seek feedback." (Warren 2016: 2) According to this definition, there are then five primary features of uptalk.

1. Uptalk is marked.
2. Uptalk has rising intonation.
3. Uptalk happens at the end of phrases.
4. Uptalk occurs on statements, not questions.
5. Uptalk is interactive.

Figure 3 below shows an example of uptalk from my data. The blue lines through the spectrogram show pitch. When describing her strengths, Speaker 13 used uptalk on several statements. Each had a fairly level pitch at the beginning of the utterance and then a steep, sustained rise at the end.

Figure 3. Spectrogram of uptalk



The red dotted line through the middle of the spectrogram in Figure 3 shows the average pitch for the entire phrase. On the word “groups,” before the uptalk begins, the speaker’s pitch is around 175 Hz. Over the next 0.7 seconds, it rises to 353 Hz.

### 1.3.2 Origins of uptalk

Though uptalk is reviled in pop linguistics and the media (DiResta 2016), it seems uptalk has been claimed by several countries as their own distinguishing characteristic. Uptalk is often associated with Australian (Guy et al. 1986) or New Zealand English (Fletcher & Loakes 2006) where it is labeled Australian Questioning Intonation (Guy et al. 1986).

Many varieties of English around the world have some form of uptalk. In North America, it is found in California as a feature of so-called “Valley Girl speak” (Armstrong et al. 2015) and in Canada as “the Canuck uptalk epidemic” (Davis 2010).

Non-interrogative rising intonation has been noted by linguists since at least 1965 (Warren 2016: 56), but I believe the literature on American English describes uptalk even before that. In Ken Pike’s work *Intonation of American English*, published in 1945,

he notes “oftentimes rising contours are somewhat polite or cheerful, and sound less brusque than falling one.” (1945: 51) and gives both interrogative and declarative examples. He also comments that rising intonations require some sort of supplementation or response either from the listener or the speaker. Pike’s observations fit well with Warren’s definition of uptalk.

In fact, Pike describes a specific intonation pattern that sounds very much like uptalk in both realization and use. He illustrates what he calls a 2-1 rising contour, or high to super-high rise (Pike 1945: 59). He emphasizes its interactive function and use by women. “Hostesses use it a great deal for friendly welcome, or cheery inquiry or to put people at ease; in fact, all of the rising contours are used occasionally in polite or cheerful contexts, but probably this one more so than the others.” Again, his example sentences with this intonation are both questions and statements. One even includes a tag-question as discussed in Section 1.1.3.

Though it is beyond the scope of this thesis, uptalk is also attested in other languages in addition to English. A similar phenomenon is noted in Japanese as early as 1955 (Abe 1955: 392).

### *1.3.3 Societal views on uptalk*

Like vocal fry, uptalk is often blamed for discrimination against women in the workplace, often unequivocally. In a blog post memorably titled “Does uptalk make you upchuck?” DiResta (2016), claims that there is “no value in using uptalk.” She says, “Uptalk renders the speaker weak, tentative, lacking conviction and authority. How can a person influence, lead, or command respect if they can’t take a stand and sound like they mean it?”

“Valley Girl speak” is frequently used as a derogatory label for English spoken by (mostly young) women from Southern California (Nycum 2018). As uptalk is associated with the speech variety, it follows that the same stereotypes would be applied to uptalk

users as to Valley Girls such as the stigma that they are vapid and superficial (Nycum 2018).

As mentioned above, Davis (2010) calls uptalk an epidemic in Canada in his newspaper editorial. In an update on his article, he calls it an infection, spreading from California to Canada to the northern United States and now to England. He bewails the fact that uptalk is accepted as normal speech now and reiterates the widespread belief that people never used uptalk before the late 1990s (Davis 2010).

These views are from, as mentioned, popular linguistics or psychology articles. They are not substantiated by research or even backed by sources, but they do offer insight on what the average person's perception of uptalk may be and how it is discussed outside the realm of linguistic study.

#### *1.3.4 Questioning intonation*

In the examples below, the period indicates a falling intonation typical of simple declarative statements. The question mark indicates a rising intonation. The word order, that of a declarative sentence, is the same in all three examples.

(2) He's at the store.

(3) He's at the store? (speaker does not know this)

(4) He's at the store? (speaker knows this)

In example (2), the speaker is communicating information to the listener and not soliciting a response. Example (3) has a propositional meaning (Guy et al. 1986); the speaker is requesting information from the listener. A more standard word order might be "Is he at the store?" The word order the speaker has chosen in example (3) may indicate incredulity or surprise, but it is still a question. In example (4), the speaker is communicating information (the whereabouts of "he") and soliciting a response from the listener to acknowledge this information. This interaction is often found in narrative

speech as listeners need to track new information in order to understand what happens next in the story (Guy et al. 1986).

This use of uptalk is clear in the speech samples I collected. In the interviews, participants would often respond to the question “What are your professional strengths and weaknesses?” with statements with rising intonation. Since they are talking about themselves, they are not requesting information but soliciting a response from the interviewer, possibly an indication of what he thinks of what they are sharing. For this reason, Guy calls utterances with uptalk metapropositional questions (1986: 26). The women are making propositions about themselves at the lexicogrammatical level but using intonation to interact with the listener (Guy et al. 1986). DiResta (2016) assumes that the reason for interacting with the listener in this way is that the speakers are weak or insecure. Alternative motivations will be discussed in Section 1.3.7.

### *1.3.5 Contrast with continuation rise*

Uptalk is sometimes equated with a continuation rise (Di Gioacchino & Crook Jessop 2011). Both are rising intonation patterns and can be used to “hold the floor” (Warren 2016: 56), indicating that the speaker is not yet yielding their turn. However, phonetically, uptalk rises are generally higher and steeper than continuation rises (Guy et al. 1986). Uptalk and continuation rises differ in meaning as well. Continuation rises retain attention on the speaker while uptalk is fundamentally interactive, involving the listener in some capacity.

### *1.3.6 Uptalk users*

Many writers, including Lakoff, assume that uptalk is only used by women (Lakoff 1973: 55; Davis 2010; Warren 2016). In a study that compared male and female speakers of American English from Southern California and Massachusetts, female participants from Southern California produced the most rises over all, but there was no

significant difference in the number of rises produced by male and female participants from Massachusetts (Armstrong et al. 2015). This study is limited by its small sample size of eight speakers from each location. Still, it seems to indicate the possibility that uptalk is not limited to female speakers of American English. Another limitation of this study is that it included all non-question rises the speakers used. Not every rise is an example of uptalk.

Linneman examined contestants on the game show Jeopardy! and found that while both men and women use uptalk, they do so in differing degrees and for different reasons (Linneman 2013). Men used less uptalk than women overall, but their use of uptalk increased when they were with female contestants, when correcting female contestants, or when they gave incorrect answers (Linneman 2013).

Linneman concludes that the higher use of uptalk when around women or addressing women is a way to mitigate the power difference between men and women. He attributes the higher use of uptalk in incorrect answers as signaling uncertainty or lack of confidence. A man who is \$10,000 behind another contestant has a greater chance of using uptalk than a man who is \$10,000 ahead (Linneman 2013).

However, this effect is reversed for women. The more successful a woman is on the show, either by virtue of being a returning champion or by being ahead of the other contestants, the more likely she is to use uptalk in her responses (Linneman 2013). He suggests that women might use uptalk as a “compensatory strategy in order to perform their gender ‘correctly.’” (Linneman 2013)

### *1.3.7 Meaning and motivation*

The suggested motivations for uptalk are varied. Ching proposes six different motivations for using uptalk: topicalization, emphasis, checking for understanding/agreement, uncertainty, deference, and politeness (Ching 1982).



Warren has a more comprehensive list of the uses of uptalk: “checking, seeking a response, sharing, qualifying, connecting, floor holding, showing surprise, signaling information structure, and committing the interlocutors to the truth value of the content.” (Warren 2016: 56) While these motivations are more far-reaching than Ching’s, Warren does not explicitly mention politeness and deference, which are prominent functions in female speech in general and uptalk specifically.

While Lakoff does not discuss uptalk directly, the motivation of politeness for the use of tag-questions also seems to apply to uptalk. “One aspect of politeness is as we have just described: leaving a decision open, not imposing your mind, or views, or claims on anyone else. Thus a tag-questions is a kind of polite statement, in that it does not force agreement or belief on the addressee.” (Lakoff 1973: 56)

As stated in the definition of uptalk in Section 1.3.1, uptalk is essentially interactional. Speakers may be motivated to use uptalk to “interact and jointly construct a discourse.” (Guy et al. 1986: 25) It pulls the listener into the discourse and checks for understanding (Guy et al. 1986).

This interactional function of uptalk may perform a similar, though diminished role, as overlapping speech in the joint construction of discourse. Turn-taking in a conversation between women does not necessarily mean one person is talking at a time (Coates 2013: 23). Instead, simultaneous speech is a cooperative strategy so that both participants in the conversation contribute to creating the text of the conversation (Coates 2013). Women use overlapping speech more often than men, which is why women tend to use more uptalk than men (See Section 1.3.6).

## 1.4 Fry

### 1.4.1 Definition of fry

Historically, vocal fry was classified as a voice disorder (Hollien et al. 1966). Vocal fry is a subcategory of creaky voice phonation. Like typical creaky voice, vocal fry is produced by constriction of the glottis and low fundamental frequency (Keating & Garellek 2015). In an early study on vocal fry, the fundamental frequencies for segments with fry were “below and level that might be expected for a normal mode of phonation.” (Hollien & Wendahl 1968: 509) Fry is also characterized by audibly distinct pulses caused by dampening the vocal tract (Hollien & Wendahl 1968: 506).

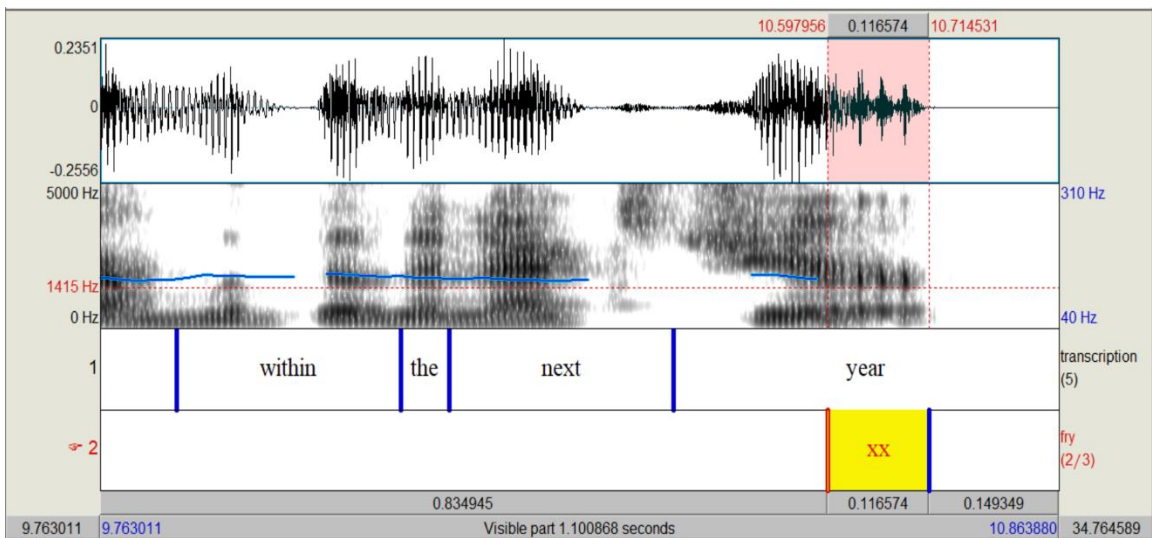
There is some discussion over whether these pulses are periodic or aperiodic. One of the distinctions Keating et al. make between vocal fry and prototypical creaky voice is that vocal fry has regular pulses and creaky voice does not (2015). In contrast, Kuang and Liberman claim that fry can be periodic, have double or triple periods, or be aperiodic (2016: 5260). Their definition of fry seems to combine Keating et al.’s definition of fry with their term “multiply pulsed voice” (2015). The difference between fry and multiply pulsed voice is whether or not  $F_0$  is low (Keating et al. 2015). Kuang and Liberman hold that the  $F_0$  does not necessarily have to be low for the speech to be classified as fry (2016: 5260). These differing opinions are probably due to the fact that Keating et al. are distinguishing between types of creaky voice and Kuang and Liberman are grouping them all under the category of fry.

For the most part, I used Keating’s definition of fry to label my data: constricted glottis, low  $F_0$ , and dampened pulses (Keating et al. 2015). However, in some instances, the duration of the fry was so short that the pulses could not be seen in the spectrogram. These examples looked more like Keating’s prototypical creaky voice, but most of what I labeled as fry in my data is periodic with regular, audible pulses. Because of the dramatic difference between creaky voice and modal voice,  $F_0$  in creaky

voice is difficult to measure with automatic pitch tracking in Praat (Boersma & Weenik 2019).

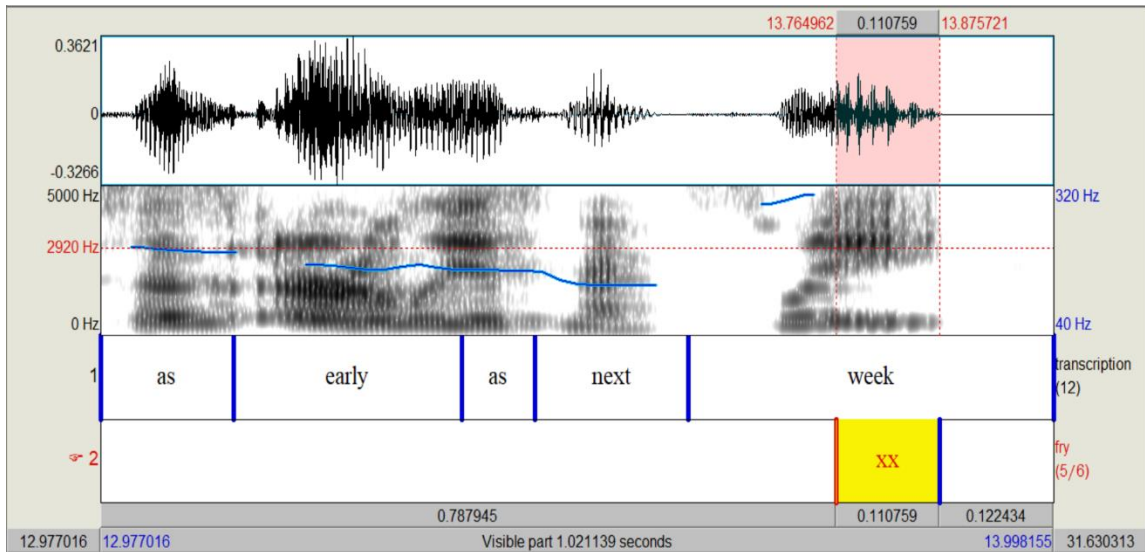
Though the measurements of F0 are inconsistent, the auditory perception of creaky voice in my data is of low pitch in all but one instance. Figure 4 below from Speaker 1's presentation shows periodic fry with low perceived pitch. The pulses can be seen in the fry at the end of the phrase.

**Figure 4. Fry with low perceived pitch**



Praat does not measure any pitch on the fry, but the auditory perception is low. In the very next phrase, shown in Figure 5, Speaker 1 uses fry on a word in the context of high pitch.

**Figure 5. Fry with high perceived pitch**



Again, Praat does not measure F0 on the part of the word that has fry, but the auditory perception is of high pitch because of the high pitch at the beginning of the word. This fry is also periodic and the pulses can be seen in the spectrogram as well as perceived audibly. Figure 5 is the only example in my data of fry occurring in the context of high pitch.

#### 1.4.2 Frog voice

In Kennard’s research on the speech styles of drill instructors, she notes a voice quality known as “Frog Voice” used by male drill instructors (Kennard 2006). Frog Voice seems to be extended use of vocal fry to sound “intimidating” and “powerful” (Kennard 2006: 149). Interestingly, female drill instructors reported that they were not allowed to use Frog Voice, and some Commands do not allow any drill instructors, male or female, to use it because it comes across as unnecessarily aggressive (Kennard 2006).

### **1.4.3 Perception**

Because vocal fry is usually produced with low fundamental frequency, the presence of artificially emulated fry conditions listeners to perceive a lower pitch (Kuang & Liberman 2016). The more fry present, the lower the perceived pitch (Kuang & Liberman 2016). This perception may motivate young female speakers to use more fry to sound like they have a lower speaking voice.

## **1.5 Conflicting motivations in young women's speech**

In the introduction, I outlined three quandaries for young women. Lower pitch may make them sound more competent (Klofstad et al. 2016) but less attractive (Collins & Missing 2003). Lower pitch may trigger vocal fry, which may, in turn, affect perceptions of competency. Uptalk may be used more by young women when trying to sound polite or check that the listener is following them (Guy et al. 1986), but the listener may interpret it as uncertainty (Warren 2016). In any conversation, there are many different motivations for how people speak. Some of these motivations may be conscious but many are also unconscious (Mills 2003). As Mills says, “speakers do not necessarily decide consciously to utter every word that they say; there are other factors, such as subconscious motivations, verbal routines, and social pressures which play a major role too.” (2003) As native speakers are not usually aware of either their own speech characteristics or the motivations behind them, there is no way to assign motivations definitely. Nonetheless, the characteristics themselves can be measured and analyzed.

For young women, the desire to perform their gender “correctly” may lead them to make certain lexical choices, use particular intonation patterns, or raise their speaking pitch. However, the desire to appear competent or authoritative may lead them to lower their speaking pitch and increase their use of vocal fry. Further entangling the

motivations behind choices in speech acts is the presence of the interlocutor. The speaker makes choices based on the perceived understanding of the listener's interpretation of these choices. This perception may or may not be accurate. Finally, the use of speech stereotypes may be construed differently based on unknown or unrecognized factors in the minds of both the speaker and the listener.

These motivations are not uniform for each person in every context. For example, in a conversation with a peer, a young woman might not be concerned about how authoritative she sounds, but she may be controlling how nice she sounds. Another woman in a conversation with a male supervisor may be more motivated to use stereotypical female speech than she would with a female supervisor because she is performing her gender to a higher degree. Some women may be concerned with how polite they sound and increase their use of uptalk. In contrast, others discard deference strategies with superiors in order to identify as a member of the in-group and lessen the social distance between them.

## **1.6 Videoconferencing**

One of the hallmarks of the COVID-19 pandemic in 2020 was the increase in the necessity of videoconferencing as more and more people worked from home instead of meeting in person. An Ipsos poll in April 2020 found that 57% of Americans they surveyed were using a video chat platform for work or social meetings (Newall & Chen 2020). Zoom (Yuan 2012), one of the most popular videoconferencing programs, had 10 million daily participants in December 2019; by April 2020, there were 300 million daily meeting participants (Iqbal 2020).

Though more and more people are using various videoconferencing platforms, there are some drawbacks to using videoconferencing in the context of linguistic research. These can be as straightforward as internet connectivity problems or other

technology issues taking the focus off the task at hand. It is difficult for people to separate technology issues from the conversation itself. Delays in transmission of teleconference calls can cause people to think that the other person in the call is not paying attention to the conversation (Schoenenberg et al. 2014). As familiarity with videoconferencing technology rises and there is wider access to high-speed internet, these interruptions become less impactful but are never fully eliminated.

One of the benefits of using videoconferencing for data collection in lieu of in-person methods is that participants can join from anywhere in the world. This access would help researchers expand their recruiting pool without incurring extra expense or needing to travel. Virtual data collection also cuts some of the cost for participants. They do not need to be in a specific place, eliminating some of the time cost in participating in studies.

While videoconferencing does seem to mimic real-life interactions, it is not the same as face-to-face conversation. In a study on “Zoom fatigue,” exhaustion caused by being on video chat for extended periods, Bailenson outlines four differences between videoconferencing and in-person conversations: eye contact, mobility, non-verbal signals, and self-view (2021: 1).

In videoconferencing calls, since only the participants’ faces are in view, all participants are receiving eye contact simultaneously for long durations (Bailenson 2021: 2). This amount of eye contact would not be possible in an in-person conversation with more than one person, and such close eye contact is usually reserved for intimate relationships, not coworkers and supervisors (Bailenson 2021: 2).

Another consequence of seeing only faces in a conversation is a change in nonverbal cues. Body language is much harder to discern, but eye movement becomes more important (Bailenson 2021: 3), leading participants to exaggerate their listening feedback, such as nodding and monitoring their other nonverbal cues more closely

(Bailenson 2021: 3). In videoconferencing, visual nonverbal cues include nodding, gaze aversion, smiling, forward leaning, facial touching, and postural matching (Croes et al. 2019: 1217). Croes et al. found that, while there were few significant differences in visual cues between virtual and real conversations, participants smiled more in videoconferencing calls than in face-to-face communication (2019).

Croes et al. also looked at vocal cues in videoconferencing (2019: 1219). Participants in their study spoke louder and varied their pitch less during videoconferencing calls than in-person conversation (Croes et al. 2019: 1223). In addition, they found that speech rate enhanced social attraction in video calls but had no effect in face-to-face conversations even though there was not a significant difference in speech rate between the two modes of communication (Croes et al. 2019: 1224). Thus, while some of the nonverbal cues are the same in video calls or in-person conversations, people may use and interpret the cues differently.

While the close-up camera view in video calls changes the range of observable gesture and nonverbal cues, it also limits the overall mobility of the participants since they are tied to one spot — generally in front of a computer (Bailenson 2021: 4). This lack of mobility most likely has a larger effect on creativity (Opezzo & Schwartz 2014) and perceived autonomy than actual communication.

Self-view may be the most important and least-considered difference for women between videoconferencing and other forms of communication. The default version for most videoconferencing platforms is to enable users to see their own video as well as the videos of others. Studies by Ingram et al. show that women tend to focus on themselves more than men do when seeing a video or mirror image of themselves (1988). When the women in the study were more focused on themselves, they had higher levels of social anxiety during the activity and more negative views of themselves (Ingram et al. 1988). Gender roles also seemed to play a part in women's



responses to the study. The highest self-focus and negative self-image scores were among the women who identified as more feminine than others (Ingram et al. 1988). This finding suggests that self-monitoring and evaluating perception by others is part of the performance of female gender identity.

A more recent study found that while viewing curated social profiles could improve self-esteem, non-edited self-view decreased self-esteem (Gonzales & Hancock 2011: 82). This study used a much smaller number of male participants than female participants and did not find that gender had a significant effect on self-esteem results (Gonzales & Hancock 2011: 81). Zoom does offer some filters that may enhance the mirror image and mitigate the effects of self-view on self-esteem or even provide a confidence boost, but no studies have been done using such filters or measuring the impact of seeing a mirror image for several hours a day.

Videoconferencing does not directly correlate with face-to-face communication. Some nonverbal signals, such as body language, are lost in video calls, and some signals change meaning or are given more importance, such as eye movement and speech rate. It is difficult to isolate any one aspect of communication from the whole act of communication, so it is necessary to keep in mind these differences between in-person and computer-mediated communication.

## CHAPTER 2 METHODOLOGY

One of the inadequacies of the research published so far on female speech is that few studies look at both production and perception based on the same data. Most do not take into account how speech changes in different contexts. There have been studies on how women change their pitch in different situations based on who is present or how much authority they want to present (Lewis 2002; Kennard 2006). There have been studies on the listeners' attitudes toward changes in pitch (Klofstad et al. 2012; Collins & Missing 2003) but neither of these studies used extended samples of speech in an attempt to isolate the effects of pitch.

For this paper, I have collected data on both production and perception of speech changes in three contexts: presentation, casual speech, and interviews. I have called participants in the production stage “speakers” to distinguish them from those participants who responded to the perception survey.

### **2.1 Stage one: Production**

The first stage of the research was recording video conference calls with pairs of cisgender female speakers using Zoom (Yuan 2012) version 5.3.1. Zoom is a good choice for linguistic data collection because the developers are aware of the impacts of delays in conversations and seek to optimize the program by keeping delays below 150 milliseconds (Pierce 2020). While most videoconferencing software allows calls to be recorded, Zoom records the audio and video of each participant separately which provides clear data, especially in the case of overlapping speech.

Speakers were recruited via social media and word of mouth. There were 18 speakers total between the ages of 18 and 35. The video calls were done with pairs of speakers who knew each other and signed up for the call together. I had previously met half of the speakers, but half of them did not know me. Before joining the video call,

speakers completed a survey supplying demographic information and consenting to the research. The research consent form included the title of the research project as “Acoustic characteristics of female speech” and gave the research goal “to understand speech patterns to promote ongoing societal equity.” The exact research questions were not disclosed to any of the participants before the video conference calls.

To simulate a higher formality context, an interviewer participated in the research. He was male, over the age of 60, and introduced himself to the participants by mentioning his title as former Chief Financial Officer of a company and his qualifications of multiple master’s degrees. He was given a short overview of the research methodology, but he was not briefed on the research questions until after the last interview. None of the speakers knew the interviewer personally. The interviewer was not present at the beginning of the videoconferencing call, joining the call after the casual and presentation speech samples had been collected. During the interview, I turned off my video so the participants could not see me, but I continued to monitor the call. The interviewer interacted with the speakers while I communicated with him via text chat.

At the beginning of the call, I met with the pair of speakers, without the interviewer present, to review the consent form and the research process. I used a script to provide some uniformity. First, speakers took turns telling a story about a favorite childhood memory or a first date. If the stories were short, they were prompted to tell another or elaborate until they had spoken for at least a minute and a half. Some stories were up to three minutes long. These stories were used as the casual speech sample. After both speakers had shared a memory, they took turns reading a paragraph from my shared screen as if they were giving a presentation to a group of people. I then turned off my video camera and changed my username to “inactive host”. Next, the interviewer joined the call and interviewed both speakers, usually switching between them after one

or two questions. The interviewer generally asked each speaker the same questions and asked follow-up questions as appropriate to attempt to simulate an actual interview. This continuity allowed for relatively controlled content of the interview samples while not overly restricting free speech.

The most interactive speech style was the interview. There was also a little interaction in the casual speech sample for some of the speakers. This interaction typically happened either at the beginning of the story, during the scene setting, or toward the end of the story, during the dramatic climax. Because Zoom records each person in the call individually, the interactive speech did not obscure the data except in the case of Speakers 1 and 2 as noted below.

### *2.1.1 Anomalies*

Most speakers joined the Zoom call on a laptop, but some used mobile devices. Speakers 1 and 2 were in the same room and were using one device for the audio until they were asked to change location after sharing the casual speech sample. Speaker 4 could not get her device's audio to work, so she called in using a phone. Her video was on, but the interviewer could not see her. Speaker 13 had loud background noise, purportedly from the computer's fan. Before the recording began, she was asked to turn the Zoom background noise filter to medium. Speaker 15 had a poor connection and was interrupted by a train. Both 15 and 16 used mobile devices. Speakers 17 and 18 both had a poor connection, and their sessions were disjointed. Speaker 17 dropped the call and had to rejoin the session.

### *2.1.2 Recording preparation*

After the calls ended, the recordings were rendered by Zoom. Each call produced several recording types: m4am m3u, and mp4 files recording the whole group and

individual m4a files for each user present in the call. If the chat box was used, there was also a txt file of the chat history.

The individual m4a files were prepared for analysis with Audacity® editing and recording software version 2.3.2 with LAME MP3 library 3.100<sup>2</sup>. I divided each recording by speech style using labels. Each label was then exported into its own mp3 file. The title for the file included the date, the participant's number, and speech type. The export settings are shown in Table 3 below.

**Table 3. Audacity settings for export**

<b>Bitrate Mode:</b>	Preset
<b>Quality:</b>	Insane 320 kpbs
<b>Variable Speed:</b>	Fast
<b>Channel Mode:</b>	Joint Stereo

Speakers 10, 13, and 17 had background noise affecting their recordings. I used the noise reduction tool in Audacity to remove some of the interference so that the recordings could be used in the study. Using the noise reduction tool did not change the pitch or fry readings for the recordings.

Both Speaker 4 and Speaker 11 were removed from the study. Because Speaker 4 had called in on the phone, the quality of the recordings was much lower than the other

---

<sup>2</sup> Audacity® software is copyright © 1999-2021 Audacity Team.  
Web site: <https://audacityteam.org/>. It is free software distributed under the terms of the GNU General Public License.  
The name Audacity® is a registered trademark.

speakers. Speaker 11 had too much interference in her recordings that could not be mitigated by Audacity's noise reduction tool.

### *2.1.3 Analysis of recordings*

I used Praat (Boersma & Weenik 2019) to prepare and analyze the recordings. Each speech sample was opened with a text grid which I used to mark and label phrase boundaries and instances of fry, rising intonation, and miscellaneous notes as shown in Table 4.

**Table 4. Text grid tiers — labels and descriptions**

<b>Phrase</b>	Label for each phrase
<b>Pitch - point tier</b>	Marked at the point of highest intensity for each vowel
<b>Fry</b>	Duration of fry within a phrase
<b>Uptalk</b>	Segments with rising intonation
<b>Flag</b>	Miscellaneous notes or excluded segments

Phrase boundaries were determined when there was a pause or a marked change in thought or intonation. Phrase boundaries were manipulated where possible to exclude filler words (um, uh, etc.). Some filler words that occurred mid-phrase could not be excluded, but they were not included in the fry measurements. Where the length of the phrase was noted, this measurement includes only the time the participant was actively speaking, not pauses between phrases.

### *2.1.4 Pitch measurements*

Because each speaker was in a unique environment using unique equipment, the analysis parameters were unique for each speaker and adjusted as necessary so that the data Praat produced matched the output of the speaker. To determine the initial

settings, the pitch range was set to 10Hz-1000 Hz, and the voicing threshold was set to 0.6. I selected the entire recording sample and recorded the minimum/maximum pitch. I looked through the sample for obvious misreadings or interference, such as extremely low or extremely high pitch readings or dramatic leaps in pitch. I manually measured the highest pitch and the lowest pitch without the misreadings. The parameters were then set for each sample about 10 hz below the minimum and 50 Hz above maximum.

Each phrase was measured for lowest pitch, highest pitch, and average pitch. From these measurements, the range of each phrase was calculated. Occasionally, analysis parameters had to be adjusted because of anomalies in the speaker's environment, computer interference noises, or misreadings in Praat. Usually these adjustments involved changing the minimum or maximum Hz range or the voicing threshold. Occasionally, the octave jump cost was raised as well. It is worth noting that when these adjustments changed the minimum or maximum pitch reading of a phrase, the effect on the mean pitch reading was negligible.

I confirmed the validity of the phrase-based pitch measurements as opposed to measuring only the pitch on each vowel using a selection of the three recordings for Speaker 1. I checked that the pitch measurements of each phrase were similar to the average F0 frequency for each vowel. This check was to determine whether or not the phrase-based measurements were affected by voiced segments that were not vowels. I added a point tier to the Praat Text Grid and added a point close to the highest intensity of the vowel. I then extracted the pitch measurement at all points and averaged the results for each phrase. These averages were similar to the phrase-based measurements — 77% of the measurements using these two methods were within 10 Hz of each other. The difference in average pitch exceeded 20 Hz in only two phrases. In the cases where the averages were different, the phrase-based measurements seemed to represent the data more accurately because longer vowels were weighted more than shorter vowels

whereas in the point method, all vowels received one measurement regardless of length. Therefore, the phrase-based method was used throughout for all speakers.

### *2.1.5 Fry measurements*

Each phrase was marked for the presence of fry as a binary measurement as well as the amount of fry in each phrase as a categorical variable as shown in Table 5 below.

**Table 5. Amount of fry in each phrase**

<b>Label</b>	<b>Description</b>
0	No fry
1	Fry on one word
2	Fry on more than one word but less than half the phrase
3	Fry on more than half of the phrase
4	Fry on the whole phrase

If one of the instances of fry occurred following or replacing a glottal stop 0.5 was added to the category label. Therefore, a phrase that had fry during more than half of the phrase and had fry replacing a glottal stop would receive a score of 3.5. The label 4.5 was not used since a phrase in category 4 had fry on the entire phrase.

I noted whether the fry occurred phrase initial, phrase medial, or phrase final.

### *2.1.6 Uptalk*

All phrases were marked first for rising intonation — the most basic component of uptalk — and then sorted into uptalk rises or other kinds of rising intonation using the five criteria given in Section 1.3.1. The markedness of the rise was determined by comparing the steepness of the rises used by that particular speaker as well as pauses following the rise to signal importance or allow for listening feedback. Because phrase



boundaries were manipulated to exclude computer interference and make measuring pitch possible, some instances of uptalk did not occur at the end of the phrase as marked in Praat, but all occurred at the end of an intonational unit.

## **2.2 Stage two: Perception**

The second part of the research was a matched guise test using nine recordings from the first stage and a questionnaire evaluating audience attitudes to speaking pitch and use of vocal fry and uptalk. I have called participants in the production stage “speakers” to distinguish them from those participants who responded to the perception survey. Speakers in the first stage did not answer the survey in the second stage. Each participant listened to nine recordings, each around a minute in length. The nine recordings were portions of the three speech styles from three speakers. The selection of the recordings is outlined below. After listening to one recording, the participants ranked the speaker using a Likert scale before continuing to the next recording. After all the recordings had been played, the participants affirmed that they did not know the speakers in the recordings and answered two questions about their attitude toward female leaders in general.

### *2.2.1 Selection of recordings*

Speakers 1, 7, and 13 were chosen for the perception survey. Each was chosen because they exhibited marked changes in one of the vocal characteristics in question and marginal or predictable changes in the others.

Speaker 1 was chosen for pitch because she had the greatest variation in pitch among the speech styles. She also had a similar percentage of phrases with some degree of fry in all speech samples and only a small amount of uptalk in the interview sample, as shown in Table 6.

**Table 6. Measurements for Speaker 1**

	<b>Mean Pitch</b>	<b>Fry</b>	<b>Uptalk</b>
<b>Speaker 1: Presentation</b>	212.75	67%	0%
<b>Speaker 1: Casual</b>	175	73%	0%
<b>Speaker 1: Interview</b>	155.6047	70%	7%

Speaker 7 was chosen for fry because she had a wider variation of percentage of phrases with fry between the speech styles than the other speakers. She had a similar average pitch for each speech style and minimal use of uptalk only in the interview style, as shown in Table 7 below.

**Table 7. Measurements for Speaker 7**

	<b>Mean Pitch</b>	<b>Fry</b>	<b>Uptalk</b>
<b>Speaker 7: Presentation</b>	198.3	60%	0%
<b>Speaker 7: Casual</b>	193.5652	30%	0%
<b>Speaker 7: Interview</b>	172.6	55%	8%

Table 7 represents the overall measurements for each characteristic for Speaker 7. The selection from the casual and interview speech used in the perception survey had slightly different percentages of fry, 35% and 45%, respectively.

Speaker 13 was chosen for uptalk because she had a noticeable increase in uptalk in the interview speech. Her pitch variation between speech styles was small, but her fry use was much higher in the presentation speech shown in Table 8 below.

**Table 8. Measurements for Speaker 13**

	<b>Mean Pitch in Hz</b>	<b>Fry</b>	<b>Uptalk</b>
<b>Speaker 13: Presentation</b>	203.3333	83%	0%
<b>Speaker 13: Casual</b>	194.86	67%	2%
<b>Speaker 13: Interview</b>	192.7945	55%	31%

Even with the variation in fry, Speaker 13 was still the best choice for the perception test because the other speakers with similar variation in uptalk had abnormal mean pitch distribution or shared more identifying information in their free speech samples.

The speakers had neutral accents, and it was possible to get selections from their recordings that were coherent and had no identifying information. Selections were taken from casual and interview speech samples that were approximately one minute long. The presentation sample was used in its entirety since it ranged from 38 seconds to 43 seconds long. Each selection was checked to ensure that the average pitch and instances of fry and uptalk were similar to what was measured for the whole sample.

I endeavored to minimize as much distracting input as I could. Any mistakes in reading the presentation prompt were removed. Filler language (such as *um*) was removed when it did not occur in the middle of a phrase. I used the noise reduction tool in Audacity and normalized all the recordings so that changes in background noise or volume would not distract the participants. I also checked the content of each recording so that nothing said would inherently provoke an emotional response from the participants or give personally identifying information about the speaker.

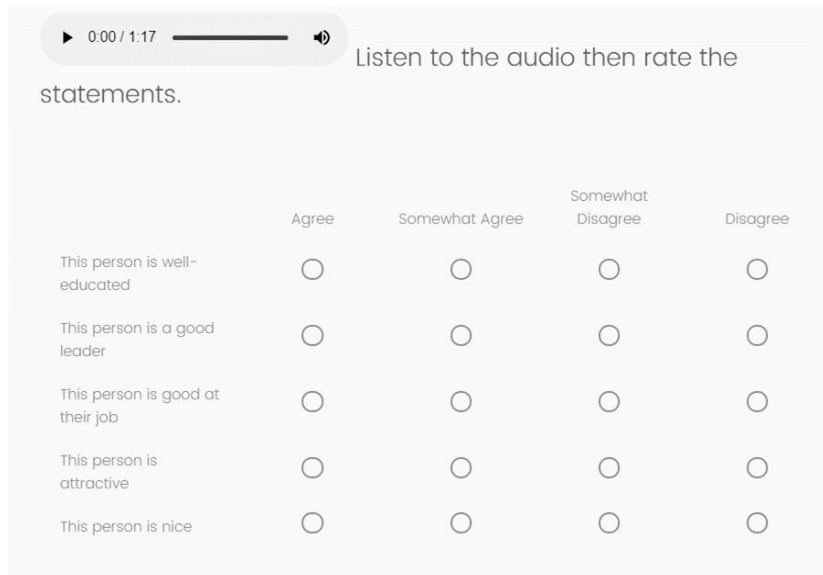
### 2.2.2 Surveys

Seven surveys were created to present the recordings in various positions to the participants. To determine the order of playing the recordings for each participant, I generated random strings of nine numbers until I got seven orders that did not have the same speaker played twice in a row or the presentation speech sample twice in a row.

A total of 45 participants successfully completed the surveys. No participants reported any difficulties in hearing or processing speech. Responses were removed from the data if the participants affirmed that they personally knew one or more of the speakers after listening to all the recordings. The only demographic information collected about the participants was their age and gender.

Each participant listened to the recording and then ranked the statements that followed using a 4-point Likert scale. The same statements followed each recording. An example from the survey is shown in Figure 6 below.

**Figure 6. Example from perception survey**



The screenshot shows a survey interface. At the top, there is an audio player with a play button, a progress bar showing 0:00 / 1:17, and a volume icon. To the right of the audio player, the text reads "Listen to the audio then rate the statements." Below this, there is a table with five rows of statements and four columns of response options: "Agree", "Somewhat Agree", "Somewhat Disagree", and "Disagree". Each cell in the table contains a radio button.

	Agree	Somewhat Agree	Somewhat Disagree	Disagree
This person is well-educated	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
This person is a good leader	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
This person is good at their job	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
This person is attractive	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
This person is nice	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

The first three statements were related to competency and authority while the last two addressed affability and attraction. Because there was no way to ensure that the participants actually listened to the audio, a timer was used on each question so they had to wait approximately the length of the recording before they could move to the next question.

The final two questions on the survey were designed to ascertain the participants' experience and attitude toward female leaders and supervisors in general. The questions and possible responses are shown in Figure 7 below.

**Figure 7. Final questions on perceptions survey**

In your experience, who are better leaders?

- Men
- Women
- Both
- Prefer not to answer

---

Have your supervisors been mostly men or women?

- Mostly men
- Mostly women
- Both
- Prefer not to answer

Though the surveys were completed anonymously, I did receive unsolicited feedback from some participants about the recordings. Some of these comments are discussed in Chapter 4.

## CHAPTER 3 RESULTS

### 3.1 Overview of analysis

This chapter shows the findings for both production and perception. For pitch, uptalk, and fry, each section first gives a description of the results followed by the statistical results and brief discussion of the findings. Further discussion of findings is in Chapter 4. The results from the perception survey are given in Section 3.5.

#### 3.1.1 *Number of phrases*

Because each speech sample was a different length, the number of phrases varied for each speaker in each speech sample. As described in Section 2.1.3, phrase boundaries were marked when the speaker paused or had a distinct change in thought or intonation. Even though the presentation speech sample was based on a script, the intonation used by each speaker was unique, and phrase boundaries were determined individually. Table 9 below shows the number of phrases in each speech style for each speaker.

**Table 9. Number of phrases in each speech style by speaker**

<b>Speaker</b>	<b>Presentation</b>	<b>Casual</b>	<b>Interview</b>	<b>Total</b>
<b>S1</b>	12	28	43	83
<b>S2</b>	12	54	74	140
<b>S3</b>	9	27	35	71
<b>S5</b>	11	47	63	121
<b>S6</b>	8	38	50	96
<b>S7</b>	10	23	51	84
<b>S8</b>	12	38	58	108
<b>S9</b>	10	38	62	110
<b>S10</b>	11	33	51	95
<b>S12</b>	13	16	51	80
<b>S13</b>	12	51	75	138
<b>S14</b>	14	58	63	135
<b>S15</b>	11	47	53	111
<b>S16</b>	12	46	58	116
<b>S17</b>	11	32	44	87
<b>S18</b>	12	62	43	117
<b>Total</b>	180	638	874	1692
<b>Average number of phrases per speaker</b>	11	40	55	106

### 3.1.2 Statistical methods

Fundamental frequency, uptalk, and fry were analyzed separately using R statistical software (R Core Team 2020). For fundamental frequency, three repeated-measures (within-speaker) analyses of variation (ANOVAs) were performed using the linear mixed-effects regression test of the lmerTest package in R. The results are shown in Section 3.2.1. For uptalk (Section 3.3.2) and fry (Section 3.4.1), the statistical tests used generalized mixed-effects regression.

**Style** was a factor variable with three levels: Presentation, Casual, and Interview. **PhraseDuration** was a continuous variable measuring the length of each phrase. **FryScalar** was a categorical variable with nine levels as shown in Section 2.1.5. The location of the fry in the phrase, whether it was initial, medial, or final, was coded as both a dependent and independent variable, **FryLocationInd** and **FryLocationDep**, respectively. Dichotomous variables marked the presence or absence of fry and uptalk in a phrase, coded separately as dependent and independent variables. These are **FryDummyDep** and **FryDummyInd** ( $\pm$  fry), **UptalkDummyInd** and **UptalkDummyDep** ( $\pm$  uptalk). Finally, each phrase had four continuous variables for F0: **f0min**, **f0max**, **f0range** (F0 max minus F0 min), and **f0mean**.

Demographic information collected for each speaker included age, education, what state they currently live in, other places they have lived, race/ethnicity, community type, profession, whether or not they currently held a supervisory position, and other languages they spoke. Each participant was sorted into a particular geographic region based on the states they had lived in. These regions are: East, North Midwest, Midwest, South, and West. The complete survey can be found in Appendix D. All responses were recoded as integers.



### 3.2 Pitch results

I set out to answer my research questions noted in the introduction with the hypothesis that speakers change their pitch depending on speech style, with interview being lowest, casual in the middle, and presentation being the highest.

$H_0$  = F0 (mean pitch) is the same for each speech style

$H_1$  = F0 (mean pitch) is lower for interview speech than casual and presentation speech

$H_2$  = F0 (mean pitch) is higher for presentation speech than casual and interview speech.

F0 = presentation > casual > interview

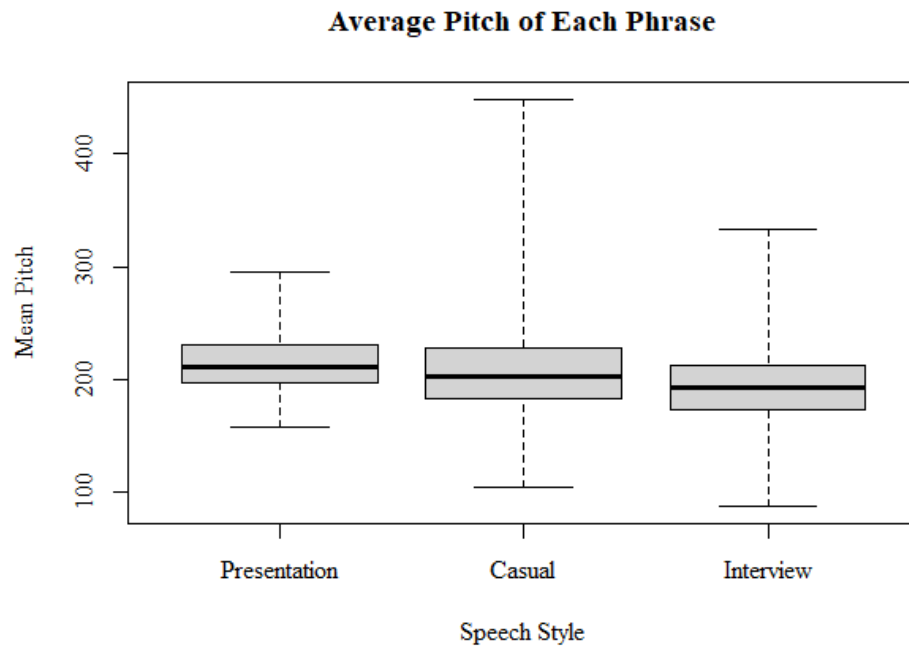
The average pitch for each style is shown in Table 10 below. The preliminary results seem to align with the first hypothesis. Obviously, statistical testing is needed to confirm this and is presented below in Section 3.2.1.

**Table 10. Average pitch for each style**

<b>Style</b>	<b>Mean Pitch</b>	<b>Standard Deviation</b>
Presentation	213 Hz	24 Hz
Casual	209 Hz	43 Hz
Interview	194 Hz	32 Hz

The casual and presentation mean pitch are quite close together, but casual speech has a larger standard deviation, as can be seen in Table 10, and greatest variation of the three speech styles, as shown in Figure 8.

**Figure 8. Average pitch by style**



While I was originally focused on average pitch across speech styles, I was surprised by the difference in range of pitch across speech styles. Casual speech has, by far, the largest variation in average pitch of each phrase. The range of pitch across speech styles had not surfaced previously in the literature though Talbot notes that women tend to use more of their vocal range than men (Talbot 2020: 26). Possible reasons for the larger range of pitch in casual speech will be discussed in Section 4.3.

### *3.2.1 Statistical tests for pitch*

A linear mixed-effects regression model was used to analyze the differences in pitch between speech styles. Four tests were run for four measures of pitch: average, minimum, maximum, and range of each phrase. Any phrases that were not measured for pitch, usually because the entire phrase had fry, were excluded from the data for these tests. Of the 1692 total phrases, 21 were excluded from the tests on pitch.

As mentioned in Section 3.1.2, **Style** was coded as a factor variable with three levels: Presentation, Casual, and Interview. **Age** and **Education** are categorical variables from the participants' responses to the survey. All of these were recoded as integers.

Three within-speaker ANOVAs were performed in four separate tests. The dependent variable in each test is pitch (**f0min**, **f0max**, **f0range**, or **f0mean**) and the independent variables are speech style, age, and education of the speakers.

$$lmer(f0 \sim (Style + Age + Education) + (1|Speaker))$$

The effect of speech style was significant for all four tests of pitch, as shown in Table 11.

**Table 11. Statistical results for pitch and style**

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
<b>F0 Mean</b>	108217	54109	2	1653.6	59.0666	<2e-16 ***
<b>F0 Range</b>	372440	186220	2	1655.72	40.2506	<2e-16 ***
<b>F0 Min</b>	15426.1	7713.0	2	1653.73	5.9051	0.002783 **
<b>F0 Max</b>	439100	219550	2	1654.24	55.7130	<2e-16 ***

Note that, while the effect of speech style was still significant for minimum pitch, it has a larger p-value than the other tests. In speech, people tend to approach their lower limit for producible pitch but not the upper limit of producible pitch, so the minimum pitch had less variation in different speech styles.

Comparisons of pitch were not statistically significant for age or education only for speech style. The results of these tests can be seen in Appendix A.

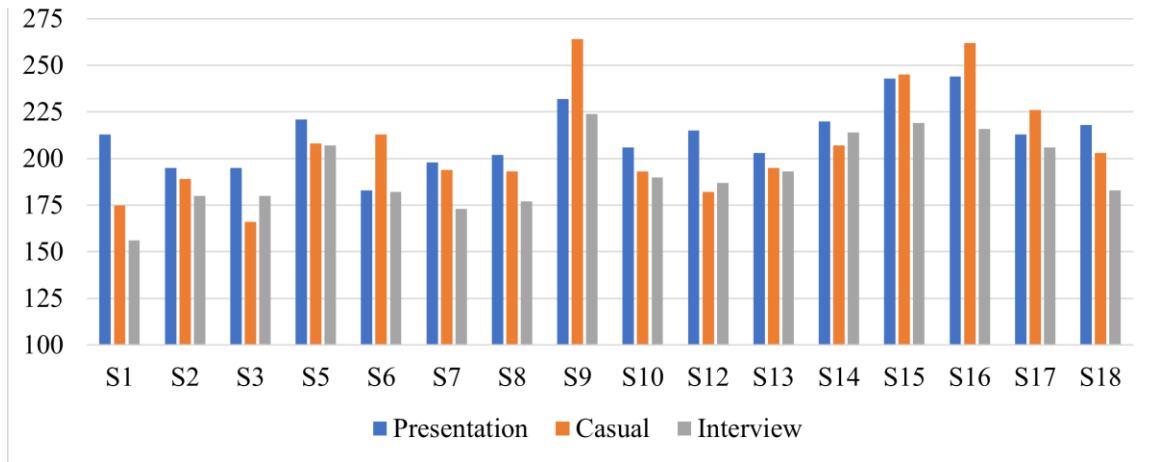
### 3.2.2 Speaker variations

Though the overall results were consistent with the first hypothesis, not all speakers followed the pattern of  $F_0 = \text{presentation} > \text{casual} > \text{interview}$ . The averages for each speaker in each speech style are shown in Table 12 and Figure 9 below.

**Table 12. Average pitch (in Hz) for each speaker by style**

	<b>S1</b>	<b>S2</b>	<b>S3</b>	<b>S5</b>	<b>S6</b>	<b>S7</b>	<b>S8</b>	<b>S9</b>
<b>Presentation</b>	213	195	195	221	183	198	202	232
<b>Casual</b>	175	189	166	208	213	194	193	264
<b>Interview</b>	156	180	180	207	182	173	177	224
	<b>S10</b>	<b>S12</b>	<b>S13</b>	<b>S14</b>	<b>S15</b>	<b>S16</b>	<b>S17</b>	<b>S18</b>
<b>Presentation</b>	206	215	203	220	243	244	213	218
<b>Casual</b>	193	182	195	207	245	262	226	203
<b>Interview</b>	190	187	193	214	219	216	206	183

**Figure 9. Average pitch (in Hz) for each speaker by style**



Speakers 3, 12, and 14 had lower average pitch in the casual speech sample than in the interview speech sample. Speakers 12 and 14 had very similar average pitch for those two styles; the difference between them was less than 10 Hz. Speaker 3, however, had much lower pitch in the casual speech than in the interview speech. It is possible that her use of fry was different in these two styles and affected her average pitch. While her percentage of phrases with fry was similar between the two styles, 85% for casual and 82% for interview, the type of fry was different. The number of phrases with each type of fry is shown in Table 13 below.

**Table 13. Amount of fry for Speaker 3 in interview and casual speech**

	0	0.5	1	1.5	2	2.5	3	3.5	4	Total
<b>Casual</b>	4	1	9	1	4	0	6	0	3	28
<b>Interview</b>	6	4	6	3	7	1	6	0	2	35

The interview style had more occurrences of fry following or replacing a glottal stop, as shown in Table 14. Types of fry for Speaker 3 in interview and casual speech below.

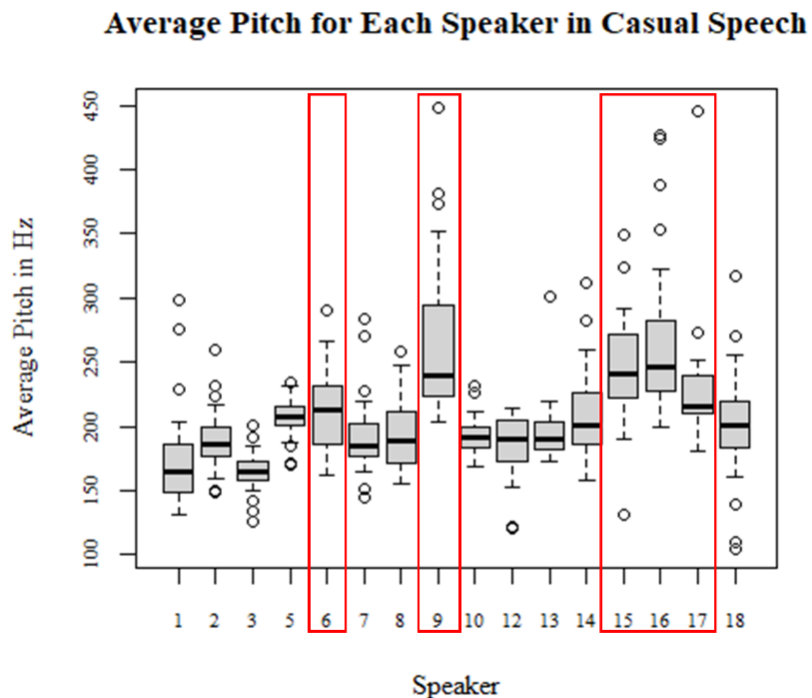
**Table 14. Types of fry for Speaker 3 in interview and casual speech**

	<b>Fry following glottal stop</b>	<b>Other fry</b>
<b>Casual</b>	2	22
<b>Interview</b>	8	21

These occurrences of fry were usually brief and had less effect on average pitch than the standard use of fry, which could explain the much lower pitch in the casual speech for Speaker 3.

Speakers 6, 9, 15, 16, and 17 had higher pitch in casual speech than in presentation speech; there are several possible reasons for this departure from the expected. The range of pitch in casual speech may be raising the average pitch for the whole speech sample. All five of these speakers had a wide range of the average pitch of phrases in casual speech, as shown in the figure below.

Figure 10. Average pitch (in Hz) for each speaker in casual speech



The outliers were more above than below the average pitch for these five speakers. Speakers 9, 15, and 16, in particular, had some phrases with extremely high pitch.

Fry may also contribute to the difference in average pitch for these speakers. For all but Speaker 17, these speakers had much higher use of fry in presentation speech than in casual speech, which may explain why the pitch in casual speech was higher than presentation speech. The percentage of phrases with fry in each speech style for these speakers is shown in Table 15 below.

**Table 15. Percentage of fry for Speakers 9, 15, 16, and 17**

<b>Speech Style</b>	<b>Speaker 9</b>	<b>Speaker 15</b>	<b>Speaker 16</b>	<b>Speaker 17</b>
<b>Presentation</b>	100.0%	72.7%	83.3%	100.0%
<b>Casual</b>	44.7%	51.1%	52.2%	93.8%
<b>Interview</b>	46.8%	37.7%	56.9%	97.7%

Speaker 17 used more fry in the presentation speech than in the casual speech, but the difference was not very large. It is more likely that Speaker 17 lowered her pitch in the presentation to sound more authoritative though the general trend was for higher pitch in the presentation in order to project.

Given the general trends and the deviations from them by individual speakers, it is most likely that a young woman will lower her pitch in interview style speech. It is somewhat likely that she will raise her pitch in a presentation.

### 3.3 Uptalk results

The percentages of phrases with uptalk for each speech style are shown in Table 16 below.

**Table 16. Percentages of phrases with uptalk by speech style**

<b>Speech style</b>	<b>Total phrases</b>	<b>Phrases with uptalk</b>	<b>Percentage of phrases with uptalk</b>
<b>Presentation</b>	180	6	3%
<b>Casual</b>	638	39	6%
<b>Interview</b>	874	125	14%



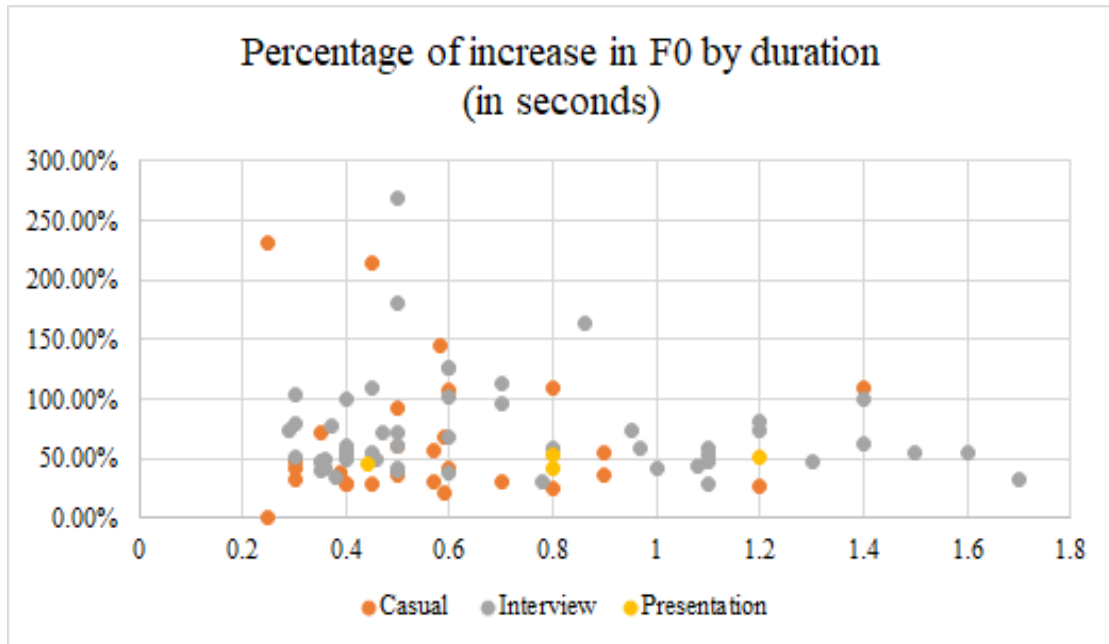
The interview style has the most phrases with uptalk, and the presentation style has the fewest phrases with uptalk though it is not much different than the casual style. The only examples of uptalk in the presentation style were produced by Speaker 6 and is discussed in Section 3.3.5.

### *3.3.1 Uptalk description*

There is no accepted acoustic definition of uptalk. Guy and colleagues described typical uptalk as a rise of “at least 40 percent, beginning on the last tonic syllable of the tone group and continuing sharply upwards through any subsequent syllables.” (1986) Warren rejects this as a definition because he believes uptalk is defined primarily by its function and not its acoustic properties.

I analyzed 85 of the phrases with uptalk from the speakers who used uptalk in both casual and interview speech. Pitch tended to drop very low right before the uptalk rise. Sometimes it dropped so low that the speaker used fry on the syllable immediately preceding the uptalk rise. I excluded the phrases with fry before the uptalk as pitch is difficult to measure accurately during syllables with vocal fry. I measured the duration of the rise in seconds and the minimum and maximum pitch of the rise in Hz. The results are shown in Figure 11 below.

Figure 11. Change in pitch during uptalk rise vs duration of rise



Generally, the steepest rises have shorter durations, but there are exceptions to this trend. As shown in Table 17 below, most of the rises overall had at least a 40% increase in pitch, meeting Guy et al.'s criteria (1986). Casual speech has the most uptalk rises that increased less than 40% in pitch.

**Table 17. Increase in F0 during uptalk rises**

Speech style	Total	Rises under 40%		Rises over 40%	
		Count	Percentage	Count	Percentage
<b>Casual</b>	29	12	41%	17	58%
<b>Interview</b>	52	6	11%	46	88%
<b>Presentation</b>	4	0	--	4	100%
<b>Overall</b>	85	18	21%	67	78%

All uptalk rises in the presentation speech were from Speaker 6. Four of them were included in this analysis. All of them had increases in pitch greater than 40% and less than 100%.

Casual and interview speech had a similar ratio of rises that more than doubled in pitch as shown in Table 18 below.

**Table 18. Uptalk rises that increased in F0 by at least 100%**

Speech style	Count	Percentage
<b>Casual</b>	6	20%
<b>Interview</b>	9	17%
<b>Total</b>	15	17%

The length of the rises varied greatly. Many rises were short, less than one second in length, and often over a single syllable. Some rises, mostly in the interview style, began early in the phrase and continued through the end. The duration of rises was slightly longer on average in the formal speech styles, as shown in

Table 19 below.

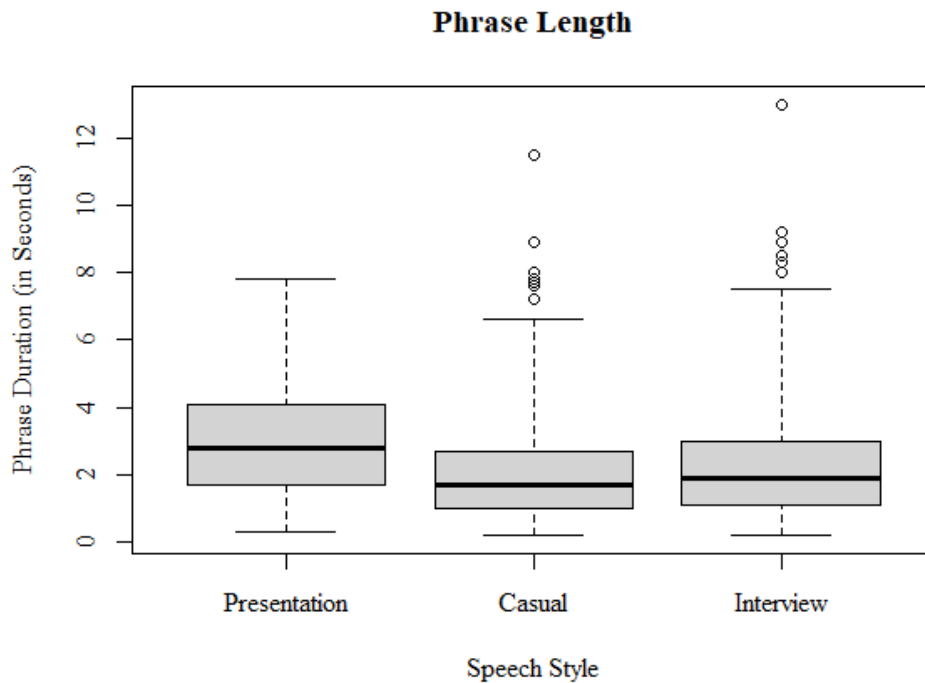
**Table 19. Average duration of uptalk rise**

Casual	Interview	Presentation
0.57 seconds	0.69 seconds	0.81 seconds

Only two uptalk rises in the casual style were longer than one second.

I measured phrase-length for 1365 phrases in the data. The averages and spread are shown in Figure 12 below.

**Figure 12. Length of phrases across speech styles**



While the average duration of each phrase was fairly similar for interview and casual speech, the formal speech styles had longer phrases overall which may be why the uptalk rises were also longer in formal speech.

### 3.3.2 Statistical results for uptalk

Since uptalk was coded as a binary variable, **UptalkDummyDep**, it was necessary to use a binomial (logit) generalized linear mixed model (Gries 2017). The same independent variables were used for the uptalk test as in the test for pitch: **Style, Age, and Education**. All 1692 phrases were included in this test.

*glmer(UptalkDummyDep ~ Style + Age + Education + (1 | Speaker))*

Again, Style was significant, and Age and Education were not, as shown in Table 20 below. Significance is marked with an \*.

**Table 20. Statistical results for uptalk**

	Estimate	Std. Error	Z value	Pr(>  z )
<b>Style</b>	0.8631	0.1546	5.582	2.37e-08 ***
<b>Age</b>	-0.4490	0.3137	-1.431	0.152
<b>Education</b>	0.1822	0.2147	0.849	0.396

The table of estimates with 95% confidence intervals is shown in Appendix B. If the interval includes 0, the null value, the null hypothesis cannot be rejected. Intervals that do not contain the null value are indicated by \*.

The statistical models confirmed the observation in Table 16. There was significantly *less* uptalk in presentation and casual style speech and significantly *more* uptalk in interview style speech.

### *3.3.3 Interview*

I expected to find the most instances of uptalk in the interview speech style for several reasons. First, the interview was the most interactive of the three speech samples. The presentation was read from a script that involved no verbal interaction from either the other speaker in the Zoom call or me. To elicit the casual speech, speakers were prompted to tell a story. The other speaker, the listener, generally interacted a little with the speaker, but by far the most back and forth interaction was in the interview. Uptalk is often used in interactive speech to check-in with the listener to see how they view the information being shared, and, consequently, the speaker sharing the information (Guy et al. 1986; Warren 2016).

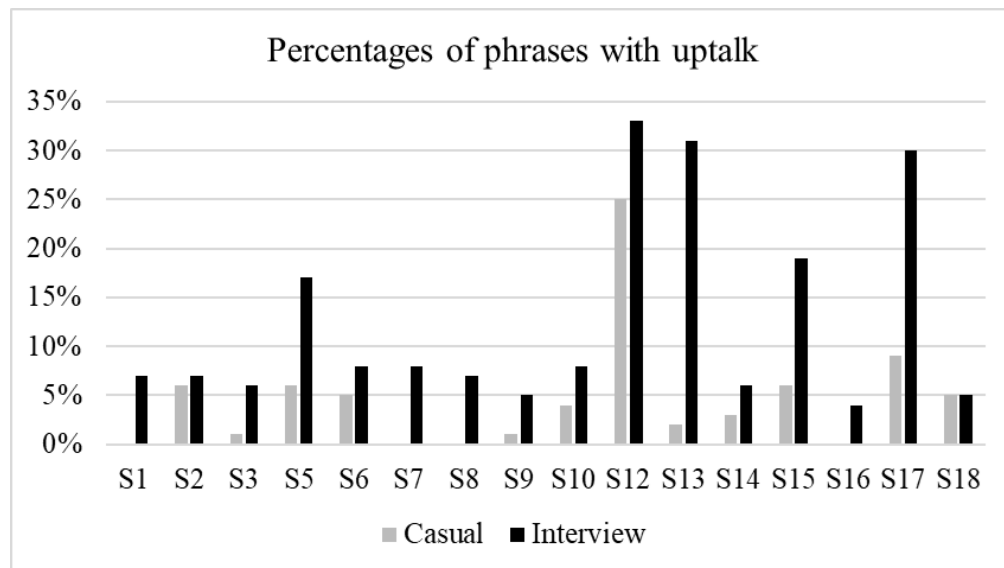
Second, the interviewer had a bigger status difference from the participant than either me, a peer with possibly no pre-established relationship, or the other speaker who was a peer with a previously established relationship. The interviewer was male, older than the speakers by at least 30 years, had no prior relationship with the speakers, and was introduced to the speakers with prestigious titles and qualifications. These status differences could prompt the speakers to use more politeness and deference strategies in their interactions with the interviewer, including but not limited to uptalk.

A third reason I expected more uptalk in the interview style is because of turn-taking and floor-holding strategies. Uptalk, while attempting to prompt a response from the listener, does not end the speaker's turn. The rise is similar to a continuation rise and can signal to the listener that the speaker is not done talking, functioning as a floor-holding tactic (Warren 2016: 56). Turn-taking is more difficult through videoconferencing because of possible internet delays and the lack of nonverbal cues (See Section 1.6). Uptalk may be one way the speakers compensate for the lack of clarity and maintain control of their turn while in a situation with an imbalance of status and power.

Finally, interviews, even ones without a stated position or goal, are always evaluative. As speakers share about themselves, they naturally want to know what the interviewer is thinking of them. Uptalk almost forces some sort of response from the listener whether a nod or verbal assent, and helps the speaker feel that the listener approves of what they are saying or at least is paying attention to them.

Overall, there were more instances of uptalk in the interview samples than the other two speech styles. There was also a wide variety of how much uptalk an individual speaker used overall and in the interview samples. Figure 13 below shows the percentages of uptalk for both casual and interview speech.

**Figure 13. Percentages of phrases with uptalk**



Twelve of the sixteen speakers used uptalk in the casual style. Three speakers did not follow the expected pattern of higher uptalk use in the interview style. Speaker 18 used the same amount of uptalk in both the casual and interview styles, and Speakers 3 and 10 used more uptalk in the casual style than in the interview. Speaker 6 had

unexpected uptalk use in the presentation style as mentioned before and will be set aside and discussed in Section 3.3.5.

Importantly, all speakers used uptalk in the interview style even if they had not used it at all previously in the call. This finding suggests that uptalk is more often used or even when the interlocutor has more status or power than the speaker.

### 3.3.4 Casual

I expected to find some amount of uptalk in the casual speech, less than in the interviews but more than in the presentation samples. This expectation was true for all the speakers except 3, 10, and 18. Speakers 3 and 18 had similar uptalk use in interview and casual speech. Speaker 10 used uptalk on 24% of the phrases in her casual speech sample compared with only 12% in the interview sample, but the actual number of phrases with uptalk was not largely different, as shown in Table 21 below.

**Table 21. Use of uptalk by Speaker 10**

	<b># of phrases</b>	<b># of phrases with uptalk</b>	<b>Percentage of phrases with uptalk</b>
<b>Casual</b>	33	8	24%
<b>Presentation</b>	51	6	12%

It is interesting to note in the casual speech samples not how often uptalk occurred but where it occurred. Uptalk was most frequently used in setting the scene of the story. The speakers seemed to be checking to make sure the listener was understanding the context and characters of the story and would be able to follow the plot as it developed. In these instances, uptalk is functioning as a discourse marker signaling background



information. This corresponds with the findings in Guy et al (1986: 43) on the higher frequency of Australian questioning intonation use in descriptions and narratives in order to check the listener's comprehension of crucial information.

The second most common use of uptalk was introducing surprising or unexpected information. For example, in describing a fun event with her family Speaker 9 used uptalk while talking about a rare treat in example (5) below.

(5) "We were even being allowed to have Pepsi?"

In casual speech, uptalk was also used to signal uncertainty about the information conveyed. Usually, this use of uptalk was accompanied by qualifying phrases such as "I'm not sure" or "I think".

### *3.3.5 Presentation*

Because uptalk is fundamentally interactive, it is expected that the presentation style speech would have the fewest occurrences in uptalk. For this speech sample, speakers took turns reading the slide in Figure 14 below shared on my screen in Zoom.

### Figure 14. Text for presentation speech

Read this paragraph as if you were presenting to a class or at your work

Based on the results that we have collected in our study, I believe that we will be seeing a vast improvement in our program within the next year. There may in fact be improvements as early as next week. The data has informed our innovative, new program design which combines the strengths of the past program with the changes that need to be made as we move into the future. I know that some have disagreed with these changes in the past, but I think that the evidence clearly shows that my approach is effective and practical. These innovative techniques, while controversial, are backed by proven results and can improve the efficiency and quality of our organization. Are there any questions?

The text was also provided beforehand in the email confirming the time for the Zoom meeting. The speakers were prompted to read the slide as if they were giving a presentation to a class or at work. Only one participant used any uptalk in the presentation sample. Speaker 6 used uptalk on 6 out of 8 phrases in her sample. Though her uptalk use was incredibly high in the presentation speech, her uptalk use in the other speech styles was relatively low compared with the other speakers, Table 22 below shows the percentage of phrases that had uptalk for Speaker 6 in each speech style and the percentage of phrases with uptalk for the other fifteen speakers in the study.

**Table 22. Uptalk use for Speaker 6 compared with other speakers**

Speech style	Speaker 6			All other speakers		
	Total phrases	Phrases with uptalk	Percentage	Total phrases	Phrases with uptalk	Percentage
<b>Presentation</b>	8	6	75%	172	0	0%
<b>Casual</b>	39	2	5%	599	37	6%
<b>Interview</b>	51	4	8%	823	121	14%

Speaker 6's use of uptalk in the presentation speech is highly irregular and does not pattern at all with the other 15 speakers. This difference suggests that her use of uptalk is idiosyncratic and not typical of young women as a group. In her interview, Speaker 6 said that she did not like confrontation or talking about her own skills. The presentation text was deliberately written as if the speaker were proposing ideas that the group may not approve. It is possible that the speaker was using uptalk to mitigate the forcefulness of her own (simulated) ideas. Deference and qualifying content, such as communicating uncertainty about the accuracy of the statement, are both motivations for uptalk use noted in Section 1.3.7.

### **3.4 Fry results**

Fry was coded multiple ways in the data, as mentioned in Section 2.1.5, in order to look at the presence, amount, and location of fry in the phrase.

Overall, the percentage of phrases with fry was highest in the presentation speech. This result is surprising because overall pitch was highest in the presentation pitch, and one of the characteristics of fry is low F0. A possible explanation of this finding is that

the pitch of fry is difficult to measure, so the fry did not affect the pitch measurements though it may affect the listener's perception of the pitch (Hollien & Wendahl 1968). The rankings of the percentages of phrases with fry by speech style for each speaker are shown in Table 23 below.

**Table 23. Rankings of percentages of fry by speech style**

<b>Speaker</b>	<b>Percentages of Fry</b>
<b>S1</b>	Casual >> Interview >> Presentation
<b>S2</b>	Presentation >> Interview >> Casual
<b>S3</b>	Casual >> Interview >> Presentation
<b>S5</b>	Presentation >> Casual >> Interview
<b>S6</b>	Presentation >> Interview >> Casual
<b>S7</b>	Presentation >> Interview >> Casual
<b>S8</b>	Presentation >> Interview >> Casual
<b>S9</b>	Presentation >> Interview >> Casual
<b>S10</b>	Interview >> Casual >> Presentation
<b>S12</b>	Presentation >> Interview >> Casual
<b>S13</b>	Presentation >> Interview >> Casual
<b>S14</b>	Presentation >> Interview >> Casual
<b>S15</b>	Presentation >> Casual >> Interview
<b>S16</b>	Presentation >> Interview >> Casual
<b>S17</b>	Presentation >> Interview >> Casual
<b>S18</b>	Interview >> Casual >> Presentation

Speakers 3, 10, and 18 are again exceptions to the general pattern of the other speakers. Both Speakers 10 and 18 had the highest percentages of fry in the interview

speech. Speaker 3 had very similar percentages of fry for interview and casual speech though the casual speech was slightly higher as noted in Section 3.2.2. Speaker 1 was also an exception to the pattern. She had very similar percentages of fry across the three speech styles, but interview and casual speech had slightly higher percentages of fry than presentation.

Speaker 9 used fry on every phrase in the presentation but on less than half of the phrases in the interview and casual speech. Speaker 2 used fry on half the phrases in the presentation, but only 26% of phrases in the casual speech and 35% of phrases in the interview speech. These findings seem to indicate that even if a person would not normally use a large amount of fry, they would often use it in a presentation.

### 3.4.1 Statistical tests for fry

Like the tests for uptalk, fry was coded as a dichotomous variable **FryDummyDep**, so a binomial generalized mixed effects model was used.

*glmer(FryDummyDep ~ Style + Age + Education + (1 | Speaker)*

The results were significant for Style as shown in Table 24 below.

**Table 24. Statistical results for fry**

	Estimate	Std. Error	z value	Pr(>  z )
<b>StylePresentation</b>	0.8161	0.19537	4.177	2.95e-05 ***
<b>StyleCasual</b>	-0.9093	0.2012	-4.52	6.19e-06 ***
<b>StyleInterview</b>	-0.8161	0.1954	-4.177	2.95e-05 ***

**StylePresentation** has a positive estimate (also called a coefficient) so there was significantly more fry in presentation speech while casual and interview speech had significantly less fry. The negative estimate for the interview style does not support my original hypothesis that there would be an increase in fry in the interview speech because it has lower pitch. Education level also seemed affect fry significantly if the speakers had some postgraduate study, but as there were only three speakers in this category, the sample is not large enough to represent the entire population. The tables with estimates and 95% confidence intervals are in Appendix C.

The amount of fry for each phrase, **FryScalar**, was not statistically significant in the different speech styles, age, or education categories, but speakers did differ from each other in how much fry they used. This difference of use among speakers seems to indicate that amount of fry is an idiosyncratic characteristic within the group of young female American English speakers or that the study size was too small for the patterns to be significant.

A total of 990 phrases had fry. The location of fry in the phrase was divided into three groups, phrase initial, phrase medial, and phrase final.

**Table 25. Location of fry in the phrase**

<b>Phrase-initial fry</b>	<b>Phrase-medial fry</b>	<b>Phrase-final fry</b>	<b>Total</b>
181	594	525	990

Presentation speech had significantly less phrase-initial fry than interview or casual speech as shown in Table 26 below.

**Table 26. Statistical results for phrase-initial fry and speech style**

	Estimate	Std. Error	z value	Pr(>  z )
<b>StylePresentation</b>	-0.77862	0.33421	-2.33	0.01982 *
<b>StyleInterview</b>	0.08397	0.17137	0.49	0.6241
<b>StyleCasual</b>	-0.08397	0.17137	-0.49	0.62415

Education had some interaction with phrase-initial fry. As shown in Table 27, there was significantly more phrase-initial fry for all education levels except for **EducationCompleted postgraduate degree**. As there were only a few speakers in each category, it is impossible to draw conclusions from these results. The effect of education on fry is an area for further research.

**Table 27. Statistical results for phrase-initial fry and education**

	Estimate	Std. Error	z value	Pr(>  z )
<b>Some postgraduate study</b>	2.82419	0.96663	2.922	0.00348 **
<b>Some college or university</b>	2.7925	0.97544	2.863	0.00420 **
<b>Completed college or university</b>	2.07635	1.01022	2.055	0.03985 *
<b>Completed postgraduate degree</b>	-2.07635	1.01023	-2.055	0.0398 *

Casual speech had significantly less phrase-medial fry, as shown in Table 28 below.

**Table 28. Statistical results of phrase-medial fry and style**

	Estimate	Std. Error	z value	Pr(>  z )
<b>StylePresentation</b>	-0.08525	0.17515	-0.487	0.62647
<b>StyleCasual</b>	-0.42208	0.11619	-3.633	0.00028 ***
<b>StyleInterview</b>	0.08525	0.17515	0.487	0.62646

Presentation speech had significantly more phrase-final fry, as shown in Table 29 below.

**Table 29. Statistical results of phrase-final fry and style**

	Estimate	Std. Error	z value	Pr(>  z )
<b>StylePresentation</b>	1.23606	0.17415	7.098	1.27e-12 ***
<b>StyleCasual</b>	-0.51374	0.30905	-1.662	0.0964
<b>StyleInterview</b>	-0.1526	0.1203	-1.268	0.2049

The larger amount of phrase-final fry may be due to the difference between free speech and read speech. In the presentation sample, participants were reading a text designed to have forceful, declarative sentences. This may have prompted them to read each sentence with more finality — falling intonation, lowered pitch, and constriction of the glottis — than they would use in free speech resulting in more phrase-final fry. Further research is needed to determine if the style-based distinctions in fry location are true in presentations that are not read directly from a script.



The tables with estimates and 95% confidence intervals for all the tests on fry are in Appendix C.

### 3.5 Perception test results

A total of 45 participants completed the survey. Table 30 below shows the total number of participants by gender.

**Table 30. Gender of survey participants**

<b>Gender</b>	<b>Count</b>	<b>Percentage</b>
<b>Male</b>	13	29%
<b>Female</b>	32	71%
<b>Total</b>	45	100%

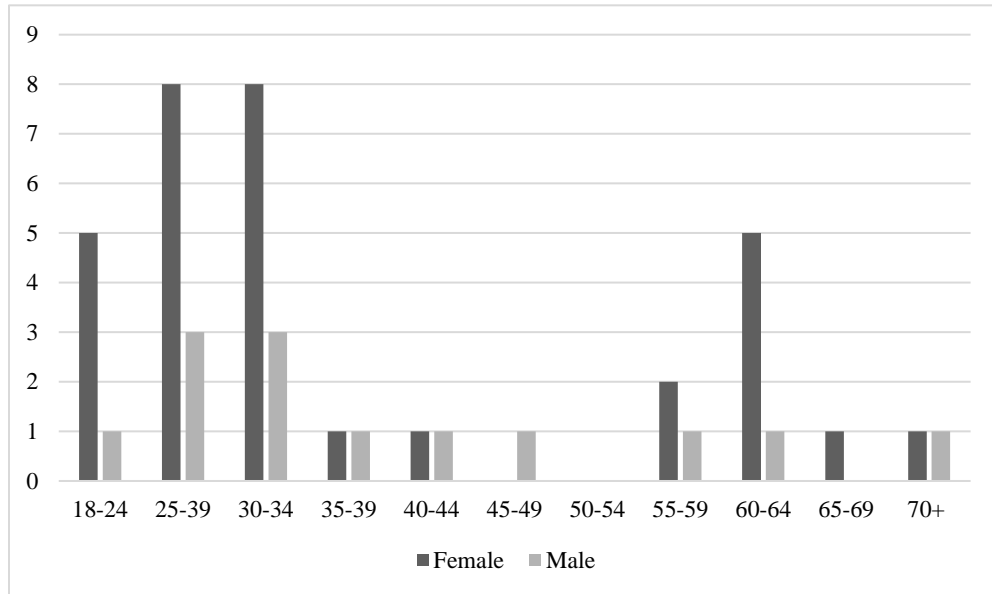
Table 31 below shows the distribution of age of the participants.

**Table 31. Age of survey participants**

<b>Age</b>	<b>Count</b>	<b>Percentage</b>
18-24	6	13%
25-29	11	24%
30-34	11	24%
35-39	2	4%
40-44	2	4%
45-49	1	2%
50-54	0	0%
55-59	3	7%
60-64	6	13%
65-69	1	2%
70+	2	4%
<b>Total</b>	<b>45</b>	<b>100%</b>

Figure 15 below shows the age of the participants grouped by gender.

**Figure 15. Age of participants grouped by gender**



Almost half of the participants in the perception survey are from the same demographic group as the speakers - women under the age of 35. This subgroup of 21 participants will be considered separately from the total in Section 3.5.6.

When asked whether men or women were better leaders in their experience, most people chose both. Nine participants, four female and five male, selected that men were better leaders. Only one person, a woman aged 25-29, said that women were better leaders. Since more women responded to the survey than men did, it is easier to see the distribution of results with percentages. The leadership preferences of the participants are in Table 32 below.

**Table 32. Gender preferences in leadership by participant gender**

Gender of participant	Prefer not to answer	Both	Men	Women
Male	0%	62%	38%	0%
Female	3%	81%	13%	3%
Total	2%	76%	20%	2%

Participants responded to the question, “In your experience, who are better leaders?” As shown, a higher percentage of male participants think that men are better leaders. Interestingly, people, whether male or female, who chose that men were better leaders had responses that were similar to the responses from female participants.

### *3.5.1 General trends*

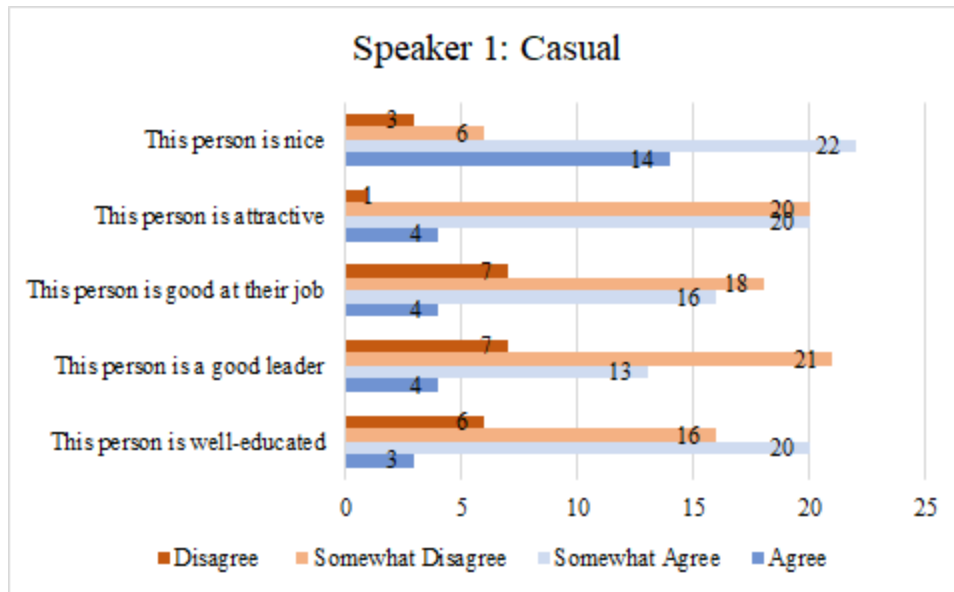
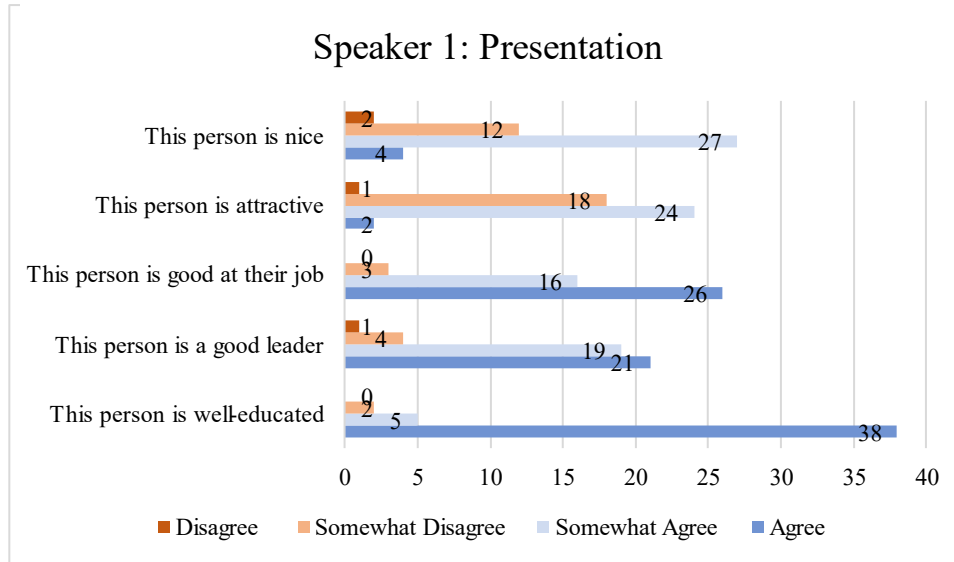
All three speakers were rated highest for well-educated in the presentation speech.

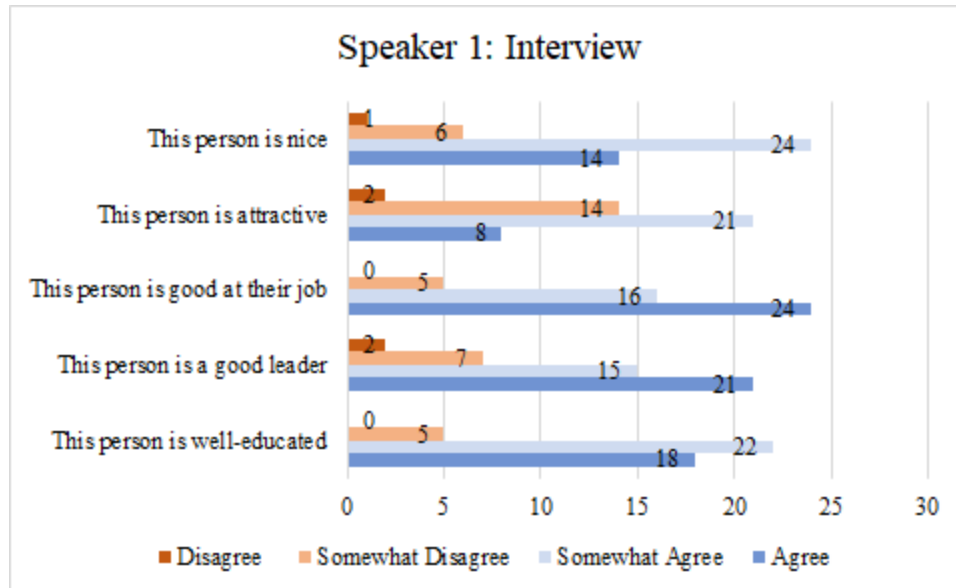
All three speakers were rated lowest for competency in their casual speech samples based on the most “Disagree” responses given for “This person is good at their job”, “This person is a good leader”, and “This person is well-educated”.

### *3.5.2 Perceptions of pitch: Speaker 1*

Speaker 1 was chosen to measure the effects of young women’s speaking pitch on the perception of the listener. She was selected because she had large differences in average pitch across the three speech styles. The results from the perception survey for Speaker 1, divided by speech style, are shown in the figure below.

**Figure 16. Survey results for Speaker 1**





Since her pitch in the presentation was the highest, it was expected that Speaker 1 would be rated as more attractive and nice based on the presentation speech because high pitch is associated with higher ratings of attractiveness (See Section 1.2.3). However, most participants rated Speaker 1 as more attractive in the interview speech, which had the lowest average pitch.

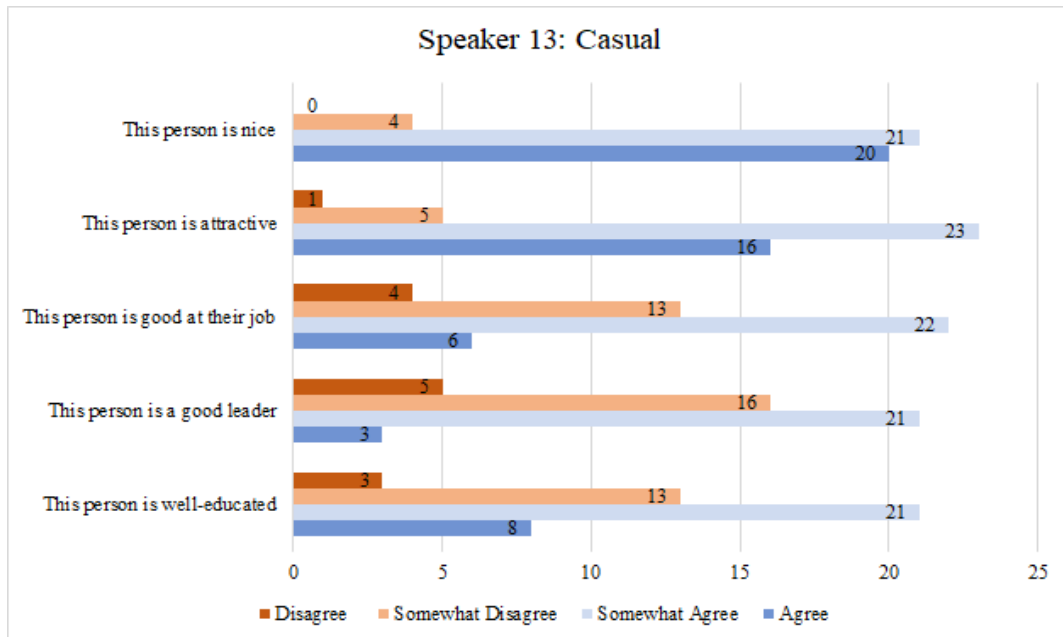
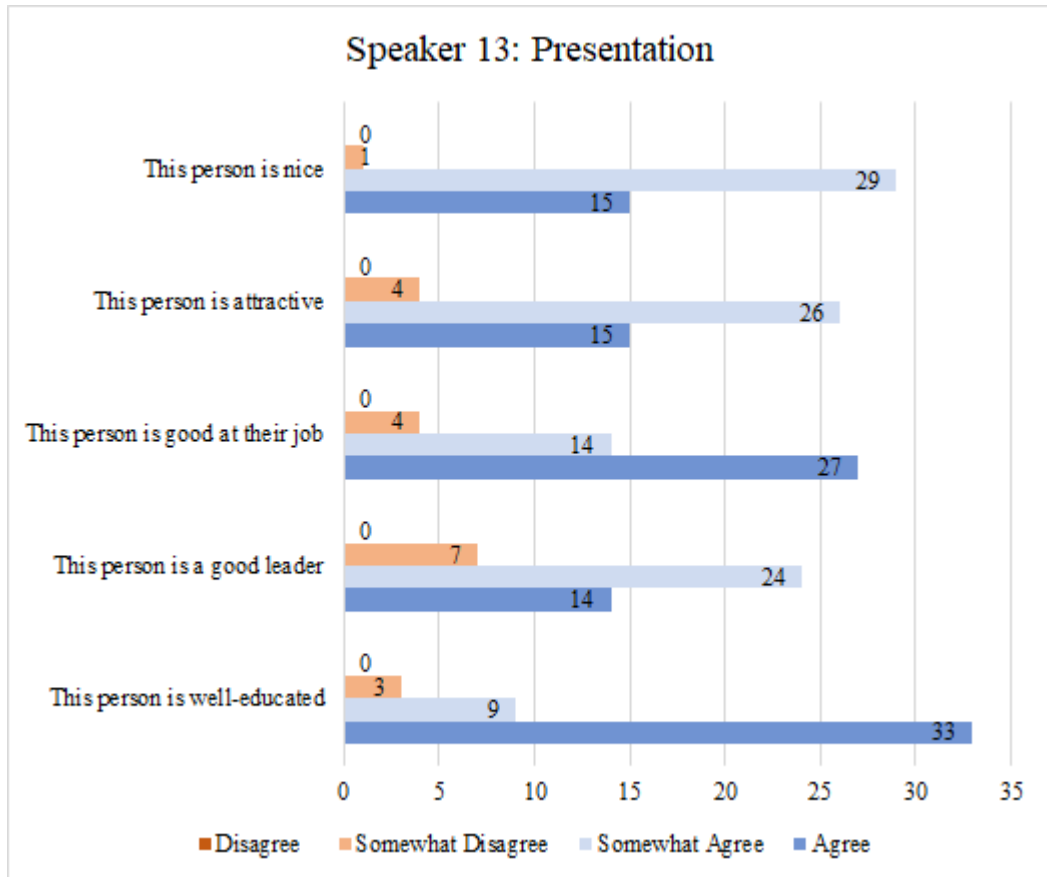
Most participants, 85%, agreed that Speaker 1 sounded well-educated in the presentation speech. This result was the strongest agreement for any speaker for any statement across all speech styles. In the casual speech, only three participants agreed that Speaker 1 was well-educated; 22 participants disagreed or somewhat disagreed. Most disagreed or somewhat disagreed that Speaker 1 was good at their job or a good leader. Attraction was split almost evenly, with three more participants agreeing than disagreeing. Speaker 1 received much lower agreement for competency statements for casual speech. It does not seem likely that pitch is affecting these changes since casual speech in the middle of Speaker 1's pitch range. Though Speaker 1 was chosen to

represent pitch, Speaker 13 may illustrate the link between high pitch and attractiveness (Collins & Missing 2003). Speaker 13 had the highest overall pitch and only two responses of “Disagree” to the statement “This person is attractive” across all speech styles. Those two responses were from the same participant, female age 30-34, and she answered “Disagree” to “This person is attractive” after every recording except Speaker 13’s presentation speech sample after which she selected “Somewhat Disagree”.

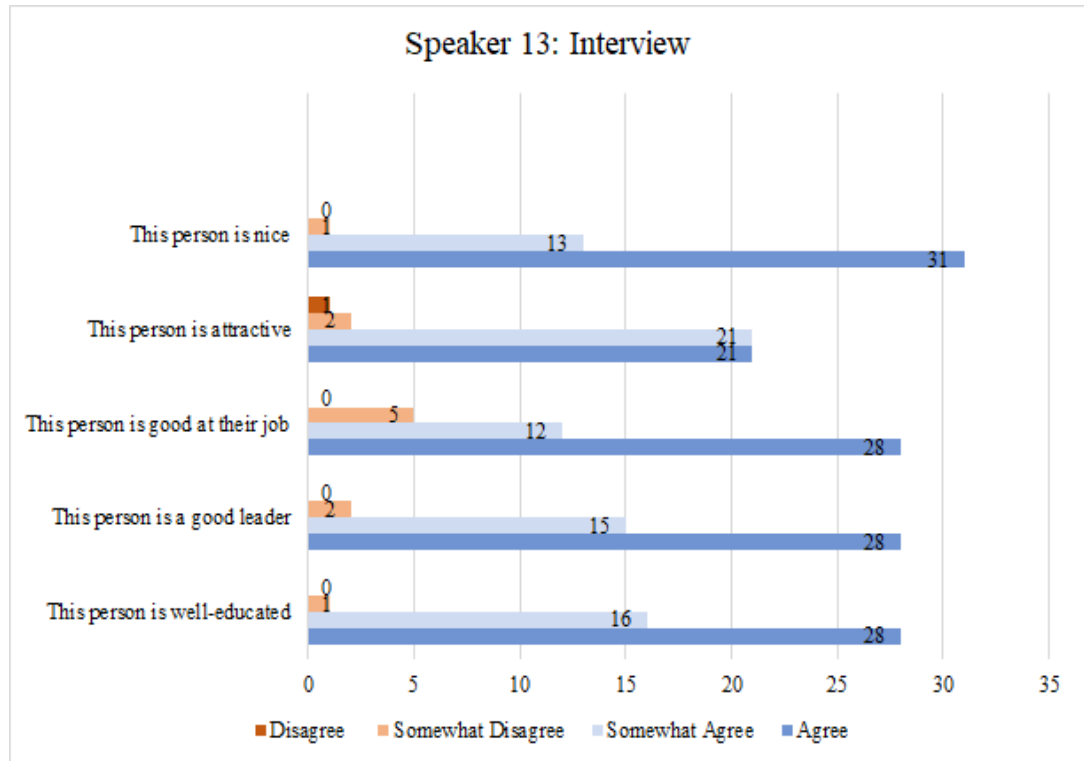
### *3.5.3 Perceptions of uptalk: Speaker 13*

Speaker 13 had very similar pitch in all her speech samples. There was around a 10 Hz difference between her interview and presentation speech. Semitones are a better estimate than Hertz of the differences in perception between frequencies. The difference between Speaker 13’s interview and presentation speech pitch was little more than one semitone. Thus, pitch should not be a significant factor affecting the listener’s perception of her across different speech styles. Speaker 13 had a high use of uptalk in the interview speech, and none were included in her presentation or casual speech samples. The results from the perception survey for Speaker 13, divided by speech style, are shown in Figure 17 below.

Figure 17. Survey results for Speaker 13





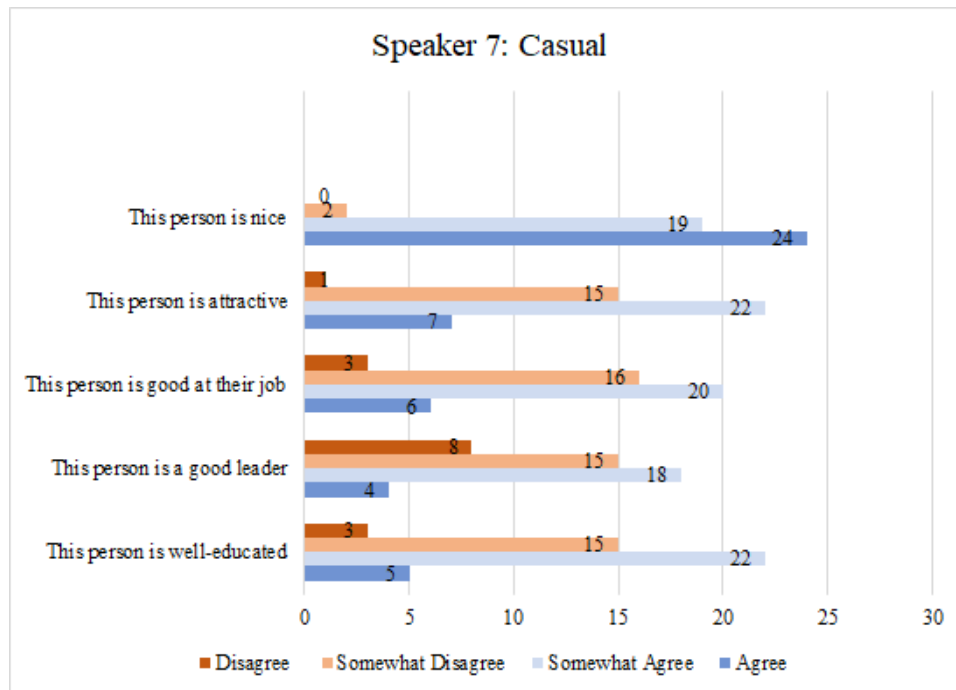
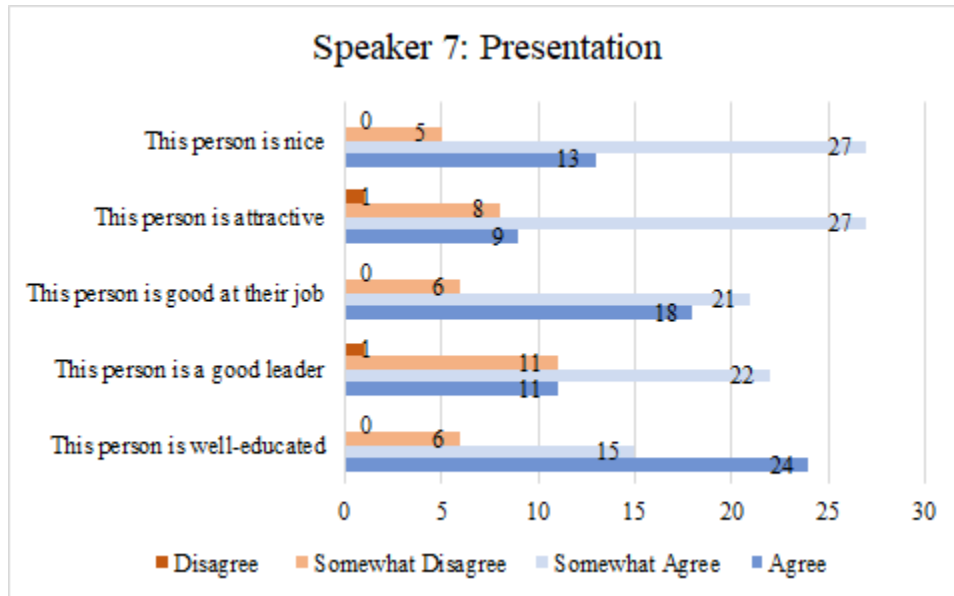


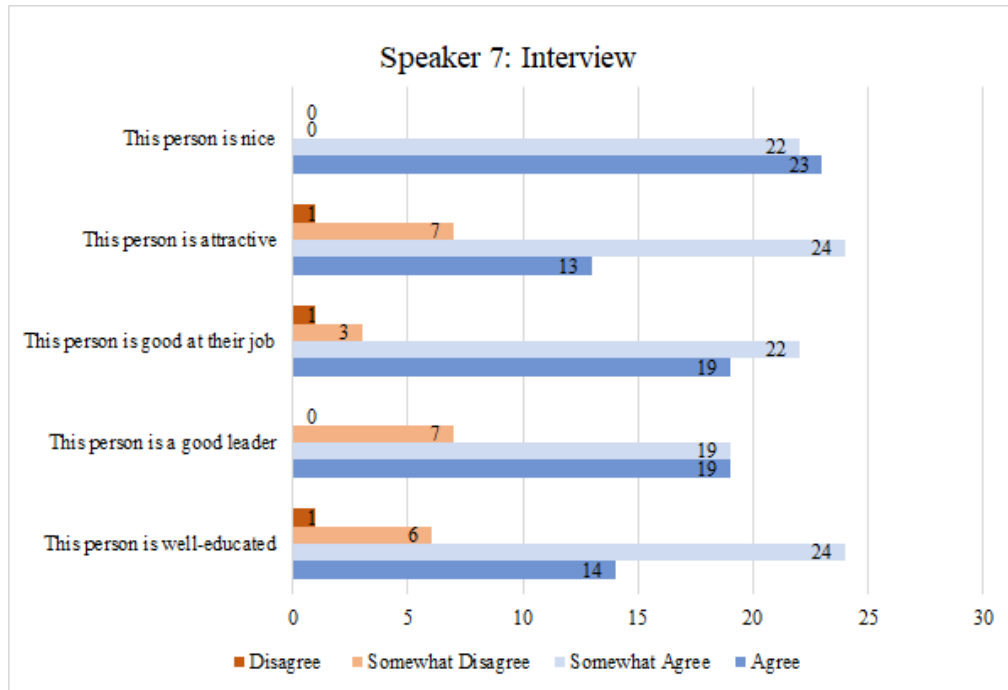
Speaker 13 received much lower ratings for competency for casual speech than interview and presentation. Since casual speech had only one phrase with uptalk, it does not seem likely that uptalk changes perceptions of competency by itself. Attraction ratings were similar across speech styles but slightly lower in casual speech.

### 3.5.4 Perceptions of fry: Speaker 7

Speaker 7 was chosen for the fry perception test. In the presentation speech, 60% of the phrases had fry; in the interview, 45% of the phrases had fry; and in the casual speech sample, only 35% of the phrases had fry. She had two phrases with uptalk in the interview sample used for the perception survey. The results from the perception survey for Speaker 7, divided by speech style, are shown in the figure below.

Figure 18. Survey results for Speaker 7





Overall, Speaker 7 was rated most highly after the interview speech, which had a 45% fry. Fry did not seem to affect perceptions of competency or attractiveness; however, there were slightly more “Agree” and “Somewhat Agree” responses to the statement “This person is nice” after interview and casual speech styles which had lower percentages of fry.

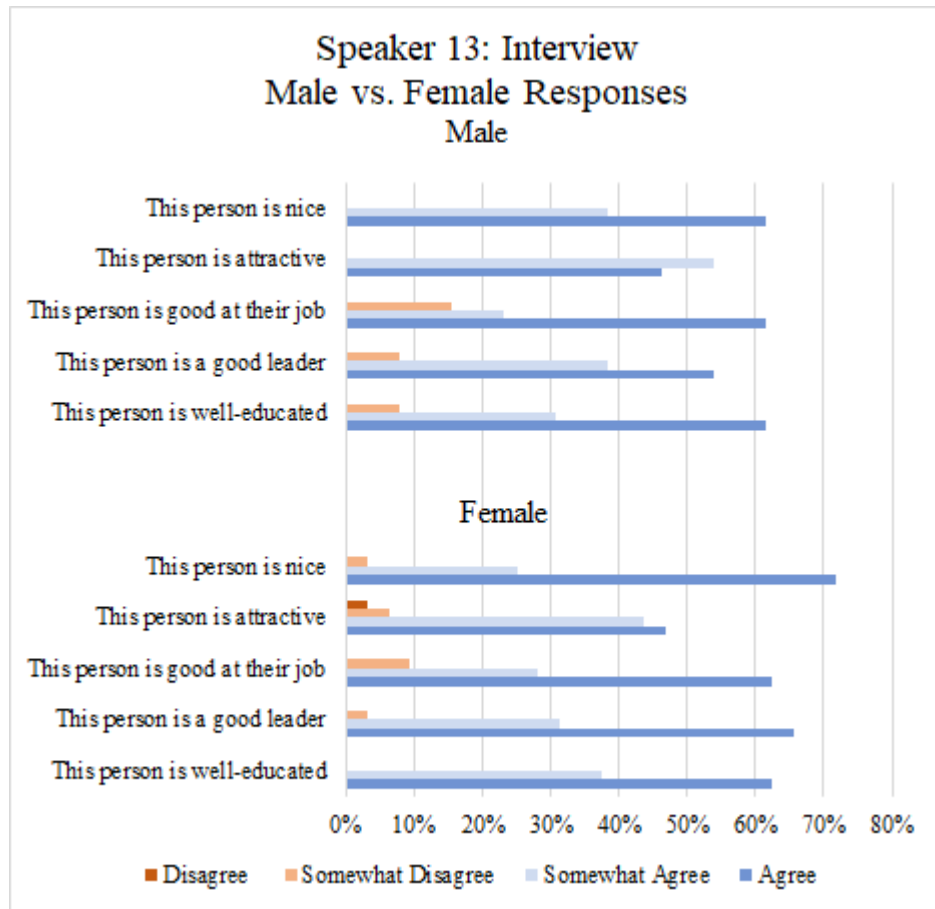
### 3.5.5 Gender

There were differences in responses to the survey based on the gender of the participants. Three of the statements in the survey were related to competency and authority and two were related to attractiveness. Men tended to choose “Disagree” or “Somewhat Disagree” more often than women for statements of competency. However, women disagreed more often than men for statements of attractiveness and niceness.

Male participants chose “Agree” less overall than women did, especially for the statement “This person is good at their job.”

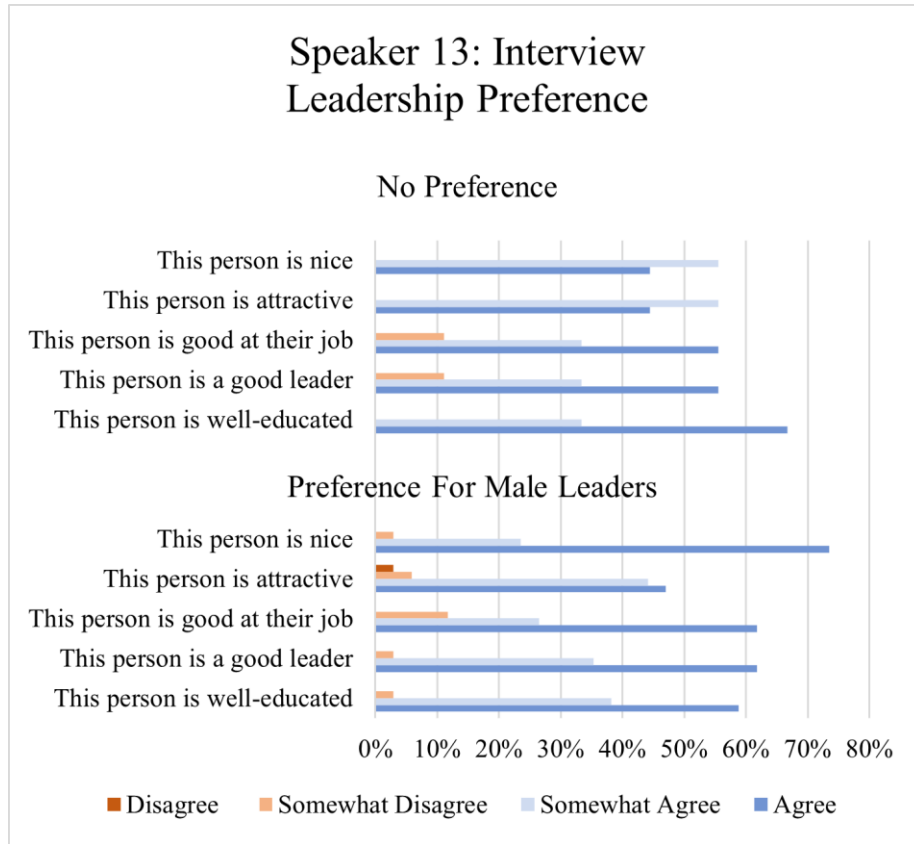
Uptalk seems to affect the perception of competency more for men than for women. Speaker 13 had uptalk use in interview speech but not in presentation speech. Women rated Speaker 13 more highly after listening to her interview speech than her presentation speech. Men had the opposite response. Male participants selected “Somewhat Disagree” more than women for competency statements in interview speech than in presentation speech. The comparison of results for men and women for Speaker 13’s interview is shown below.

**Figure 19. Responses for Speaker 13 by gender**



As mentioned, both men and women who had a preference for male leaders patterned similarly to women in their responses on the survey, as shown in Figure 20 below.

**Figure 20. Results for Speaker 13 by leadership preference**



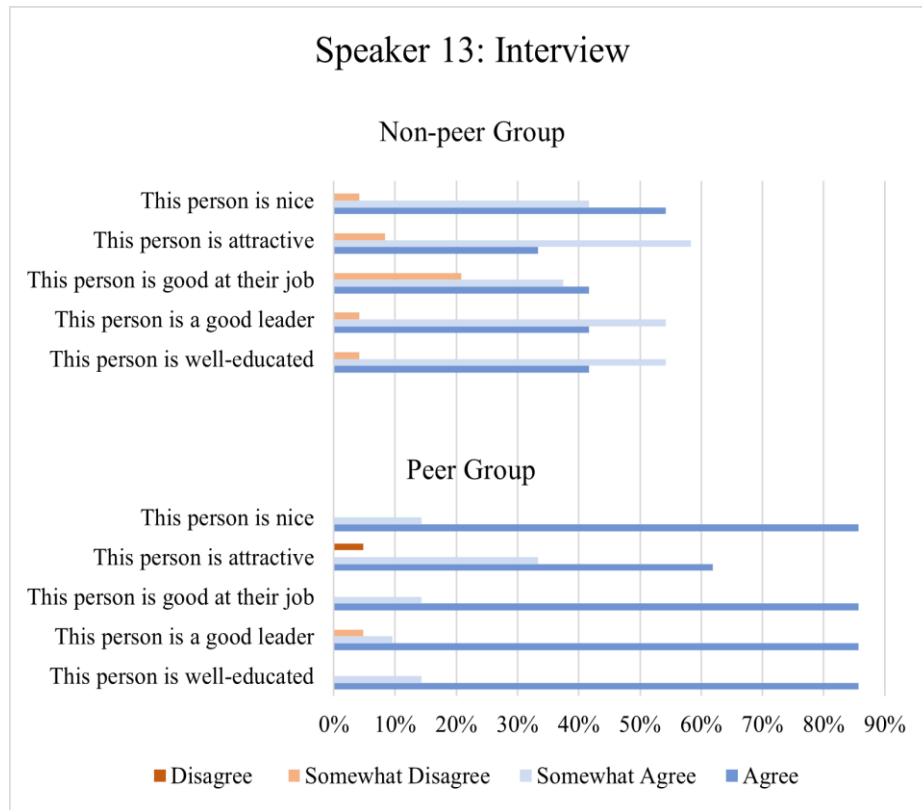
Participants who did not have a preference for male or female leaders tended to answer “Disagree” more for statements of competency in casual speech styles for all three speakers than participants who preferred male leaders.

### 3.5.6 Peer group perception

Women under the age of 35, the same demographic group as the speakers, patterned similarly to the other participants except in two cases.

First, young women seem to be less negatively affected by Speaker 13’s use of uptalk in interview speech. Speaker 13 received much higher responses of “Agree” for almost all statements from the peer group than from the participants who were not part of the peer group, as shown in the figure below. The only exception was for the statement “This person is attractive” which was only slightly higher.

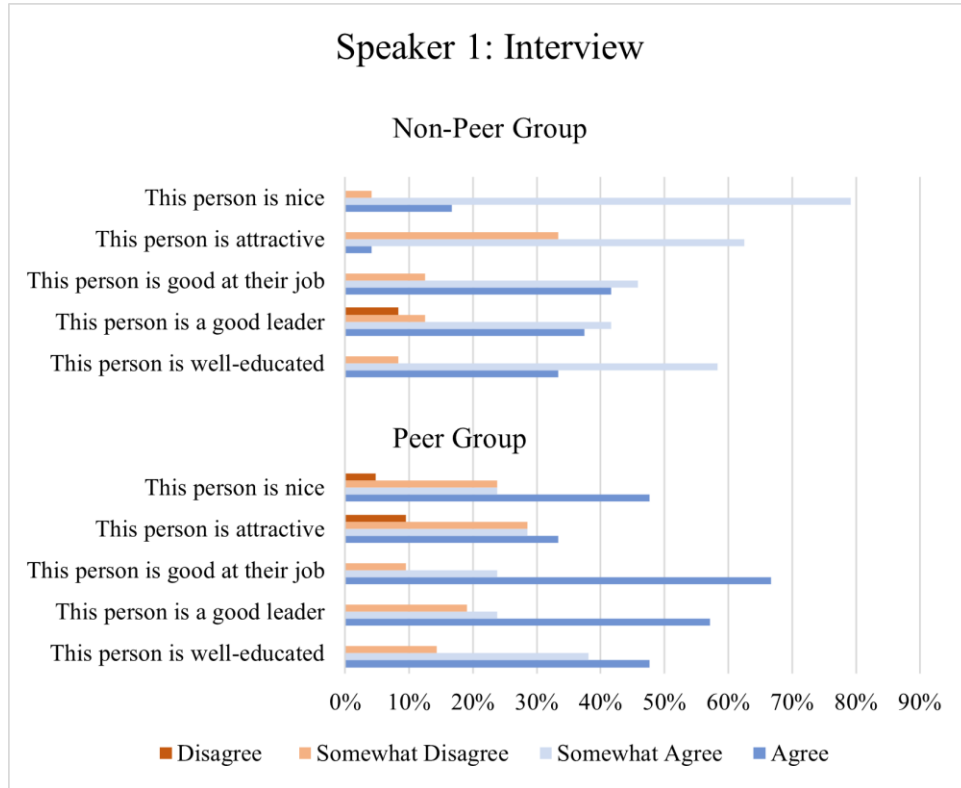
**Figure 21. Survey responses for Speaker 13: Interview by peer group**



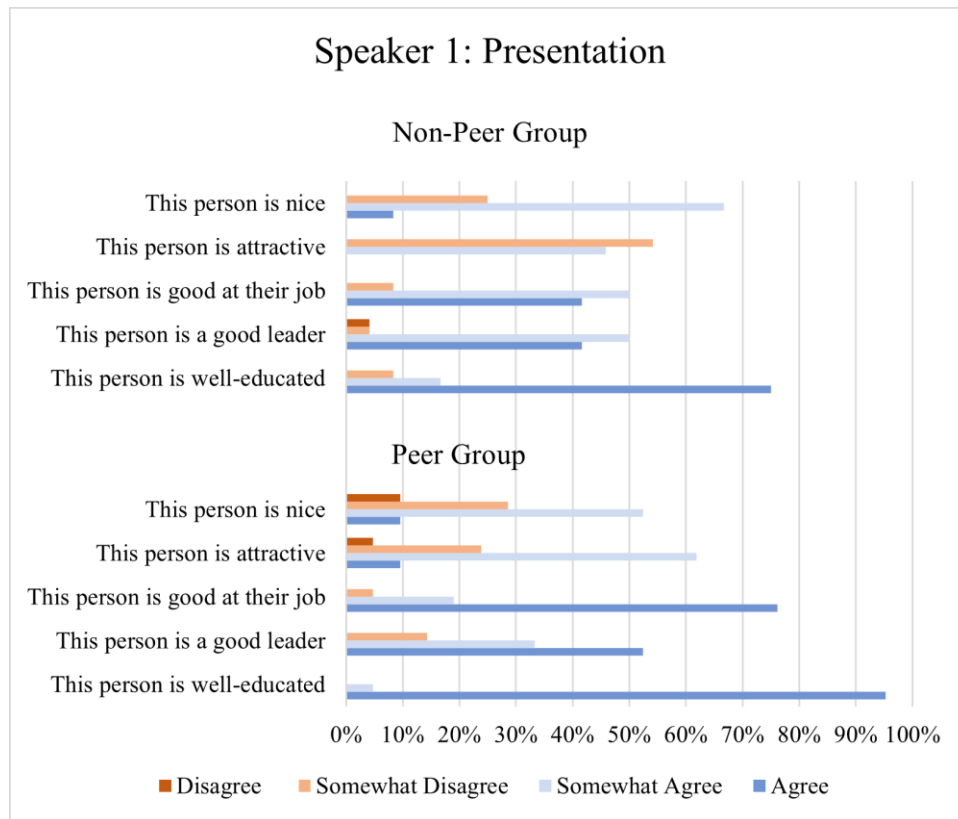
In Figure 19 in Section 3.5.5, the results from Speaker 13’s interview speech sample are divided by gender. The peer groups responses in Figure 21 above are even more positive than women’s responses as a whole which in turn were more positive than responses from men. These results indicate that young women view uptalk less negatively than other people. They do not seem to believe that uptalk use makes someone seem less competent or authoritative.

The second difference in peer group responses was the result for Speaker 1. Speaker 1 was rated higher for leadership in the interview and presentation styles by the peer group than by the non-peer group, as shown in the figures below.

**Figure 22. Responses for Speaker 1: Interview by peer group**



**Figure 23. Responses for Speaker 1: Presentation by peer group**



In the interview speech, the peer group had more variation in responses for statements “This person is nice” and “This person is attractive” than the non-peer group. This variation seems to indicate that young women are less affected by pitch as an indicator of a young woman’s attractiveness.



## CHAPTER 4 DISCUSSION

### 4.1 Summary of findings

In English, women's and men's speech differ in many ways, including lexical choice, intonation, and the use of tag-questions. These differences can be motivated by the desire to perform gender, behaving in a particular way to emphasize one's social gender. Often, gender performance involves using stereotypes to meet the audience's expectation of what someone of that gender does, whether or not those stereotypes are things a person of that gender would naturally do.

In this thesis, I have described and analyzed the understanding and use of three aspects of speech by young female American English speakers in different formality contexts. I specifically analyzed changes in pitch, uptalk use, and vocal fry across three speech styles: a presentation, a casual narrative, and an interview. The different contexts were simulated through prompts designed to elicit different types of speech and through the presence of an interlocutor of higher status than the speakers. All the conversations happened over Zoom, a videoconferencing platform.

A survey was done to measure the effects of pitch, uptalk, and vocal fry use on the listener's perception of the speaker. Three speakers were chosen who had noticeable changes in one of the speech characteristics in question. Each participant in the survey listened to three recordings from each speaker and rated them using a Likert scale to respond to different statements. The statements covered competency, authority, and attractiveness. In the following sections, I will discuss my findings regarding the production and perception of pitch, uptalk, and vocal fry.

#### 4.1.1 *Pitch*

Young women changed their pitch in different speech styles. In this study, presentation speech had the highest average pitch, casual speech was in the middle, and

the interview had the lowest average pitch. The effect of style on pitch was statistically significant for average pitch, minimum pitch, maximum pitch, and pitch range.

Raised pitch in the presentation may be motivated by the speaker's desire to project their voice and be understood. There is a correlation between raised pitch and amplitude for women (Kennard 2006).

Lowered pitch in the interview may be motivated by accommodation. The women may be attempting to mimic the speaking pitch of the male interviewer. Lowered pitch could also be motivated by a desire to seem authoritative and competent. Both men and women prefer voices with lower pitch when considering leadership capacity and qualities of strength, competency, and honesty (Klofstad et al. 2012). Lowered pitch may also be motivated by active non-performance of gender. Since there are more men than women in high-profile professional leadership positions (Catalyst 2020), young women may choose to emulate male speech in professional environments such as an interview setting.

This thesis is not exhaustive. It could be that there are other motivations behind the lowered speaking pitch in the interview style. Shevchenko found that upper class men and women used lower pitch than people from middle or lower class (1999). It could be that women are imitating upper class speech patterns and not mimicking male speech when they lower their pitch in the interview. I did not collect any information on class, so this motivation is outside the scope of this study and an area for further research.

#### *4.1.2 Uptalk*

Young women used significantly more uptalk when talking to a male interviewer than in casual conversation or a presentation. The presentation recordings contained significantly less uptalk than the casual speech and interviews. Most uptalk rises were less than 1 second in duration and had a 40% or more increase in pitch from the

beginning of the rise to the end. Often, the speaker used vocal fry immediately preceding an uptalk rise which increased the perception of the steepness of the rise.

Men are more likely than women to interpret uptalk as detracting from the speaker's competency or leadership skills. For men, the age of the participants did not change the overall results of the perception survey for uptalk. However, women over the age of 35 had slightly more negative perceptions of uptalk than women under the age of 35.

#### *4.1.3 Fry*

There was significantly more fry in the presentation than in the other speech styles. Even speakers who did not use fry very often in casual or interview speech had fry on most phrases in the presentation recording. The amount of fry used was not significantly affected by speech style, but different speakers used different amounts of fry. Presentation speech had more phrase-final fry than the other speech styles and less phrase-initial fry. Casual speech had less phrase-medial fry.

The difference in fry location is probably because the speakers read a paragraph for the presentation speech, whereas the other speech styles were free conversation. The presentation speech had strong declarative statements and few sentences that could begin with glottal stops, where most phrase-initial fry occurs.

The use of fry did not seem to affect listener's perceptions of the speaker's competency or attractiveness. However, higher amounts of vocal fry may negatively impact listener's judgments on how nice a person seems.

## **4.2 Limitations of this research**

### *4.2.1 Number of speakers and participants*

The relatively small number of speakers limits this study. Data from only 16 speakers was used for the analysis. From such a small sample, it is impossible to generalize about the whole population of young women speaking American English. However, related studies have used smaller numbers of participants. Lewis (2002) had 12 female subjects, and Kennard (2006) only had four, so this study exceeds the sample size of previous research.

There were only 45 participants in the perception survey total and only 13 men. This relatively small sample is difficult to draw conclusions from, but the results were fairly consistent for all men who participated in the survey. Expanding the survey is an area for further research.

### *4.2.2 Race*

One speaker identified as Latinx/Hispanic and two did not disclose their race. The other 13 speakers identified as White. Therefore, no conclusions can be made about the effect of the speaker's race on any of the variables discussed here.

### *4.2.3 Videoconferencing*

Because all the speakers' data was gathered remotely, I had little control over the quality of recordings. All the speakers joined using different devices in a variety of conditions. Sometimes they were interrupted by internet connectivity issues or by other people in their environments. The differences in quality between the recordings were not usually enough to make a large impact on the data analysis, but two speakers had to be removed from the study after their data had already been collected because their recordings were not high enough quality.

As mentioned in Section 1.6, videoconferencing does not completely approximate in-person interaction. Using videoconferencing to collect data is both a limitation and a benefit of this study. It is a limitation because any comparisons between this research and other studies that took place in-person must be done cautiously with careful attention to the differences between face-to-face and computer-mediated conversation. It is beneficial because it adds to the scant body of research using videoconferencing, which is becoming more prevalent.

#### *4.2.4 Naturalness and uniformity as competing values*

Some studies on speech prioritize isolating the particular characteristic in question. For example, one study on pitch used only vowels in order to control for all other features (Collins & Missing 2003). While using only vowels does eliminate noise in the study, it is not natural speech and so perceptions may not accurately reflect perceptions of pitch in natural speech. I chose to prioritize naturalness in my data collection.

This introduced some noise into the study. Speakers' accents, use of idioms and some filler words, story-telling ability, and interview skills were not completely controlled for.

Only the presentation style used read speech. For the other two speech styles, I attempted to control the content by having similar topics and questions, but each speaker had their own responses. For the casual speech sample, the speakers were prompted to tell a story of a first date or childhood memory. One drawback to this method of soliciting data is that there are different topics so the content and style is not as uniform as it could be if only one topic was allowed. Some of the speakers who chose to tell about a childhood memory acted out dialogue in their stories and used a higher pitch for those phrases to represent themselves as children. However, the quoted speech was brief and did not dramatically increase the overall average pitch of the samples.

### 4.3 Range of pitch

In Section 3.2, I noted that casual speech had much larger pitch variation than the other two speech styles. Presentation speech had the least amount of variation.

In Lewis' study on the effect of interlocutors on pitch, she found that pitch range was significantly larger when female subjects were speaking with a female interlocutor regardless of whether the interlocutor was a peer or had a higher status than the subject (2002). She attributes this difference primarily to the high value women place on solidarity and camaraderie (Lewis 2002).

If women are motivated by solidarity, they may accommodate more to the speaking pitch of the interlocutor. This accommodation may be part of the reason pitch variation was higher in casual speech in my data than in the interview. In the interview, the speakers may be attempting to accommodate to the male interviewer through mimicry, lowering their average pitch and thus diminishing their range. While accommodation may account for the pitch variation in the interview speech, this does not satisfactorily explain why casual speech had a much larger range.

Even though another speaker was present during the recording, the speaker was instructed to tell a story in the casual speech sample, not to have a conversation. The other person present did not speak very much, so it seems unlikely that the speaker would be accommodating to them. It seems more likely that the range of pitch in the casual speech was motivated by dramatic storytelling or the comfort level of the speakers and the range of pitch in the interview speech was limited by accommodation.

Croes et al. found that video-mediated communication had less pitch variation than face-to-face communication (2019). Because Lewis did not report her results for range of pitch in Hz and used speech samples that were 40-60 seconds long (2002: 59), not the measurements of each phrase, it is not possible to draw a direct comparison between her research and my findings. However, she did report means and standard

deviation of pitch. With female interlocutors, the mean was 192 Hz with a standard deviation of 23, and with male interlocutors, the mean was 188 Hz with a standard deviation of 21 (Lewis 2002: 61). The standard deviation of pitch in my data was much larger for both interview (SD = 32) and casual speech (SD = 43) though it was similar in presentation speech (SD = 24). However, as I mentioned above, this does not discount the findings of Croes et al since my data is not a direct comparison to Lewis'. The changes in the range of pitch across speech styles and modes of communication with interlocutors of different genders is an area of further research.

#### **4.4 Attitudes about uptalk**

Many people have strong opinions about uptalk, comparing it to infectious disease or blaming it for why young women are not taken seriously (Davis 2010; DiResta 2016). Even in conducting this survey, I received unsolicited feedback from a participant with good intentions who summed up popular beliefs about uptalk.

“On delivery, one thing that can undermine your authority (and a personal pet peeve of mine) is the raising of the pitch of the voice at the end of non-question sentences — a fairly recent cultural phenomena (within [the] last 20 years). It's as if the speaker is looking for validation or encouragement and lacks confidence in what she/he is saying. Speak with authority, like you are deeply rooted and standing firm.” (Anonymous, personal communication)

Here this participant reiterates the false belief that uptalk is a recent development. This is certainly not the case (Pike 1945: 59). The participant also shares the opinion that uptalk use makes a speaker seem unsure of themselves. Male participants seemed to interpret uptalk differently than female participants; however, the effect is small. Speaker 13 was used in the perception survey to measure attitudes towards uptalk. The sample selected from the interview speech had uptalk on 35% of the phrases. Her other

speech samples had no uptalk. Men rated her lower in competency than women did after the recording with uptalk. Yet, more people agreed that she was a good leader and good at her job after listening to the interview speech than the presentation or casual speech.

Perhaps uptalk bothers people in the abstract when they pay attention to it, but does not actually cost young women, or other people using uptalk, their perceived authority or competency. People who discourage young women from using uptalk assign it only one or two particular meanings, usually uncertainty or excessive deference. When potential uses for uptalk from Ching (1982) and Warren (2016) are combined, there are, at least, a dozen different motivations for uptalk, such as topicalization, emphasis, checking for understanding, deference, politeness, floor-holding, signaling surprise, continuation, connecting with or seeking a response from the listener, verifying agreement, and qualifying the truthfulness or certainty of a statement. Uptalk may indicate uncertainty or insecurity, but those are certainly not the only, or even primary, uses of uptalk. Assuming that uptalk can only communicate uncertainty is a misconception of its various uses and a misunderstanding of the intentions of the speaker using it. Women may use uptalk in ways distinct from men (Linneman 2013) to perform gender, but as men and women speak with one another, men need to understand the range of meaning in uptalk to avoid groundless prejudice about the competency of young women based on valid intonation use.

#### **4.5 Further research**

Two findings in my study are areas for further research. The first is the much larger range of pitch in casual speech over the two formal speech styles. This could be motivated by the speech style, the gender or status of the interlocutor, or some



unknown factor. The second finding that has not been previously discussed in detail is the use of uptalk as a discourse feature in story-telling.

Two limitations of my study need to be expanded. There seems to be some interaction between the education level of the speaker and the use of vocal fry, but the number of speakers was not sufficient to draw definite conclusions. Second, since I did not have a racially diverse group of speakers and did not collect information on race from the survey participants, the impact of race on both production and perception is an area for further study.

#### **4.6 Conclusion**

The goals of this study were to examine the changes young women make in their speech when they want to appear more competent and professional and how these speech adjustments affect the way young women are perceived.

The women in this study had lowest average pitch in interview speech and the widest range of pitch in casual speech. The findings on pitch were not conclusive but seemed to confirm Collins and Missing's conclusion that women with higher pitched speaking voices are judged as more attractive (Collins & Missing 2003).

The speakers used more uptalk in interview speech than in presentation or casual speech. If they used uptalk in casual speech, it was most frequently during scene setting or the dramatic climax of the narrative. It was interesting to note the contrast between the male and female listeners perception of uptalk. The male listeners felt the use of uptalk demonstrated a lack of competence, whereas the women's perception was neutral.

The women in this study had the most vocal fry in presentation speech. Using vocal fry with either male or female listeners may negatively impact their perceptions of the niceness of the speaker.

Women are faced with many conflicting motivations in their use of pitch, uptalk, and fry. In order to sound authoritative and competent, they may make changes in their speech that sacrifice attractiveness and niceness. In attempting to sound polite or engage the listener, they may come across as insecure. My research adds to the conversation about these complex issues, so that the context and motivations behind the use of pitch, uptalk, and vocal fry by young women can be understood. It is important that users of American English understand these speech characteristics in order to understand each other better.

## **APPENDICES**

## APPENDIX A

These tables show the results for linear mixed regression models for pitch. Three within-speaker ANOVAs were performed in four separate tests. The dependent variable in each test is pitch (**f0min**, **f0max**, **f0range**, or **f0mean**) and the independent variables are speech style, age, and education of the speakers.

**Appendix Table 1. Linear regression results for mean F0**

`f0test1 <- lmer(f0mean ~ (Style + Age + Education) + (1|Speaker))`

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
<b>Style</b>	108217	54109	2	1653.6	59.0666	<2e-16 ***
<b>Age</b>	394	197	2	10.0	0.2150	0.8101
<b>Education</b>	1349	450	3	10.0	0.4908	0.6965

**Appendix Table 2. Linear regression results for range of F0 (max-min)**

`f0test2 <- lmer(f0range ~ (Style + Age + Education) + (1|Speaker))`

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
<b>Style</b>	372440	186220	2	1655.72	40.2506	<2e-16 ***
<b>Age</b>	12889	6444	2	10.03	1.3929	0.2925
<b>Education</b>	10905	3635	3	10.04	0.7857	0.5287

**Appendix Table 3. Linear regression results for min F0**

f0test3 <- lmer(f0min ~ (Style + Age + Education) + (1|Speaker))

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
<b>Style</b>	15426.1	7713.0	2	1653.73	5.9051	0.002783 **
<b>Age</b>	2445.6	1222.8	2	9.97	0.9362	0.424042
<b>Education</b>	6174.2	2058.1	3	9.97	1.5757	0.256221

**Appendix Table 4. Linear regression results for max F0**

f0test4 <- lmer(f0max ~ (Style + Age + Education) + (1|Speaker))

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
<b>Style</b>	439100	219550	2	1654.24	55.7130	<2e-16 ***
<b>Age</b>	1773	886	2	10.04	0.2249	0.8025
<b>Education</b>	1610	537	3	10.04	0.1362	0.9362

## APPENDIX B

Below is a table of estimates with upper and lower limits of 95% confidence intervals for logistic regression models for uptalk. Intervals that do not contain the null value, 0, are marked with an asterisk.

**Appendix Table 5. Logistic regression table of estimates for uptalk**

<b>Variable</b>	<b>Estimate</b>	<b>Lower Limit</b>	<b>Upper Limit</b>
Presentation	-1.5983095	-2.4401596	-0.7564594*
Casual	0.6805187	-0.20301172	1.564049
Interview	1.5983093	0.75645261	2.440166*
Age18-24	-0.1815821	-0.8835138	0.52034964
Age25-29	0.1815821	-0.52034932	0.8835135
Age30-35	-0.8629521	-2.03013205	0.3042279
Some college or university	-0.5484989	-1.1764978	0.0795001
Completed college or university	0.5484984	-0.07950194	1.1764988
Some postgraduate study	0.4443196	-0.37957362	1.2682127
Completed postgraduate degree	0.4405644	-1.06086122	1.9419899

## APPENDIX C

Below are tables of estimates with upper and lower limits of 95% confidence intervals for logistic regression models for fry. Intervals that do not contain the null value, 0, are marked with an asterisk.

**Appendix Table 6. Logistic regression results for fry**

<b>Variable</b>	<b>Estimate</b>	<b>Lower limit</b>	<b>Upper limit</b>
Presentation	0.816099	0.433181	1.199017*
Casual	-0.90932	-1.30366	-0.51498*
Interview	-0.8161	-1.19902	-0.43318*
Age18-24	0.43984	-0.77354	1.653219
Age25-29	0.664045	-0.37927	1.707364
Age30-35	-0.66404	-1.70735	0.379261
Some college or university	1.09821	-0.47996	2.676375
Completed college or university	0.383683	-1.28595	2.053318
Some postgraduate study	1.944212	0.374054	3.514369*
Completed postgraduate degree	-0.38369	-2.05329	1.285918

**Appendix Table 7. Logistic regression results for phrase initial fry**

<b>Variable</b>	<b>Estimate</b>	<b>Lower limit</b>	<b>Upper limit</b>
Presentation	-0.77862	-1.43367	-0.12358*
Casual	0.694653	0.02492	1.364386*
Interview	0.778623	0.123579	1.433667*
Age18-24	-0.51757	-1.60651	0.571362
Age25-29	0.10869	-0.66616	0.88354
Age30-35	0.517573	-0.57135	1.6065
Some college or university	2.792508	0.880603	4.704413*
Completed college or university	-0.71615	-1.41457	-0.01773*
Some postgraduate study	2.82419444	0.92956103	4.7188278*
Completed postgraduate degree	-2.79251	-4.70436	-0.88065*

**Appendix Table 8. Logistic regression results for phrase-medial fry**

<b>Variable</b>	<b>Estimate</b>	<b>Lower limit</b>	<b>Upper limit</b>
Presentation	-0.08525	-0.42854	0.258046
Casual	-0.42208	-0.64981	-0.19435
Interview	0.085246	-0.25805	0.428539
Age18-24	0.521673	-0.35295	1.396298
Age25-29	0.544536	-0.20233	1.291399
Age30-35	-0.54454	-1.29141	0.202337
Some college or university	-0.51962	-1.19609	0.156855
Completed college or university	-0.4761	-1.02134	0.069127
Some postgraduate study	0.519614	-0.15685	1.196083
Completed postgraduate degree	-0.66907	-1.83056	0.492407

**Appendix Table 9. Logistic regression results for phrase-final fry**

<b>Variable</b>	<b>Estimate</b>	<b>Lower limit</b>	<b>Upper limit</b>
Presentation	1.083498	0.730806	1.436191
Interview	-0.15256	-0.38844	0.083314
Casual	0.152562	-0.08331	0.388438
Age18-24	-0.51374	-1.11947	0.091986
Age25-29	0.63638	-0.10069	1.373448
Age30-35	-0.63638	-1.37345	0.100687
Some college or university	0.528615	-0.61137	1.668596
Completed college or university	0.057242	-1.14509	1.259571
Some postgraduate study	0.788771	-0.34286	1.920402
Completed postgraduate degree	-0.05725	-1.25958	1.145084



## APPENDIX D

### Copy of survey used to recruit speakers.

Q1 If you are a woman ages 18-35 living in the U.S., we need your help! Through our research on speech patterns, we want to understand and promote societal equity. If you'd like to participate, we ask that you and a friend sign up for a 30-minute Zoom call and answer a few questions about yourself in a short questionnaire.

1. You and your friend decide on a time when you are available to meet during the weeks of October 12th and 19th - Monday, Wednesday, or Thursday between 4pm and 8pm ET. You can sign up for a specific time in the questionnaire.
2. Answer the questionnaire. This will take about 5 minutes.
3. After you complete the questionnaire, you'll get an email with more information and a Zoom link. Save the link! Download Zoom if you haven't already.
4. When it's time for the call, open the Zoom link. You and your friend will get a chance to ask any questions you may have. Then you'll chat with each other, read a paragraph, and talk to an interviewer. The Zoom call will be recorded, and parts of the audio (not video) may be shared with other participants, but your name will not be shared.
5. You're done! Thank you for promoting linguistic knowledge in the world!

Ready to get started?

#### Q2 [Consent Form](#)

I have read, understood, and if desired, printed a copy of, the consent form and desire of my own free will to participate in this study.

Yes

No

Q3 How old are you?

- 17 or under
- 18-24
- 25-29
- 30-35
- 36 +

Q4 Do you currently have any difficulties hearing, processing, or producing speech or language?

- Yes
- No

Q5 What is your gender assigned at birth?

- Male
- Female
- Prefer not to say

Q6 What is your gender identity?

- Male
- Female
- Transgender
- Other/Prefer not to say

Q7 Where do you live?

- USA
- I do not live in the USA

Q8 What state do you live in?

---

Q9 Have you lived in another state for more than 5 years?

- Yes
- No

Q10 Where else have you lived?

---

Q11 What is your race/ethnicity?

- Other/Prefer not to say
- White
- Black or African American
- Asian
- American Indian and Alaska Native
- Native Hawaiian and Other Pacific Islander
- Latinx/Hispanic

Q12 Do you live in a...

- Rural area
- Town/village
- City
- Urban area

Q13 How long have you lived in your current state?

- 0-4 years
- 5-9 years
- 10-14 years
- 15-19 years
- 20+ years

Q14 Have you previously lived in another country?

- No
- Yes

Q15 Where else have you lived?

---

Q16 What is your current level of education?

- No high school/secondary school
- Some high school/secondary school
- Completed high school secondary school
- Some college/university
- Completed college/university
- Some postgraduate school
- Completed postgraduate school

Q17 What is your profession?

---

Q18 Do you supervise people at your work?

- Yes
- No

Q19 Is English your first language?

- Yes
- No

Q20 When did you begin learning English?

---

Q21 Do you speak languages other than English?

Yes

No

Q22 What other languages do you speak?

---

## REFERENCES

- Abe, Isamu. 1955. Intonational patterns of English and Japanese. *WORD* 11(3). 386–398. (doi:10.1080/00437956.1955.11659567)
- Abitbol, Jean & Abitbol, Patrick & Abitbol, Béatrice. 1999. Sex hormones and the female voice. *Journal of Voice* 13(3). 424–446. (doi:10.1016/S0892-1997(99)80048-4)
- Apple, William & Streeter, Lynn A. & Krauss, Robert M. 1979. Effects of pitch and speech rate on personal attributions. *Journal of Personality and Social Psychology*. US: American Psychological Association 37(5). 715–727. (doi:10.1037/0022-3514.37.5.715)
- Armstrong, Meghan E & Piccinini, Page E & Ritchart, Amanda. 2015. The phonetics and distribution of non-question rises in two varieties of American English. *The Scottish Consortium for ICPhS 2015*, 5. Glasgow, UK: the University of Glasgow.
- Bailenson, Jeremy N. 2021. Nonverbal overload: A theoretical argument for the causes of Zoom fatigue. *Technology, Mind, and Behavior*. PubPub 2(1). (doi:10.1037/tmb0000030) (<https://tmb.apaopen.org/pub/nonverbal-overload/release/1>) (Accessed April 16, 2021.)
- Boersma, Paul & Weenik, David. 2019. *Praat: Doing phonetics by computer [Computer program]*. (<http://www.praat.org/>)
- Britannica, The Editors of Encyclopaedia. 2013. Castrato. *Encyclopedia Britannica*. Encyclopaedia Britannica. (<https://www.britannica.com/art/castrato>) (Accessed March 31, 2021.)
- Bryant, Gregory A. & Haselton, Martie G. 2009. Vocal cues of ovulation in human females. *Biology Letters* 5(1). 12–15. (doi:10.1098/rsbl.2008.0507)
- Catalyst. 2020. *Historical list of women CEOs of the Fortune Lists: 1972-2020*. New York.
- Ching, Marvin K. L. 1982. The question intonation in assertions. *American Speech*. [Duke University Press, American Dialect Society] 57(2). 95–107. (doi:10.2307/454443)
- Coates, Jennifer. 2013. Language, gender and career. In Mills, Sara (ed.), *Language and gender: Interdisciplinary perspectives*. Routledge.
- Collins, Sarah A. & Missing, Caroline. 2003. Vocal and visual attractiveness are related in women. *Animal Behaviour* 65(5). 997–1004. (doi:10.1006/anbe.2003.2123)
- Croes, Emmelyn A. J. & Antheunis, Marjolijn L. & Schouten, Alexander P. & Kraemer, Emiel J. 2019. Social attraction in video-mediated communication: The role of

- nonverbal affiliative behavior. *Journal of Social and Personal Relationships*. SAGE Publications Ltd 36(4). 1210–1232. (doi:10.1177/0265407518757382)
- Dabbs, James M. & Carr, Timothy S. & Frady, Robert L. & Riad, Jasmin K. 1995. Testosterone, crime, and misbehavior among 692 male prison inmates. *Personality and Individual Differences* 18(5). 627–633. (doi:[https://doi.org/10.1016/0191-8869\(94\)00177-T](https://doi.org/10.1016/0191-8869(94)00177-T))
- Dabbs Jr., James M. & Mallinger, Alison. 1999. High testosterone levels predict low voice pitch among men. *Personality and Individual Differences*. Netherlands: Elsevier Science 27(4). 801–804. (doi:10.1016/S0191-8869(98)00272-4)
- Davis, Hank. 2010. The uptalk epidemic. *Psychology Today*. (<http://www.psychologytoday.com/blog/caveman-logic/201010/the-uptalk-epidemic>) (Accessed March 6, 2021.)
- de Klerk, Vivian. 1991. Expletives: Men only? *Communication Monographs*. Routledge 58(2). 156–169. (doi:10.1080/03637759109376220)
- Di Gioacchino, Martina & Crook Jessop, Lorena. 2011. Uptalk: Towards a quantitative analysis. *Toronto Working Papers in Linguistics* 33.
- DiResta, Diane. 2016. Does uptalk make you upchuck? *DiResta Communication Inc.* (Blog.) (<https://www.diresta.com/knockoutpresentationsblog/resources/articles/does-uptalk-make-you-upchuck>) (Accessed March 4, 2021.)
- Dohle, G. R. & Smit, M. & Weber, R. F. A. 2003. Androgens and male fertility. *World Journal of Urology* 21(5). 341–345. (doi:10.1007/s00345-003-0365-9)
- Eckert, Penelope & McConnell-Ginet, Sally. 2003. *Language and gender*. Cambridge ; New York: Cambridge University Press.
- Fishman, Pamela. 1983. Interaction: The work women do. *Language, gender, and society*, 89–101. Rowley, Mass: Newbury House.
- Fletcher, Janet & Loakes, Deborah. 2006. Patterns of rising and falling in Australian English. 42–47. University of Auckland, New Zealand: Australian Speech Science & Technology Association Inc.
- Gonzales, Amy & Hancock, Jeffrey. 2011. Mirror, mirror on my Facebook wall: Effects of exposure to Facebook on self-esteem. *Cyberpsychology, behavior and social networking* 14. 79–83. (doi:10.1089/cyber.2009.0411)
- Gries, Stefan Th. 2017. Quantitative methods in linguistics. 14.
- Guy, Gregory & Horvath, Barbara & Vonwiller, Julia & Daisley, Elaine & Rogers, Inge. 1986. An intonational change in progress in Australian English. *Language in Society*. Cambridge University Press 15(1). 23–51.



- Hollien, H. & Wendahl, R. W. 1968. Perceptual study of vocal fry. *The Journal of the Acoustical Society of America* 43(3). 506–509. (doi:10.1121/1.1910858)
- Hollien, Harry & Moore, Paul & Wendahl, Ronald W. & Michel, John F. 1966. On the Nature of Vocal Fry. *Journal of Speech and Hearing Research*. American Speech-Language-Hearing Association 9(2). 245–247. (doi:10.1044/jshr.0902.245)
- Holmes, Janet. 1990. Hedges and boosters in women's and men's speech. *Language & Communication* 10(3). 185–205. (doi:10.1016/0271-5309(90)90002-S)
- Ingram, Rick E & Cruet, Debra & Johnson, Brenda R & Wisnicki, Kathleen S. 1988. Self-focused attention, gender, gender role, and vulnerability to negative affect. *Journal of Personality and Social Psychology* 55(6). 967–978.
- Iqbal, Mansoor. 2020. Zoom revenue and usage statistics (2020). *Business of Apps*. (<https://www.businessofapps.com/data/zoom-statistics/>) (Accessed April 17, 2021.)
- Keating, Patricia & Garellek, Marc. 2015. Acoustic analysis of creaky voice. Portland, OR. (Poster presented at the LSA Annual Meeting, Portland, OR.)
- Keating, Patricia & Garellek, Marc & Kreiman, Jody. 2015. Acoustic properties of different kinds of creaky voice. *Proceedings of the 18th International Congress of Phonetic Sciences*, 5. Glasgow, UK: the University of Glasgow.
- Kennard, Catherine Hicks. 2006. *Gender and command: A sociophonetic analysis of female and male drill instructors in the United States Marine Corps*. University of Arizona. (Dissertation.)
- Klofstad, Casey A. & Anderson, Rindy C. & Peters, Susan. 2012. Sounds like a winner: voice pitch influences perception of leadership capacity in both men and women. *Proceedings of the Royal Society B: Biological Sciences*. Royal Society 279(1738). 2698–2704. (doi:10.1098/rspb.2012.0311)
- Klofstad, Casey A. & Nowicki, Stephen & Anderson, Rindy C. 2016. How voice pitch influences our choice of leaders. *American Scientist* 104(5). (<https://www.americanscientist.org/article/how-voice-pitch-influences-our-choice-of-leaders>) (Accessed March 31, 2021.)
- Kramer, Ernest. 1963. Judgment of personal characteristics and emotions from nonverbal properties of speech. *Psychological Bulletin*. American Psychological Association 60(4). 408–420. (doi:10.1037/h0044890)
- Kuang, Jianjing & Liberman, Mark. 2016. The effect of vocal fry on pitch perception. *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 5260–5264. Shanghai: IEEE. (doi:10.1109/ICASSP.2016.7472681) (<http://ieeexplore.ieee.org/document/7472681/>) (Accessed March 9, 2021.)

- Lakoff, Robin. 1973. Language and woman's place. *Language in Society* 2(1). 45–80.
- Lass, Norman J. & Brown, William S. 1978. Correlational study of speakers' heights, weights, body surface areas, and speaking fundamental frequencies. *The Journal of the Acoustical Society of America* 63(4). 1218–1220. (doi:10.1121/1.381808)
- Lewis, Julie Anne. 2002. *Social influences on female speakers' pitch*. Berkely, CA: University of California.
- Linneman, Thomas J. 2013. Gender in Jeopardy!: Intonation variation on a television game show. *Gender & Society*. SAGE Publications Inc 27(1). 82–105. (doi:10.1177/0891243212464905)
- Mills, Sara. 2003. *Gender and politeness*. Cambridge University Press.
- Montepare, Joann M. & Zebrowitz-McArthur, Leslie. 1987. Perceptions of adults with childlike voices in two cultures. *Journal of Experimental Social Psychology* 23(4). 331–349. (doi:10.1016/0022-1031(87)90045-X)
- Newall, Mallory & Chen, Emily. 2020. Coronavirus prompts increased use of video chat platforms for work, connection | Ipsos. *Ipsos*. (<https://www.ipsos.com/en-us/news-polls/coronavirus-prompts-increased-use-of-video-chat>) (Accessed September 18, 2020.)
- Nowaczyk, Ronald H. 1982. Sex-related differences in the color lexicon. *Language and Speech*. SAGE Publications Ltd 25(3). 257–265. (doi:10.1177/002383098202500304)
- Nycum, Reilly. 2018. In defense of Valley Girl English. *The Compass* 1(5). 23–29.
- Oppezzo, Marily & Schwartz, Daniel L. 2014. Give your ideas some legs: The positive effect of walking on creative thinking. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 40(4). 1142–1152. (doi:10.1037/a0036577)
- Pierce, David. 2020. Zoom conquered video chat — now it has even bigger plans. *Protocol*. (<https://www.protocol.com/zoom-videoconferencing-history-profit>) (Accessed April 17, 2021.)
- Pike, Kenneth L. 1945. *The intonation of American English*. Ann Arbor, MI: University of Michigan Press.
- Pipitone, R. & Gallup. 2008. Women's voice attractiveness varies across the menstrual cycle. *Evolution and Human Behavior - EVOL HUM BEHAV* 29. 268–274. (doi:10.1016/j.evolhumbehav.2008.02.001)
- R Core Team. 2020. *R: A language and environment for statistical computing*. Vienna, Austria. (<https://www.R-project.org/>.)
- Schoenenberg, Katrin & Raake, Alexander & Koeppe, Judith. 2014. Why are you so slow? – Misattribution of transmission delay to attributes of the conversation

partner at the far-end. *International Journal of Human-Computer Studies* 72(5). 477–487. (doi:10.1016/j.ijhcs.2014.02.004)

Schwab, Klaus & Crotti, Robert & Geiger, Thierry & Ratcheva, Vesselina & World Economic Forum. 2020. *Global gender gap 2020 insight report*. Geneva: World Economic Forum.

Shevchenko, T I. 1999. The sociocultural value of F0 variation in British and American English. *Proceedings of the 14th International Congress of Phonetic Sciences*, 1609–1612. San Francisco, CA: University of California Press.

Talbot, Mary. 2020. *Language and gender*. Third. Cambridge, UK: Polity Press.

Warren, Paul. 2016. *Uptalk: The phenomenon of rising intonation*. Cambridge University Press.

Yuan, Eric. 2012. *Zoom*. San Jose, CA: Zoom Video Communications, Inc. (<https://zoom.us/>)